2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

An Iterative Partition-Based Moving Horizon Estimator for Large-Scale Linear Systems

René Schneider*† , Holger Scheu*‡ and Wolfgang Marquardt*§
*AVT - Process Systems Engineering, RWTH Aachen University, 52056 Aachen, Germany
† Email: rene.schneider@avt.rwth-aachen.de
‡ Email: holger.scheu@avt.rwth-aachen.de
§ Email: wolfgang.marquardt@avt.rwth-aachen.de

***Abstract*— We transfer the ideas behind sensitivity-driven distributed model predictive control (c.f. Scheu and Marquardt, 2011) to the moving horizon state estimation problem and present a novel decentralized state estimation algorithm, namely, sensitivity-driven partition-based moving horizon estimation (S-PMHE). We discuss convergence and optimality of S-PMHE for the case of given positive-definite arrival cost weights. Finally, we demonstrate the method on a numerical example.**

I. Introduction

State estimation of linear systems is a classical problem in modern control theory. While traditional methods such as the Kalman filter [1] and the Luenberger observer [2] are straight forward to implement and do not pose any significant numerical demand, the more recently introduced optimization-based methods such as the moving horizon estimator (MHE) [3], [4] can be computationally more demanding, in particular for large-scale systems. Spatial decomposition can be considered to reduce the computational demand in moving horizon estimation, similar to distributed model-predictive control, which has gained broad attention in the last years [5]. In this paper, we develop a novel sensitivity-driven moving horizon estimation scheme for large-scale linear systems.

According to Farina et al. [6], distributed control systems can be classified according to their communication topology and the rate of information exchange, also called data delivery. In case of state estimation, classification also has to account for the specification of the part of the overall state vector which is estimated by each of the subsystems. In particular, concerning the topology of the communication network, Farina et al. [6] distinguish between *all-to-all communication* where all subsystems communicate with each other, and *neighbor-to-neighbor communication* where interaction between the subsystems can be described by a directed graph and each subsystem communicates only with its "neighbor" subsystems. Information can be exchanged either at a given transmission rate (*continuous data delivery*), or when a given event occurs (*event-driven data delivery*), or as a response to an explicit request from other subsystems (*observer-initiated data delivery* or *request-reply model*). Finally, any state of the overall system may be part of the state vector of only one or more than one subsystem. We refer to the latter as the case of overlapping states. *Distributed estimation* refers to the extreme case of overlapping states, where every subsystem estimates the whole state vector of the system, while *partition-based estimation* is used in all other cases [6]. State estimation methods have been proposed in the literature for all these different types of estimation schemes. We refer to Farina et al. [6] for an excellent review of the different methods.

The structured large-scale linear systems we are interested in in this work are motivated by chemical or energy process systems which consist of a large number of coupled subsystems reflecting the process units of a mass and/or energy conversion plant. The couplings between the subsystems refer to material, energy or information flow. Consequently, the large-scale model of such plants can be naturally decomposed into a set of interconnected subsystems. The decentralized state estimator considered in this work exploits this decomposition and assumes all-to-all communication, continuous data delivery and multiple information exchange between the estimators for the subsystems in every sampling interval. Further, in order to minimize the computational effort in each of the decentralized state estimators, the estimated states in the subsystems are non-overlapping. Thus, the state estimation algorithm we are going to propose in this work can be classified as an iterative, partition-based moving-horizon estimator with non-overlapping states. While many types of coordination strategies could be envisioned in case of partition-based filtering, a sensitivity-driven coordination algorithm [7] is adapted here for the first time in the context of state estimation, since this algorithm has already turned out to be very competitive among various algorithms in the context of distributed model-predictive control [8].

In order to put our research into the context of previous work, we briefly review those decentralized estimation schemes for large-scale systems which are most relevant for the types of process problems we are interested in and for the type of solution method we are going to present. These works include [9], [10], [11], [12], [13]. Hassan et al. [9] were among the first to present a decentralized version of the Kalman filter. In their two-level concept, the Kalman filter gain is computed for every subsystem in an iterative manner. The authors show that this approach exhibits more favorable numerical properties than a centralized Kalman filter applied to the whole system, because the filter calculations are performed on blocks of subsystems equations of lower dimension than the whole system. A reduction of the computational complexity of centralized Kalman filtering has also been the motivation behind the work of Hashemipour et al. [10] and Rao and Durrant-Whyte [11]. However, in contrast to [9], these classical works assume that each subsystem has full knowledge of the overall plant dynamics, i.e., the model of the complete system is available to all the estimators for the subsystems. This is also assumed in the sensor network application presented by Necoara et al. in [12]. These authors further presented a systematic framework for solving coupled estimation problems. Unfortunately, their framework does not apply to the general optimization problem presented here which is typically coupled in both the objective function and in the constraints, and in which the decision variables of the subsystems are distinct. Finally, the recent approach of Farina et al. [13] is probably the closest in spirit to the method proposed in this paper. These authors apply a moving horizon strategy to a partition-based state estimation problem. The optimization-based moving horizon framework is selected by these authors to systematically incorporate constraints at the expense of an increase in computational cost compared to the Kalman filter [3]. Three variants of a moving horizon partition-based estimation (PMHE) algorithm are presented and their convergence properties are studied assuming absence of measurement noise. One of their variants, called PMHE2, comes closest to the method proposed in this work, since it also assumes all-to-all communication. However, in constrast to PMHE2, the method proposed here does not compute the weights or covariances in a decentralized manner and does not necessarily require that every subsystem has access to the dynamic model of the complete system. Further, while all of the PMHE variants presented in [13] are non-iterative, the one proposed in this work relies on iterations in order to mimic the behavior of the centralized MHE and to derive the same optimal state estimate asymptotically.

The remainder of this paper is structured as follows: In the next section, the investigated linear system of interest is introduced and the centralized MHE problem is formulated in its original and in a decomposed form. The novel decentralized estimation algorithm is developed in Section III. Convergence to the optimal solution of the central estimation problem is studied in Section IV. A case study is presented in Section V, and the paper concludes with a summary and an outline of further research in Section VI.

II. Problem formulation

We consider the discrete-time linear time-invariant system

$$x^\diamond(k+1) = Ax^\diamond(k) + w^\diamond(k), \quad x^\diamond(0) = x_0^\diamond, \tag{1a}$$

$$y^\diamond(k) = Cx^\diamond(k) + v^\diamond(k). \tag{1b}$$

$x^\diamond(k) \in \mathbb{R}^n$ and $y^\diamond(k) \in \mathbb{R}^p$ are the state and measurement vectors of the physical system at time t_k, where k refers to the time index, i.e. $t_k = t_0 + k\Delta t$, and Δt is the sampling time. Analogously, $w^\diamond(k) \in \mathbb{R}^n$ and $v^\diamond(k) \in \mathbb{R}^p$ are the zero-mean Gaussian process and measurement noise with covariance matrices $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{p\times p}$, respectively. Finally, $x_0^\diamond$ refers to the initial condition and $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{p\times n}$ are the system and output matrices.

In order to estimate the state $x^\diamond$ of system (1) from the available measurements $y^\diamond$, we formulate the following MHE problem, at the current time $t_{k'}$ [3]:

$$\min_{\Delta x(k^0),x,w,v} \frac{1}{2}\Bigg(\left\|\Delta x(k^0)\right\|^2_{\tilde{P}} \tag{2a}$$

$$+ \sum_{k=k^0}^{k'-1} \|w(k)\|^2_{\tilde{Q}} + \sum_{k=k^0}^{k'} \|v(k)\|^2_{\tilde{R}} \Bigg)$$

$$\text{s.t. } 0 = -\Delta x(k^0) + x(k^0) - \bar{x}(k^0), \tag{2b}$$

$$0 = -w(k) + x(k+1) - Ax(k), \quad k = k^0, \dots, k'-1, \tag{2c}$$

$$0 = -v(k) + y^\diamond(k) - Cx(k), \quad k = k^0, \dots, k', \tag{2d}$$

where we consider a moving estimation window of maximum length $K\Delta t$, i.e. processing up to $K+1$ measurement samples $y^\diamond(k)$ from system (1), $k \in \{k^0,\dots,k'\}$, $k^0 = k' - K$. In the beginning, i.e. when $k' < K$, the problem is solved on a growing horizon, processing only those measurements up to time $t_{k'}$, i.e. $k^0 = 0$ in (2). In the problem above, we have further introduced the a priori estimate of the state at the beginning of the horizon, $\bar{x}(k^0)$. To be more compact, we have also defined $x = \langle x(k^0),\dots,x(k')\rangle$, $w = \langle w(k^0),\dots,w(k'-1)\rangle$ and $v = \langle v(k^0),\dots,v(k')\rangle$, where the bracket notation $\langle a_1,\dots,a_m\rangle$ indicates $[a_1^T \dots a_m^T]^T$. Finally, $\tilde{P}$, $\tilde{Q}$ and $\tilde{R}$ are weighting matrices, assumed to be positive definite. In that case, problem (2) is a strictly convex quadratic program (QP) since the only variable that is not directly entering the objective function, i.e., x, can be eliminated from the problem by means of Eq. (2c). Nevertheless, we will keep x as an optimization variable for the sake of a clearer presentation. Finally, the solution of (2), i.e. the vectors $\Delta x(k^0)$, x, w, and v, gives us an optimal estimate of $x^\diamond$, $w^\diamond$, and $v^\diamond$ of system (1), respectively.

In Eq. (2a), the weighting matrices are typically chosen as the inverse of the covariance matrices of $\Delta x(k^0)$, $w(k)$ and $v(k)$, i.e. such that $\tilde{P}(k) = P^{-1}(k)$, $\tilde{Q} = Q^{-1}$ and $\tilde{R} = R^{-1}$. This way, compatibility with the stochastic framework of Kalman filtering can be achieved [3]. To this end, the optimal specification of the first term in Eq. (2a) is of particular concern in MHE. It reflects past measurement information and is called *arrival cost*. For linear systems, it is well know that the parameters associated with the arrival cost, i.e. $\tilde{P}(k)$ and $\bar{x}(k^0)$, should be updated using the Kalman filter equations. Then, MHE yields the same minimum-variance state-estimates as the Kalman filter [3].

The possibility of adding physical constraints, such as non-negativity of concentrations or temperature, is one of the main advantages of MHE. While we do not consider such constraints here, we believe that our method can be extended to allow additional equality and inequality constraints in the future.

Next, we partition the system model of problem (2) into N subsystems, referenced by index i. Therefore, all vectors are partitioned analogously to the state vector as $x(k) = \langle x_1(k), \ldots, x_N(k)\rangle$, and the matrices are decomposed as $A = [A_{ij}]_{i,j=1,\ldots,N}$, $C = [C_{ij}]_{i,j=1,\ldots,N}$, $\tilde{P} = [\tilde{P}_{ij}]_{i,j=1,\ldots,N}$, $\tilde{Q} = [\tilde{Q}_{ij}]_{i,j=1,\ldots,N}$ and $\tilde{R} = [\tilde{R}_{ij}]_{i,j=1,\ldots,N}$ where the submatrices have the dimensions $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $C_{ij} \in \mathbb{R}^{p_i \times n_j}$, $\tilde{P}_{ij} \in \mathbb{R}^{n_i \times n_j}$, $\tilde{Q}_{ij} \in \mathbb{R}^{n_i \times n_j}$ and $\tilde{R}_{ij} \in \mathbb{R}^{p_i \times p_j}$.

In the MHE problem (2), all the subsystems are coupled in both, the objective function (2a) as well as in the constraints (2c) and (2d). In order to solve this estimation problem in a decentralized manner, it is rewritten in terms of N subproblems. To this end, problem (2) is first reformulated as

$$\min_{\boldsymbol{x},\boldsymbol{w},\boldsymbol{v}} \sum_{i=1}^{N} \Phi_i \tag{3a}$$

$$\text{s.t. } \Phi_i = \frac{1}{2}\sum_{j=1}^{N}\left(\boldsymbol{w}_i^T \tilde{\mathcal{Q}}_{ij}\boldsymbol{w}_j + \boldsymbol{v}_i^T \tilde{\mathcal{R}}_{ij}\boldsymbol{v}_j\right), \tag{3b}$$

$$0 = -\boldsymbol{w}_i + \boldsymbol{x}_i - \bar{X}_i(k^0) - \sum_{j=1}^{N}\boldsymbol{A}_{ij}\boldsymbol{x}_j, \tag{3c}$$

$$0 = -\boldsymbol{v}_i + \boldsymbol{y}_i^\diamond - \sum_{j=1}^{N}\boldsymbol{C}_{ij}\boldsymbol{x}_j, \tag{3d}$$

$$\forall i \in \{1,\ldots,N\}.$$

This formulation relies on the defined vectors

$$\begin{aligned}
\boldsymbol{x}_i &= \left\langle x_i(k^0),\ \ldots,\ x_i(k')\right\rangle,\\
\boldsymbol{x} &= \left\langle \boldsymbol{x}_1,\ \ldots,\ \boldsymbol{x}_N\right\rangle,\\
\boldsymbol{w}_i &= \left\langle \Delta x_i(k^0),\ w_i(k^0),\ \ldots,\ w_i(k'-1)\right\rangle,\\
\boldsymbol{w} &= \left\langle \boldsymbol{w}_1,\ \ldots,\ \boldsymbol{w}_N\right\rangle,\\
\boldsymbol{v}_i &= \left\langle v_i(k^0),\ \ldots,\ v_i(k')\right\rangle,\\
\boldsymbol{v} &= \left\langle \boldsymbol{v}_1,\ \ldots,\ \boldsymbol{v}_N\right\rangle,\\
\boldsymbol{y}_i^\diamond &= \left\langle y_i^\diamond(k^0),\ \ldots,\ y_i^\diamond(k')\right\rangle,
\end{aligned}$$

and matrices

$$\begin{aligned}
\tilde{\mathcal{Q}}_{ij} &= \operatorname{diag}\left(\tilde{P}_{ij},\ \tilde{Q}_{ij},\ \ldots,\ \tilde{Q}_{ij}\right),\\
\tilde{\mathcal{R}}_{ij} &= \operatorname{diag}\left(\tilde{R}_{ij},\ \ldots,\ \tilde{R}_{ij}\right),\\
\bar{X}_i(k^0) &= \left\langle \bar{x}_i(k^0),\ 0,\ \ldots,\ 0\right\rangle,\\
\boldsymbol{C}_{ij} &= \operatorname{diag}\left(C_{ij},\ \ldots,\ C_{ij}\right),\\
\boldsymbol{A}_{ij} &= \begin{bmatrix} 0 & & & \\ A_{ij} & \ddots & & \\ & \ddots & \ddots & \\ & & A_{ij} & 0 \end{bmatrix}.
\end{aligned}$$

Further introducing the symbols

$$\begin{aligned}
\boldsymbol{z}_i &= \langle \boldsymbol{x}_i,\ \boldsymbol{w}_i,\ \boldsymbol{v}_i\rangle,\\
\boldsymbol{z} &= \langle \boldsymbol{z}_1,\ \ldots,\ \boldsymbol{z}_N\rangle,\\
\mathcal{T}_{ij} &= \operatorname{diag}\left(0,\ \tilde{\mathcal{Q}}_{ij},\ \tilde{\mathcal{R}}_{ij}\right),\\
X_i &= \left\langle \bar{X}_i(k^0),\ 0,\ -\boldsymbol{y}_i^\diamond\right\rangle,\\
\mathcal{I}_i &= \begin{bmatrix} I & -I & 0\\ 0 & 0 & -I\end{bmatrix},\\
\mathcal{A}_{ij} &= \begin{bmatrix} \boldsymbol{A}_{ij} & 0 & 0\\ \boldsymbol{C}_{ij} & 0 & 0\end{bmatrix},
\end{aligned}$$

we finally obtain a compact formulation of Eq. (2), i.e.,

$$\min_{\boldsymbol{z}} \sum_{i=1}^{N} \Phi_i \tag{4a}$$

$$\text{s.t. } \Phi_i = \frac{1}{2}\sum_{j=1}^{N} \boldsymbol{z}_i^T \mathcal{T}_{ij}\boldsymbol{z}_j, \tag{4b}$$

$$0 = c_i, \qquad \text{where} \tag{4c}$$

$$c_i = \left(\mathcal{A}_{ii} - \mathcal{I}_i\right)\boldsymbol{z}_i + X_i + \sum_{\substack{j=1\\ j\neq i}}^{N}\mathcal{A}_{ij}\boldsymbol{z}_j, \tag{4d}$$

$$\forall i \in \{1,\ldots,N\}.$$

Eq. (4) is a quadratic program (QP) which is still coupled in the objective function and in the constraints. An algorithm to solve this problem in a decentralized manner will be presented in the next section.

III. Partition-based Moving-Horizon Estimation

In [7] and [14], a distributed method has been proposed to solve strictly convex QPs of type (4). It is shown, that the unique global minimizer of such QPs may be found as the solution to the following sequence of QPs as $l \to \infty$:

$$\boldsymbol{z}_j^{[l+1]} = \arg\min_{\boldsymbol{z}_i}\ \Phi_i^* \tag{5a}$$

$$\begin{aligned}
\text{s.t. } \Phi_i^* = &\frac{1}{2}\boldsymbol{z}_i^T\ \mathcal{T}_{ii}\ \boldsymbol{z}_i + \frac{1}{2}\sum_{\substack{j=1\\ j\neq i}}^{N}\boldsymbol{z}_i^T\ \mathcal{T}_{ij}\ \boldsymbol{z}_j^{[l]}\\
&+\left[\sum_{\substack{j=1\\ j\neq i}}^{N}\left.\frac{\partial\,\Phi_j}{\partial \boldsymbol{z}_i}\right|_{\boldsymbol{z}^{[l]}} - \left(\lambda_j^{[l]}\right)^T\left.\frac{\partial\, c_j}{\partial \boldsymbol{z}_i}\right|_{\boldsymbol{z}^{[l]}}\right](\boldsymbol{z}_i - \boldsymbol{z}_i^{[l]}),
\end{aligned} \tag{5b}$$

$$0 = (\mathcal{A}_{ii} - \mathcal{I}_i)\boldsymbol{z}_i + X_i + \sum_{\substack{j=1\\ j\neq i}}^{N}\mathcal{A}_{ij}\boldsymbol{z}_j^{[l]}, \tag{5c}$$

$$\forall i \in \{1,\ldots,N\},$$

where the upper index $[l]$ indicates the l-th iteration towards the optimal solutions $\boldsymbol{z}_j, j = 1,...,N$, and the variables $\lambda_j^{[l]}$ are the so-called Lagrange multipliers, where $l = 1,...,L$. In practice, the number of iterations L is selected according to a stopping criterion. In this work, we apply a fixed number of iterations L. Thus, a solution for problem (4) may be obtained (i) by including first order sensitivities of both, the individual objective functions and the constraints in the adapted objective functions Φ_i^* of each subsystem, and (ii) by an iterative (parallel) solution strategy. We will refer to this method as sensitivity-driven partition-based moving horizon estimation (S-PMHE), if the data of this problem are used as introduced in Section II.

We finish this section with two remarks. First, note that convergence of S-PMHE to the solution of the central MHE can be guaranteed asymptotically if both, the weighting matrices $\tilde{P}$, $\tilde{Q}$ and $\tilde{R}$ as well as the data X_i, $i \in 1, \ldots, N$, used by both schemes are identical at every sampling time. Second, we have stated above that MHE is only equivalent to the Kalman filter if the arrival cost is updated in a specific way [3]. In particular, the initial state covariance matrix $\tilde{P}$ and the initial state estimate $\bar{x}(k^0)$ must be inferred from previous horizons, requiring knowledge of the overall dynamic plant model, if no specific measures are taken. To pragmatically avoid this issue at the current state of research, we assume for now that the arrival cost parameters are given and return to this open issue at a later point in time. If the optimal arrival cost were available at every sampling time, the S-PMHE algorithm would converge to the optimal Kalman filter estimates without knowledge of the overall dynamic plant model.

IV. Convergence and Optimality Analysis

Next, we study the convergence and optimality properties of the S-PMHE method. Because of the similar problem structure, the convergence analysis proceeds exactly as presented in [14]. It is adjusted to the estimation problem and is re-stated here for completeness. Thus, the necessary conditions of optimality are stated for the decomposed QP (5). The Lagrangian function for each subsystem i is given as

$$\mathcal{L}_i = \Phi_i^*(\boldsymbol{z}) - \lambda_i^T c_i(\boldsymbol{z}). \tag{6}$$

The necessary conditions of optimality are

$$\begin{aligned}\frac{\partial \mathcal{L}_i}{\partial \boldsymbol{z}_i} = 0 = \mathcal{T}_{ii}\boldsymbol{z}_i^{[l+1]} + \sum_{\substack{j=1\\ j\neq i}}^{N} \mathcal{T}_{ij}\boldsymbol{z}_j^{[l]} \\ - \sum_{\substack{j=1\\ j\neq i}}^{N} \mathcal{A}_{ji}^T\lambda_j^{[l]} - (\mathcal{A}_{ii} - \mathcal{I}_i)^T\lambda_i^{[l+1]}\end{aligned} \tag{7}$$

and

$$\begin{aligned}\frac{\partial \mathcal{L}_i}{\partial \lambda_i} &= 0 = c_i(\boldsymbol{z}) \\ &= (\mathcal{A}_{ii} - \mathcal{I}_i)\boldsymbol{z}_i^{[l+1]} + X_i + \sum_{\substack{j=1\\ j\neq i}}^{N} \mathcal{A}_{ij}\boldsymbol{z}_j^{[l]}.\end{aligned} \tag{8}$$

We define a mapping ζ_i with $(\boldsymbol{z}_i^{[l+1]}, \lambda_i^{[l+1]}) = \zeta_i(\boldsymbol{z}^{[l]}, \lambda^{[l]})$, where $\lambda = \langle \lambda_1, \ldots, \lambda_N \rangle$ is the aggregated vector of Lagrange multipliers. Aggregating (7) and (8) for all i results

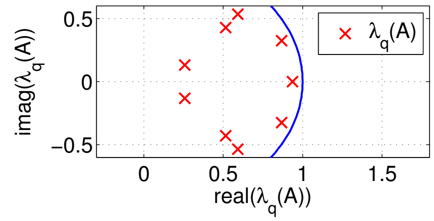

Fig. 1. Eigenvalues of the system considered in the case study. The solid line indicates the stability bound.

in the mapping $(\boldsymbol{z}^{[l+1]}, \lambda^{[l+1]}) = \zeta(\boldsymbol{z}^{[l]}, \lambda^{[l]})$ given by

$$\begin{bmatrix}\boldsymbol{z}^{[l+1]}\\ \lambda^{[l+1]}\end{bmatrix} = \underbrace{\Xi_{\mathrm{d}}^{-1}\Xi_1\begin{bmatrix}\boldsymbol{z}^{[l]}\\ \lambda^{[l]}\end{bmatrix} + \Xi_{\mathrm{d}}^{-1}\Xi_0}_{=\zeta(\boldsymbol{z}^{[l]}, \lambda^{[l]})} \tag{9a}$$

with

$$\Xi_{\mathrm{d}} = \begin{bmatrix}\mathcal{T}_{\mathrm{d}} & (\mathcal{I}_{\mathrm{d}} - \mathcal{A}_{\mathrm{d}})^T\\ (\mathcal{I}_{\mathrm{d}} - \mathcal{A}_{\mathrm{d}}) & 0\end{bmatrix}, \tag{9b}$$

$$\Xi_1 = \begin{bmatrix}(\mathcal{T}_{\mathrm{d}} - \mathcal{T}) & (\mathcal{A} - \mathcal{A}_{\mathrm{d}})^T\\ (\mathcal{A} - \mathcal{A}_{\mathrm{d}}) & 0\end{bmatrix}, \tag{9c}$$

$$\Xi_0 = \begin{bmatrix}0\\ X_0\end{bmatrix}. \tag{9d}$$

The matrices introduced in this mapping are defined as

$$\begin{aligned}\mathcal{A}_{\mathrm{d}} &= \mathrm{diag}(\mathcal{A}_{11}, \ldots, \mathcal{A}_{NN}),\\ \mathcal{I}_{\mathrm{d}} &= \mathrm{diag}(\mathcal{I}_1, \ldots, \mathcal{I}_N),\\ \mathcal{T}_{\mathrm{d}} &= \mathrm{diag}(\mathcal{T}_{11}, \ldots, \mathcal{T}_{NN}),\\ \mathcal{A} &= (\mathcal{A}_{ij})_{i,j=1,\ldots,N},\\ \mathcal{T} &= (\mathcal{T}_{ij})_{i,j=1,\ldots,N},\\ X_0 &= \langle X_1, \ldots, X_N \rangle.\end{aligned}$$

Applying the contraction mapping theorem [15] leads to the sufficient convergence condition

$$\gamma = \|\Xi_{\mathrm{d}}^{-1}\Xi_1\| < 1. \tag{10}$$

If condition (10) holds and if problem (4) is strictly convex, then the solution of S-PMHE converges asymptotically to the unique and optimal solution of the original problem (4) [14]. This directly leads to conditions for nominal stability of S-PMHE: If S-PMHE satisfies the convergence condition (10) and the centralized MHE (2) is an asymptotically stable observer for the nominal system (1) as defined in [16], i.e. when $w^\diamond(k) = v^\diamond(k) = 0$, $\forall k$ in (1), then S-PMHE is an asymptotically stable observer for the nominal system as well. Sufficient conditions for nominal asymptotic stability of the centralized MHE (2) can be found in [16].

V. Case Study

For illustration of the suggested S-PMHE method, we apply it exemplarily to estimate the state of the autonomous discrete-time linear system (1) with

$$A = \begin{bmatrix} A_{11} & \mathbf{0} & A_0 \\ A_0 & A_{22} & \mathbf{0} \\ \mathbf{0} & A_0 & A_{33} \end{bmatrix}, \quad C = \begin{bmatrix} C_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_0 \end{bmatrix},$$

$$A_{11} = \begin{bmatrix} 0.7 & 0.5 & 0 \\ 0 & 0.7 & 0.5 \\ 0.1 & -0.6 & 0.5 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 0.7 & 0.5 & 0 \\ 0 & 0.7 & 0.5 \\ -0.1 & -0.1 & 0.6 \end{bmatrix},$$

$$A_{33} = \begin{bmatrix} 0.7 & 0.5 & 0 \\ 0 & 0.7 & 0.5 \\ -0.1 & -0.3 & 0.1 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.3 & 0 & 0 \end{bmatrix},$$

$$C_0 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Thus, the variable dimensions of the subsystems are $n_i = 3$ and $p_i = 1$, $i = 1, 2, 3$. The noise covariance matrices are $Q = \text{diag}(Q_{11}, Q_{22}, Q_{33})$ and $R = \text{diag}(R_{11}, R_{22}, R_{33})$ with $Q_{ii} = I \in \mathbb{R}^{3\times3}$, and $R_{ii} = 0.1,\ i = 1, 2, 3$. Finally, the initial conditions are given as

$$x_0^\diamond = [3, 0.1, 0.1, -2, 0.1, 0.1, 1, 0.1, 0.1]^T.$$

As in [14], the system consists of three cascaded subsystems with a feedback representing a recycle. Hence, there is a strong coupling between the variables of each subsystem, but a weaker coupling between the different subsystems represented by the zero submatrices and the sparse submatrices A_0. Note that both, the overall system as well as the individual subsystems are observable and stable. While stability of the dynamic system is not required for successful state estimation, it is just more convenient to study open-loop stable systems. The eigenvalues of A are shown in Fig. 1 for reference.

The (centralized) MHE problem for this case study is

$$\begin{aligned} \min_{\Delta x(k^0), x, w, v} \frac{1}{2} \Bigg(& \left\| \Delta x(k^0) \right\|_{\tilde{P}}^2 \\ & + \sum_{k=k^0}^{k'-1} \| w(k) \|_{\tilde{Q}}^2 + \sum_{k=k^0}^{k'} \| v(k) \|_{\tilde{R}}^2 \Bigg), \\ \text{s.t. } x(k^0) &= \bar{x}(k^0) + \Delta x(k^0), \\ x(k+1) &= Ax(k) + w(k),\ k = k^0, \ldots, k'-1, \\ y^\diamond(k) &= Cx(k) + v(k),\ k = k^0, \ldots, k', \end{aligned}$$

where $\tilde{Q} = Q^{-1}$, $\tilde{R} = R^{-1}$, $k^0 = k' - K$, $K = 4$, and $k' = 0, 1, \ldots, 50$. For the computation of the arrival cost, we use a constant weighting matrix $\tilde{P} = \text{diag}(\tilde{P}_{11}, \tilde{P}_{22}, \tilde{P}_{33})$ with $\tilde{P}_{ii} = 0.1,\ i = 1, 2, 3$, and a simple strategy for guessing the initial state of the estimator. We use $\bar{x}(k^0) = x(k^0)$, where $x(k^0)$ is taken from the solution of the optimization problem solved on the previous horizon. Until the horizon has reached its full length, i.e. for $k' \le K$, we set $\bar{x}(k^0) = x_0^\diamond$.

TABLE I

CONVERGENCE CRITERION: LIPSCHITZ CONSTANT γ

k'	0	1	2	3	≥ 4
$\gamma(k')$	0	0.32	0.45	0.48	0.55

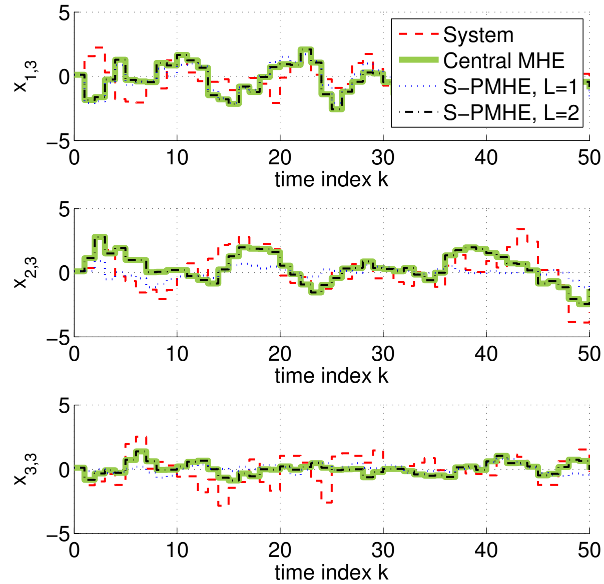

Fig. 2. Typical trajectories of states and state estimates

First, we analyze the mapping ζ for the system considered. The corresponding Lipschitz constant γ is given in Table I. It increases as the horizon length is growing initially. This can be explained by the structural changes of the matrices involved in the computation of the Lipschitz constant. As the horizon length increases, so does the dimension of Ξ_d and Ξ_1. Additional rows and columns then influence the value of the Lipschitz constant. However, $\gamma(k') < 1,\ \forall k'$. Furthermore, the matrices $\tilde{P}$, $\tilde{Q}$ and $\tilde{R}$ are positive definite. In light of the discussion earlier in the paper, this implies that the problem is strictly convex. Thus, according to Eq. (10), the S-PMHE method is convergent and applicable for the case study considered.

Next, we compare the new partition-based estimator to a centralized MHE with identical horizon length in order to judge its performance.

TABLE II

ESTIMATION PERFORMANCE: ABSOLUTE ESTIMATION ERROR e, RELATIVE PERFORMANCE Φ_{REL} AND AVERAGE COMPUTING TIME $\bar{t}$ FOR THE ESTIMATORS CONSIDERED.

Method	L	e	Φ_{rel} [%]	$\bar{t}$ [s]
Cen. MHE	–	11.93	0	0.39
S-PMHE	1	13.18	10.48	3×0.22
S-PMHE	2	11.93	0	3×0.40

For S-PMHE, two fixed iteration numbers, i.e., $L = 1, 2$, are chosen. The two decentralized S-PMHE schemes and the centralized MHE are implemented and applied to the case study. Figure 2 shows representative trajectories for each of the subsystems. It can be seen that the centralized MHE estimates the system states reasonably well, though it does not converge due to the presence of process and measurement noise. Clearly, the S-PMHE scheme with only one iteration does not result in comparable estimation quality. However, if two iterations are employed, the state estimates determined by MHE and S-PMHE are almost indistinguishable. This exemplary result not only demonstrates the capability of the S-PMHE algorithm to converge asymptotically to the centralized MHE solution, but also reveals – in full accordance with previous experience with a similar algorithm in the context of control [14] – that only few iterations are typically sufficient for high performance of the decentralized scheme. A more detailed performance assessment is given in Table II. It shows the absolute estimation error, computed as

$$e = \sum_{k=0}^{50} \|x^{\diamond}(k) - x(k)\|,$$

and a relative performance measure, computed as $\Phi_{\text{rel}} = (e - e_{\text{ref}})/e_{\text{ref}}$, where e_{ref} is the absolute estimation error of the reference method, i.e. the centralized MHE. These numbers emphasize that S-PMHE is able to compete with a centralized MHE at a very low number of iterations. In addition, Table II provides the average computing time $\bar{t}$ for each of the methods implemented. All estimators employ the standard Matlab QP solver `quadprog` on an Intel Core i3-2100 desktop machine. Only one core has been assigned to Matlab for either the MHE and the S-PMHE schemes, in order to prevent automatic parallelization of the code. For S-PMHE, a parallel solution of the QP is straightforward by using different processors for the solution of each subproblem. Hence, the average computing time may then be further reduced for the S-PMHE schemes. For instance, the average computing time of the S-PMHE algorithm with $L = 1$ iteration, implemented on three cores, is $\bar{t} = 0.22s$, 43% less than that of the centralized MHE.

VI. Conclusions

In this paper, a novel iterative partition-based moving horizon estimator for structured large-scale systems has been presented. Similar to the sensitivity-driven distributed model predictive control method (S-DMPC) developed in [7] and [14], the novel algorithm relies on a sensitivity-based coordination of the subproblems of the partitioned optimization problem. It is consequently called sensitivity-driven partition-based moving horizon estimator (S-PMHE). Iteratively, every subsystem estimates a distinct, non-overlapping part of the overall state vector. The iteration eventually converges to the optimal state estimate obtained with a centralized MHE. Hence, all properties of MHE are inherited by S-PMHE. In contrast to many estimators suggested in literature so far, the estimators of the subsystems in S-PMHE only need to know the part of the overall system dynamics which they cover.

While we did not explicitly consider constraints in this paper, we are confident that they can be included easily along the lines of [7]. This extension is necessary to fully exploit the potential of MHE in contrast to traditional state estimation methods based on Kalman filters for large-scale structured systems. In terms of performance, further improvements are also expected from different initializations. For instance, less iterations may be required if the best estimates from the previous sampling time are used to initialize the variables of the current horizon. Finally, future work will focus on decentralized computation of the arrival cost. Currently, the optimal arrival cost must be known at every sampling time in order to obtain the same state estimates as with the Kalman filter.

Despite these open challenges, the promising results obtained in the case study make us expect that a generally applicable and high-performant decentralized constrained state estimator can be developed for structured large-scale systems as they typically arise in the context of process and energy systems.

VII. Acknowledgement

This research has been partially supported by Cybernetica AS, Norway, through the project "Model identification methods for online monitoring, control and economic optimization".

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, Series D, pp. 35–45, 1960.

[2] D. G. Luenberger, "Observing the state of a linear system," *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74–80, 1964.

[3] K. R. Muske, J. B. Rawlings, and J. H. Lee, "Receding horizon recursive state estimation," in *Proceedings of the 1993 American Control Conference*, 1993, pp. 900 –904.

[4] H. Michalska and D. Mayne, "Moving horizon observers and observer-based control," *IEEE Transactions on Automatic Control*, vol. 40, no. 6, pp. 995–1006, June 1995.

[5] R. Scattolini, "Architectures for distributed and hierarchical model predictive control - a review," *Journal of Process Control*, vol. 19, pp. 723–731, 2009.

[6] M. Farina, R. Scattolini, J. Garcia, J. Espinosa, and J. Rawlings, "Report on the state of the art in distributed state and variance estimation, and on preliminary results on disturbance modelling for distributed systems," European FP7 project HD-MPC, Tech. Rep., 2010. [Online]. Available: http://www.ict-hd-mpc.eu/deliverables/hd_mpc_D_5_1.pdf

[7] H. Scheu and W. Marquardt, "Sensitivity-based coordination in distributed model predictive control," *Journal of Process Control*, vol. 21, no. 5, pp. 715–728, 2011.

[8] I. Alvarado, D. Limon, D. Muñoz de la Peña, J. Maestre, M. Ridao, H. Scheu, W. Marquardt, R. Negenborn, B. De Schutter, F. Valencia, and J. Espinosa, "A comparative analysis of distributed MPC techniques applied to the HD-MPC four-tank benchmark," *Journal of Process Control*, vol. 21, no. 5, pp. 800 – 815, 2011.

[9] M. Hassan, G. Salut, M. Singh, and A. Titli, "A decentralized computational algorithm for the global Kalman filter," *IEEE Transactions on Automatic Control*, vol. 23, no. 2, pp. 262–268, apr 1978.

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

[11] B. Rao and H. Durrant-Whyte, "Fully decentralised algorithm for multisensor Kalman filtering," *IEE Proceedings D (Control Theory and Applications)*, vol. 138, no. 5, pp. 413–420, 1991.

[12] I. Necoara, V. Nedelcu, and I. Dumitrache, "Parallel and distributed optimization methods for estimation and control in networks," *Journal of Process Control*, vol. 21, no. 5, pp. 756 – 766, 2011.

[13] 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.

[14] H. Scheu and W. Marquardt, "Distributed model-predictive control driven by simultaneous derivation of prices and resources," in *Proceedings of the 18th IFAC World Congress*, 2011, pp. 398–403.

[15] H. K. Khalil, *Nonlinear Systems*. Prentice Hall. Inc., 2000.

[16] C. V. Rao, J. B. Rawlings, and J. H. Lee, "Constrained linear state estimation–a moving horizon approach," *Automatica*, vol. 37, no. 10, pp. 1619–1628, 2001.