Moving Horizon Estimators for Large-Scale Systems

J. Garcia, and J. Espinosa

Research Group on Automatic (GAUNAL) ,Universidad Nacional de Colombia, Facultad de Minas. Cra 80 No 25-223, Medellin, Colombia

Abstract: In this report, a review on state estimation schemes applied to large-scale systems is made. The attention is focused on Moving Horizon Estimation (MHE) schemes due to the addressing of the estimation problem in an optimal way, and it inherent capability to handle the process constraints. Moreover, the cost function can be proposed unlike other optimal estimation schemes like those based on Kalman Filters. Therefore, contributions on state estimation schemes applied to large-scale systems are described, in order to outline its merits and limitations. Finally, open problems are listed with the aim to prepare a basis for future contributions.

Keywords: Moving horizon estimator, distributed estimation, large-scale systems, nonlinear programming, constraint addressing.

1. INTRODUCTION

Moving Horizon Estimation is becoming an important tool to estimate the state and parameters of any plant. The main added value is the possibility to address the constraints directly into the formulation of the optimization problem. As a main drawback, it must be pointed out the difficulty to handle a numerical complex problem, each sample time, as a consequence of the solution of a nonlinear optimization problem with constraints. Many authors have focused on the improvement of the numerical issues in which MHE are involved [Binder et al (2005), Kang (2006), Jørgensen et al (2004), Kraus et al (2006), Zavala et al (2007), Zavala et al (2008)]. Therefore, as the computational advances allow the use of powerful processors, and the research on numerical issues in nonlinear optimization is increasing, the development of new efficient estimation schemes for large-scale and complex systems are becoming in a feasible research area.

On the other hand, large-scale systems have been characterized by complex systems with hard nonlinearities, uncertainties, large number of variables and dynamics with different response times. Its decomposition into a set of smaller systems have been researched for many decades [Siljak and Vukcevic (1976), Sanders et al (1978), Mahil and Bommaraju (1992)]. This kind of decomposition looks for more manageable subsystems with the aim to organize them, possibly in a hierarchical way. As there are more subsystems, local control and estimation becomes more attractive, but the plant coordination and information exchange between subsystems turns more complicated. A large scale example is the traffic in a city, where sensors are collecting information about traffic conditions, and this information is distributed to the coordination system, allowing people to select alternate routes avoiding traffic jams. Similar problems may be weather prediction, total estimation of the state in a big industrial plant, etc.

Later, it will be shown why the state estimation of large-scale systems remains a challenge. The lack of a general methodology do not allow the design of efficient and fast schemes in order to guarantee convergence and stability of the estimations in a desired way. The paper is organized as follows: In Section 2, the problem statement of MHE is presented. Moreover some of the most relevant publications on MHE field are shown in order to establish the current status of the estimation scheme. The first part of Section 3 describes some general approaches to estimate the state in a large-scale system. In the next subsection, some published schemes of MHE applied to large-scale systems are also reviewed as the main subject in this note. Advantages and drawbacks are analyzed in all reviewed approaches. Finally some conclusions and open problem are presented in Section 4.

2. MOVING HORIZON ESTIMATORS

MHE strategies was born as a dual problem of the Model Predictive Control (MPC). Despite MPC and MHE procedures are quite similar, MPC technology was developed first in petroleum industry due to the dynamic complexity of the processes and the need of improved control strategies, whereas MHE theory was developed first in academia [Algöwer et al (1999)].

The basic strategy of MHE reformulate the estimation problem as a quadratic problem using a moving, fixed-size estimation window. The fixed-size window is needed to bound the computational effort to solve the problem. This is the principal difference of MHE with batch estimation problem (or full information estimator) [Findeisen (1997), Algöwer et al (1999), Rao (2000)]; once a new measurement is available, the oldest one is discarded, using the concept of window shifting. Moreover, the main advantage of MHE in comparison with another estimation schemes (like the Kalman Filter-based) is the direct constraint addressing inside the optimization problem, and the possibility to propose the cost function. However, as MHE is a limited memory filter, stability and convergence issues arise.

2.1 Problem statement

Assume a system modeled by means of the following nonlinear difference equation:

$$\begin{aligned} x_{k+1} &= f(x_k) + g(x_k, w_k) \\ y_k &= h(x_k) + v_k \end{aligned} \tag{1}$$

where some constraints are imposed over the state variables and disturbances as follows:

$$x_k \in \mathbb{X},\ w_k \in \mathbb{W} \tag{2}$$

with x_k and y_k the state and output at k sample respectively, w_k is the disturbance or model uncertainty, and v_k is the noise in the measured variables. Also, $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ with $g(\cdot, 0) = 0$, and $h : \mathbb{R}^n \to \mathbb{R}^p$. Finally it is assumed that $\mathbb{X}$ and $\mathbb{W}$ are closed with $0 \in \mathbb{W}$.

As it is desired estimate the state of the system described by Ec (1), first we define the full information estimator (FIE), and then the MHE general statement is shown. Consider an optimal constrained estimation problem:

$$(P1) : \phi_T^* = \min_{x_0,\ \{w_k\}_{k=0}^{T-1}} \phi_T(x_0, \{w_k\}_{k=0}^{T-1}), \tag{3}$$

where x_0 is the initial state, and T is the current discrete time. The problem is subject to the following constraints:

$$\begin{aligned} x_k \in \mathbb{X} \text{ for } k = 0, ..., T \\ w_k \in \mathbb{W} \text{ for } k = 0, ..., (T-1) \end{aligned} \tag{4}$$

with the cost function as:

$$\phi_T(x_0, \{w_k\}_{k=0}^{T-1}) \triangleq \sum_{k=0}^{T-1} L(w_k, v_k) + \Gamma(x_0) \tag{5}$$

Notice that, the most important constraint is the model (1). It is assumed the stage cost function as $L : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}_+$ and the initial penalty as $\Gamma : \mathbb{R}^n \to \mathbb{R}_+$. The solution to $(P1)$ at T, it is denoted by $\left\{\widehat{x}_{k|T-1}\right\}_{k=0}^{T}$ and disturbance sequence by means of $\left\{\widehat{w}_{k|T-1}\right\}_{k=0}^{T-1}$. It is pointed out that the noise is obtained from the difference between the measured outputs and the model output: $v_k = y_k - h(x_k)$. The problem $(P1)$ is referred as the full information problem (FIP) or batch state estimation problem (BSEP) in the most relevant literature [Rao (2000), Rao and Rawlings (2000), Rao et al (2001), Rao et al (2003), Findeisen (1997)].

As the problem $(P1)$ gets more information as time goes, the optimization become more complicated, due to the availability of new information. Therefore, as the problem has T stages, the computational complexity increases at least as a linear function of time. Unless the problem is linear, unconstrained, and it has a quadratic cost function (in which case the solution is equivalent to the obtained with the Kalman filter), the problem becomes infeasible to be solved on line. One way to avoid the last problem is getting a new optimization problem with a fixed window data, that is, a fixed dimension optimal problem by means of a moving horizon approximation.

Consider again the problem $(P1)$, with the cost function rewritten in an alternative way:

$$\begin{aligned} \phi_T(x_0, \{w_k\}_{k=0}^{T-1}) &= \sum_{k=0}^{T-1} L(w_k, v_k) + \Gamma(x_0) \\ &= \sum_{k=0}^{T-N-1} L(w_k, v_k) + \sum_{k=T-N}^{T-1} L(w_k, v_k) + \Gamma(x_0) \end{aligned} \tag{6}$$

where N is the moving horizon. The last cost function can be rearranged in a more compact function, and then the problem can be reformulated as:

$$(P1)^* : \min_{x_{T-N},\ \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} L(w_k, v_k) + \Xi_{T-N}(x_{T-N}) \tag{7}$$

with

$$\Xi_{T-N}(\overline{x}) = \min_{x_0,\ \{w_k\}_{k=0}^{T-N-1}} \sum_{k=0}^{T-N-1} \left\{ \phi_{T-N}(x_0, \{w_k\}_{k=0}^{T-N-1}) \right\} \tag{8}$$

subject to constraints:

$$\begin{aligned} x_k \in \mathbb{X} \text{ for } k = 0, ..., T-N, \\ w_k \in \mathbb{W} \text{ for } k = 0, ..., (T-N-1) \end{aligned} \tag{9}$$

Henceforth in this note, the initial penalty Ξ_j is referred as the *arrival cost*. Note that, the arrival cost must have all the past information over the process, in order to establish an equivalence between $(P1)$ and $(P1)^*$. In fact, some authors assume this term as an equivalent statistic that summarize the past data [Algöwer et al (1999), Rao (2000)]. The previous statement is not always true, because of the exact computation of the arrival cost is only possible if the system is linear, the cost function is quadratic, and there are not constraints. Otherwise, the arrival cost must be approximated. So the computation of the arrival cost is still a challenge for the scientific community. To see a graphical comparison between FIE and MHE, see Fig 1.

2.2 About the latest developments on MHE

This subsection deals with the most relevant results on MHE in the last years. First, stability issues in a discrete framework are presented. Then the framework is changed to a continuous time. Finally, contributions on numerical algorithms to solve the numerical optimization involved on MHE problem are also presented. Interesting MHE applications can be found in Alamir and Corriou (2003), Alamir and Sheibat-Othman (2007), Haverbeke et al (2008), Kawohl et al (2007), and Kupper et al (2008).

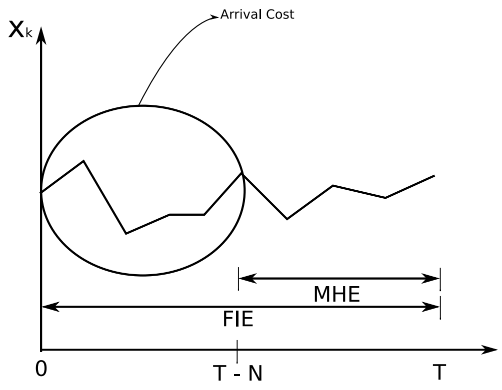

Fig. 1. Full Information and Moving Horizon formulations [Algöwer et al (1999)]

MHE Stability in a discrete framework. Stability contributions in a discrete framework have been presented in Algöwer et al (1999), Ling and Ling (1999), Rao (2000), Rao and Rawlings (2000), Rao et al (2001), Rao et al (2003), Alessandri et al (2008), and Qu and Hahn (2008). Allgöwer et al. [Algöwer et al (1999)] show the historical path of MHE procedures until the end of 90's. First, the contribution of Jang and coworkers [Jang et al (1986)] is outlined as the earliest MHE technique, but the disturbance effect and constraints addressing are ignored, focussing only on the estimation of the initial state of the system. Then, some important works are listed with the aim to show their theoretical importance, advantages, and drawbacks. Some authors investigated the logical extension of MPC to MHE, nonlinear data reconciliation, the use of inequality constraints (that is the cornerstone of the MHE procedures), stability, and convergence issues among others. The remainder of this publication presents the optimization problem statements for full information estimation and MHE for both, linear and nonlinear models subject to constraints. Then, a result on estimator stability is shown in a deterministic way for linear systems: the proposition says that the Kalman Filter covariance becomes a lower bound to the arrival cost, and then stability is guaranteed if (C, A) is observable and all the covariance matrices are positive definite. Finally, it is pointed out the challenge to formulate the MHE strategy in a nonlinear framework. The main difficulty is to solve online a nonlinear optimization problem with constraints.

In Ling and Ling (1999) it was shown a stability result for certain receding horizon recursive state estimator based on the choice of the state estimation horizon. The disadvantage of this publication lies in its limited application: only linear systems. Later, Rao et al. [Rao and Rawlings (2000)] present some fundamentals of MHE theory. The main difference with Algöwer et al (1999) are the MHE strategies with guaranteed stability, also for nonlinear systems. These results are based on the Lyapunov theory and are concerned with the search of a global bound for the arrival cost if the system is linear. It is pointed out that the stability results are mainly based on some assumptions made over the terms involved in the cost function, the fact that the system state remains inside the constraints, and the system is incrementally observable. In the nonlinear case, the bounding of the arrival cost is more challenging. Rather than attempt to bound the arrival cost, a strategy to generate a monotone decreasing sequence that bounds the optimal cost function was proposed, and based on these sequences the initial penalization is calculated. Further results on stability of MHE with linear models and constraints are presented in Rao et al (2001) extending the published analysis made in Rao and Rawlings (2000). In fact, these results are based on the "equivalence" found between the MHE with the full information estimator. Assuming the state respects the constraints all the time, it is guaranteed that the full information estimator is stable, and then conditions to reformulate the stability analysis for MHE are made. These conditions are only sufficient, limiting the result to some systems. Finally, the paper presents some smoothing strategies. Later, Rao and coworkers presented a nonlinear MHE procedure (NMHE) in Rao et al (2003). This is based on a linearization of the nonlinear model around certain operating point. Once the previous linearization is performed an approximation of the arrival cost is found as a function of the covariance matrix of the Extended Kalman Filter (EKF). Further results on stability of this approach are discussed here. The performance of this approach is tested first with a linear model compared with a Kalman Filter (KF). Then the performance of the scheme is tested with a nonlinear model with an EKF as a benchmark. Despite the results, this contribution points out the need to develop more research on NMHE.

In a recent work [Alessandri et al (2008)], a MHE strategy based on the least-squares optimization is performed, taking explicitly into account the measurement noises and disturbances uncertainties, both of them with unknown statistics. Unlike previous contributions, it is considered a bounded optimization expected error, therefore a strict optimization is not necessary. Another difference is the inclusion of a new term in the cost function, that is, a term that penalizes the deviation between the predicted initial value with the estimated one. Stability is proven over the estimation error dynamics. Finally numerical results are reported and compared with an EKF.

A novel computation of arrival cost is performed in Qu and Hahn (2008). This was done by means of the Unscented Kalman Filter (UKF), published by Julier and Uhlmann [Julier and Uhlmann (2004)], instead of EKF. This last publication demonstrates the superiority of the UKF compared with the EKF since the parameter statistics are found in a more precise way. Therefore, any arrival cost computation with UKF gives better results than the one computed with EKF. This result gives consequently convergence and then stability as it was demonstrated previously in Rao and Rawlings (2000).

Contributions in a continuous framework. Contributions in a continuous framework are presented in Alamir (2007), Alamir and Calvillo-Corona (2002), Mayne and Michalska (1992), and Michalska and Mayne (1995). Alamir and Calvillo-Corona [Alamir and Calvillo-Corona (2002)] presented further results on MHE using the continuous time approach. The main idea is as follows: with a compact set S_0 of possible initial states, it assumed the existence of certain high-gain receding horizon observer with guaranteed convergence if the initial estimation error is lower than the observability radius associated to S_0. This idea is derived assuming uniformly local observability of all system in each (t, x) pair, that is in most cases hard to guarantee. The analysis is made assuming a quadratic cost function penalizing only the output error. This publication does not mention anything about constraint handling.

Mayne and Michalska developed a MHE with the aim to get a new receding horizon structure: model based regulator plus a receding horizon observer [Mayne and Michalska (1992), Michalska and Mayne (1995)]. This is done first, building an MH observer in a deterministic way, with a simple cost function that only penalizes the output deviation. In this procedure, measurement noise and disturbance uncertainty initially are not taken into account. In fact, once the observer procedure is shown in absence of noise, a sufficient condition based in a deteriorated optimization is made in order to tolerate certain bounded measurement noise. Finally, a stabilizing composed strategy is derived from the convergence of the MHE, but it is not part of the present review.

Numerical issues. Explicitly numerical issues are studied in Kang (2006). A Moving Horizon Numerical Observer (MHNO) is developed both for linear and nonlinear systems [Kang (2006)]. The key result is different of previous approaches and it is based on the optimization and the integration errors. This is done by minimizing certain cost function that penalizes only the deviation in the output. Then, previous minimum is improved with an approximate Newton method based on an approximate of the Hessian matrix. As an advantage it can be outlined that the method is flexible since any numerical and optimization method can be applied. The main disadvantage is the lack of penalization of the disturbances in the cost function. The feasibility of the method is demonstrated with certain bounds on the errors.

3. STATE ESTIMATION IN LARGE-SCALE SYSTEMS

Large-scale systems deal with more complicated process than is usual. A concrete definition of this kind of systems is not agreed by the scientific community, but some concrete features can be pointed out: high or infinite dimension, multi-objects, model uncertainty (due to randomness, fuzzyness, etc), special architectures, large number of variables, etc. Due to the high dimensionality, it is impossible (or uneconomical) to carry out some centralized calculation. It is mandatory to decompose the large-scale system into several coupled subsystems. Therefore, estimation and control procedures become a hard task. As it will be seen, large-scale monitoring and control is still a challenge. Moreover, large-scale state estimation with MHE procedures is still an unexplored work area.

3.1 A review on state estimation schemes applied to large-scale systems

It is important to explain some relevant schemes applied in order to solve this problem. The identification of what has been done and what remain a challenge in this area is mandatory in order to propose a future contribution. Therefore, a review of state estimation schemes applied to large-scale systems can be classified in two main groups: purely decentralized state estimators and distributed or partially decentralized estimators. As it was stated earlier, the centralized framework become a performance reference, instead of a possibility to implement in practice.

Fully decentralized schemes deal with a set of subsystems which perform control and/or estimation tasks in an autonomous way, that is, there is not information exchange between subsystems. Some works about estimation based on decentralization of linear systems was published in Siljak and Vukcevic (1976), Sanders et al (1978), Chen and Lu (1988), and Mahil and Bommaraju (1992). First, Siljak and Vukcevic (1976) address process stabilization and state estimation designs for a set of input-decentralized and output-decentralized linear models respectively. As a main disadvantage it must be outlined that the duality concept is defined only for linear systems. If the system is nonlinear the design must be performed with other additional assumptions. Another disadvantage is the impossibility to address constraints. Finally, the performance of this straightforward scheme was not even tested in simulation nor in real applications. In a subsequent publication [Katti (1982)], it is shown by means of a counterexample that the previous scheme does not always work.

In a later paper [Sanders et al (1978)], the authors considered some specific structures to develop certain state observers for large-scale linear systems. The performance of such structures are evaluated with different levels of information exchange between local agents. The analysis was made first considering noisy channels between decoupled systems and additive disturbances in the state equation. An important idea taken of this paper is whatever structure is chosen, there exist important parameters that could modify the performance of the state estimation in one or another way. The main contribution of this paper is the use of the *interaction measurement of noise crossover level* between two different structures with the aim to compare the performance. The comparison of each structure is made taking into account two levels of information: process model information and measurement information. Note that it is possible to limit the last one to a desired level. Perhaps the lack of generality is the main limitation of this contribution, since it can only be applied to linear models. Also, the analysis becomes intractable as more subsystems are considered because the results are obtained when there are two subsystems. Finally this scheme does not take into account constraints.

On the other hand distributed state estimation schemes deals with the capability of local agents (local filters) to intercommunicate between them in some way, in order to improve the overall estimation. A generic distributed state estimator can be depicted in Fig. 2. The interconnection network is a generalized channel to perform the information exchange between subsystems, in a desired fashion. x_{1e}, x_{2e},..., x_{ne} are the local state estimates of each subsystem. Moreover, a mathematical description of the overall linear large-scale system, and then a decomposition to subsystems can be written as:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{10}$$

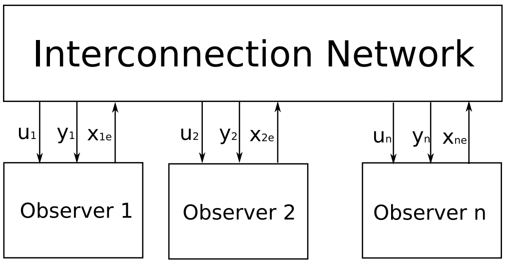

Fig. 2. Distributed observer structure

$$S_i : \dot{x}_i = A_{ii}x_i + B_{ii}u_i + \sum_{j=1,j\neq i}^{N} A_{ij}x_j + \sum_{j=1,j\neq i}^{N} B_{ij}u_j$$
$$y_i = C_{ii}x_i \tag{11}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$ are the state, input, and output of the whole system respectively. A, B, and C are the state, input, and output matrices of the large-scale system, and S_i with $i = 1, ..., N$ is each decomposed subsystem, N the number of subsystems. Note that the model dependence with other subsystems is shown in a explicitly way. In an analogous manner, the nonlinear counterpart can be described mathematically as follows:

$$\begin{aligned} \dot{x} &= f(x) + g(x,u) \\ y &= h(x) \end{aligned} \tag{12}$$

$$\begin{aligned} S_i : \dot{x}_i &= f_i(x_i, ..., x_{i-1}, x_i, x_{i+1}, ..., x_N, u) \\ y_i &= h_i(x_i) \end{aligned} \tag{13}$$

where $i = 1, ..., N$. Moreover f, g, and h are vector fields with real valued functions as components. In this equation, it can be seen that the dependence with other subsystems is not generic nor explicit.

Then, in a distributed framework some publications can be outlined: Abdel-Jabbar (1998), Vadigepalli and Doyle (2003), and Vadigepalli and Doile III (2003). Two critical issues are further discussed in Abdel-Jabbar (1998): distributed state estimation in large-scale systems, and its parallel computer implementation. Moreover a complete design including a distributed design of the state estimation of the whole system based on a parallel program and a fully distributed control is presented and tested in two complex chemical examples. The parallel simulation is performed by means of a dynamic block Jacobi-like iteration scheme. Moreover, in order to account the fact that different subsystems may have different time responses it is defined an iteration time greater than all the sample times in order to make the intercommunication only at the end of this time. In fact, while the iteration time is running an interpolation method e.g. cubic splines approximate internally the states. As additional this paper studied the effect of observer gains on the convergence of the estimation, and the parallelization of the scheme.

In Vadigepalli and Doyle (2003) a structural analysis of plantwide processes for application of multi-rate distributed estimation and control is presented. First the paper discuss the need to implement a distributed estimation scheme in such a way that it performs as the centralized optimal structure. The main idea presented in this paper is to transform the global estimation and control problem to a fully distributed and decentralized structure using some results on multisensor data fusion. This algorithm first published in Mutambara (1998) provides the required scalability and at the same time retains the global optimal performance. An extension to multi-rate systems is published in Vadigepalli and Doile III (2003). The scalable distributed methodology is presented in two steps: first an appropriated model decomposition in N nodes is made and then distributed estimators and control are designed locally. In Fig. 3 it can be seen a generic distributed estimation structure. The decomposition procedure can be summarized as follows: first identification of the units in the process are made in an heuristic way to get the so called computation nodes. Later the discretized model is obtained from the continuous one, taking into account the sample time as a critical parameter to preserve the sparseness of the state matrix. Then the overlapping states are identified at each node based on the discrete-time global state transition matrix if available. Otherwise the use of the plant flowsheet in a heuristically way is needed. The computational load at each node must be revaluated continuously in order to redistribute the nodes when it is necessary. If the communication is excessive due to large number of overlapping states a new process decomposition must be made again with a lower sampling time. On the other hand, the estimation task is performed with a distributed Kalman filter. This is done in two steps as usual: first a prediction is made and then an update in the estimation is achieved. Moreover, the overall estimation step consists of three stages: (i) local estimation, (ii) internodal communication, and (iii) assimilation to produce a global estimate. Control issues are not outlined here. Finally, as key issues we outline: local state vector at node i is related to the global state vector by means of a linear nodal transformation matrix. In this procedure only orthonormal transformations are taken into account for simplicity; excessive partitioning of the system results in increased communication load while reducing computational burden on each node; the sample time needs to be sufficiently small to preserve the structure and sparseness in the state transition matrix. The structural analysis of two plantwide processes (an industrial reaction-separation and a pulp mill process) are shown. The main drawback of this contribution is that the estimation procedure by a distributed Kalman filter does not allows constraint handling. Moreover, the linear framework limits the global operation of a nonlinear plant.

In a subsequent publication [Vadigepalli and Doile III (2003)], the authors show the multirate distributed and decentralized approach for large-scale processes commented in Vadigepalli and Doyle (2003). This contribution is an extension of the algorithm proposed in Mutambara (1998) for multirate complex systems. Theoretical results are presented in Vadigepalli and Doyle (2003). As a difference, the plantwide estimator and control designs are presented in this paper for the separation-reaction process described in Luyben et al (1997). In this application, the effect of the sample time on the sparseness of the transition matrix is shown. Also, some topologies are presented in order to test the computational load of the distributed nodes and intercommunication issues. For the estimation problem, the prediction error covariance

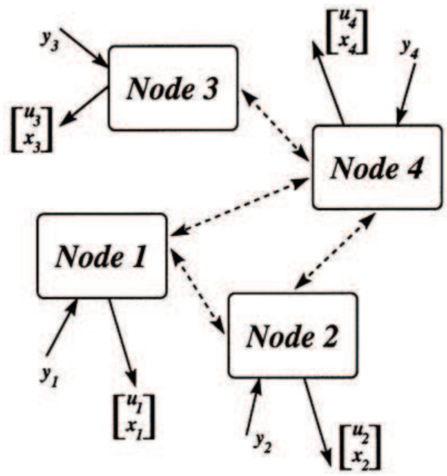

Fig. 3. Typical Distributed Estimation and Control Network

matrices P_i involved in prediction and estimation steps are computed offline. Simulation results are performed in a parallel framework using MATLAB/SIMULINK. The main contribution of this paper is the distributed scheme with a centralized performance guaranteed for estimation purposes. On the other hand as main drawback it can be pointed out that the estimation procedure by a distributed Kalman filter does not allows constraint handling. Also the linear framework limits the global operation of a nonlinear plant.

An interesting application is performed in Mjaavatten and Foss (1997). In this paper a Topology-Based Diagnosis is described in an analytical redundancy framework. In fact, some observers are designed locally in a decentralized way, taking into account the normal plant operation. The modular estimation method used is the proposed by Sanders [Sanders et al (1974)]. Despite, its contribution is not focused directly on estimation of large-scale systems, they propose an interesting viewpoint about the modularization of plantwide processes. This allows to use any agent (equipment) as an information box and then using it in other applications. The central subject of this paper is the development of a modular process representation suitable for the design of flexible operator support systems. The methodology is tested in an operator support system in a fertilizer plant. As an advantage, it is stated the modularization concept. Some drawbacks can be outlined: first the estimation is made in a linear framework, then only a restricted operation must be taken into account. Also the decentralization scheme guarantees local estimation of the state of any agent but issues like optimality and the relationship between the decentralized scheme and an optimal centralized estimator are not discussed. Finally there is no constraint handling.

3.2 MHE schemes applied to large-scale systems

MHE has been proved as a powerful strategy to estimate states, parameters, and disturbances in an optimal way, allowing constraints handling and taking into account the measurement noises and state uncertainties. Also, it must be pointed out the versatility of the scheme, because it consider some punctual penalties inside the cost function. The main drawback, is the computational burden generated by solving at each sample time an optimization problem with constraints that in most cases is a nonconvex problem. Another difficult issue is the approximation made to perform the receding horizon procedure by means of the arrival cost.

The reviewed works are classified in order of its main contribution. First, numerical issues for MHE algorithms has been presented in Jørgensen et al (2004), Kraus et al (2006), Zavala et al (2007), Zavala et al (2008), and Biegler and Zavala (2009). In the first publication only numerical methods for constrained linear MHE with quadratic functions are shown. The main contribution of this work is the use of KKT (Karush-Kuhn-Tucker) system as a direction finder in all of primal active set, dual active set, and interior-point algorithms, improving the optimization time. This KKT system may be solved in a recursive way by means of a Ricatti equation, even if there are constraints. The paper title mentioned the usefulness of the contribution to large-scale systems, but there are not explicit elements.

Later in Zavala et al (2007), Zabala and coworkers proposed a fast computational framework for Large-Scale MHE in a centralized fashion. First it is assumed a quadratic functional cost penalizing the moving initial condition, the parameters, and the output deviation. Constraints are tackled as appropriate barrier terms in the cost function. Once this cost function is built, then an IPOPT algorithm [Watcher and Biegler (2006)] is used for the solution of the NLP (nonlinear program). This IPOPT algorithm follows a full-space, primal-dual barrier approach and applies Newton method to the KKT conditions derived to the cost function and then find the desired solution. Finally a shifting strategy is presented, but a coupling between NMPC and MHE is needed. Therefore the last idea can not be applied at free-control systems. A successful application of the scheme is presented in a full-scale low-density polyethylene process, taking into account all 294 differential and 64 algebraic states variables.

One of the latest result was published in Biegler and Zavala (2009). An integrating framework for plantwide dynamic optimization based on fast calculations of MHE and NMPC is discussed. The theoretical fundamentals of these fast algorithms was published in Zavala et al (2008). The main idea is to integrate the real-time optimization and the control at the higher level of decision-making, that is, in the scheduling and planning levels. On-line dynamic optimization deals with solving two nonlinear problems each sample time: first the current state of the process and the model parameters must be estimated from the model measurements using a MHE scheme. Next, with the updated state, optimal values of the manipulated variables are calculated by means of an NMPC. The results for the estimation procedure was shown in Zavala et al (2008).

Application cases. Application cases can be found in Kraus et al (2006), Hedengren and Allsford (2007), and Samar and Gorinevsky (2006). An interesting application was published in Kraus et al (2006). A MHE algorithm is applied to the well-known Tennessee Eastman benchmark problem with the aim to test its performance and computational effort. The procedure is first based on a least-squares optimization problem solved numerically by the direct multiple shooting technique. A drawback is the multiple iteration done in order to solve the optimization problem at each sampling time. This would be a problem in an hypothetical real-time implementation. A possible solution to the last problem is also presented in this paper, and it is based on the application of a single generalized Gauss-Newton iteration. The procedure separates the estimation into a separation phase (it does not need the latest measurement) and the short estimation phase (it is executed once the latest measurement is available) as in Zavala et al (2007).

4. OPEN PROBLEMS AND CONCLUSIONS

This paper consider three important subjects: MHE as an estimation procedure, estimation in large-scale systems in general and finally MHE schemes applied to large-scale systems. In the first subject, it must be outlined the few theoretical results compared with the dual problem, that is, MPC. In fact, it was analyzed some relevant papers showing the lack of maturity of this area. Some important results on MHE with linear and nonlinear models are reviewed. Numerical issues are described in those papers. However, further research on stability, convergence and numerical optimization for nonconvex problems is needed.

There exist more research about large-scale state estimation in general. The main results are based on decoupling, decentralization and distribution of large-scale systems. Issues like model reduction, computational parallel formulation, subsystem communication arise. However those reviewed results has the same drawback: impossibility to handle constraints and in most cases difficulties to filter the noise and tackle the state uncertainty.

As it is well known, there exists a lot of strategies to estimate the state in a nonlinear system. However, the main drawback of almost all of these strategies is the lack of generality to be applied in any nonlinear system. Some examples are the differential geometric approach proposed in Krener and Isidori (1983), high-gain Luenberguer-like observers, Kalman Filters and its derived (Extended Kalman Filter, Unscented Kalman Filter, Ensemble Kalman Filter, etc). All of them has successful applications, but each of them fail under certain conditions. On the other hand, MHE seems to solve the previous problems, even if large-scale systems are concerned.

The challenge is focused on the design of MHE schemes applied to large-scale systems in a distributed fashion. Therefore, there is a need to formulate fast and precise nonlinear optimization solvers. Also, stability and convergence proves of those schemes must be derived in order to get new feasible procedures. Moreover, computational issues must be tackled by means of accurate parallel computational programs, taking advantage of the distributed structure.

ACKNOWLEDGEMENTS

This research has received funding from European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n. 223854.

REFERENCES

N. Abdel-Jabbar, C. Kravaris, and B. Carnahan (1998). A Partially Decentralized State Observer and Its Parallel Computer Implementation. *Ind. Eng. Chem.* (37), 2741-2760.

M. Alamir (2007). Nonlinear Moving Horizon Observers: Theory & Real-Time Implementation. In Nonlinear Observers and Applications, Gildas Besancon (Ed). Lecture Notes on Communication and Information Science. Springer-Verlag-Series.

M. Alamir, and L. A. Calvillo-Corona (2002). Further Results on nonlinear Receding -Horizon Observers. *IEEE Trans. on Auto. Cont.* **47**(7), 1184-1188.

M. Alamir, and J. P. Corriou (2003). Nonlinear Receding-Horizon State Estimation for Dispersive Adsorption Columns with Nonlinear Isotherm. *J. Process. Contr.* **13,** 517-523.

M. Alamir, N. Sheibat-Othman, and S. Othman (2008). On the use of Nonlinear Receding-Horizon Observers in Batch Terpolymerization Processes. *International Journal of Modelling, Identification and Control.* **4**(1), 79-88.

A. Alessandri, M. Baglietto, and G. Battistelli (2008). Moving-Horizon State Estimation for Nonlinear Discrete-Time Systems: New Stability Results and Approximation Schemes. *Automatica,* **44,** 1753-1765.

F. Allgöwer, T. A. Badgwell, J. S. Qin, J. B. Rawlings and S. J. Wright (1999). Nonlinear Predictive Control and Moving Horizon Estimation - An Introductory Overview. Proceedings of the Advances in Control - Highlights of ECC '09 (pp: 391-449). London, Springer.

L. T. Biegler, and V. Zabala (2009). Large-Scale Nonlinear Programming using IPOPT: An Integrating Framework for Enterprise-wide Dynamic Optimization. *Comp. and Chem. Eng,* **33,** 575-582.

T. Binder, L. Blank, W. Dahmen, and W. Marquardt (2005). An Adaptive Multiscale Method for Real-Time Moving Horizon Optimization. *Proc. of the ACC,* Chicago, USA.

B-S. Chen, and H-C. Lu (1988). State estimation of large-scale systems. *Proc. of Inter. Jour. of Cont,* **47**(6), 1613-1632.

C. F. Cheng (1997). Design of Robust Observation Schemes for Uncertain Large Scale Systems. *IEEE Proc. Cont. Theory,* **144**(5), 369-374.

P. Feiler, R. P. Gabriel, J. Goodenough, R. Linger, T. Longstaff, R. Kazman, M. Klein, L. Northrop, D. Schmidt, K. Sullivan, K. Wallnaw (2006). *Ultra-Large-Scale Systems: the Software Challenge of the Future.* Carnellie Mellon University.

K. Findeisen (1997). Moving Horizon State Estimation of Discrete-Time Systems. Master's Thesis, University of Wisconsin-Madison, USA.

V. Gupta (2006). Distributed Estimation. Class Notes. Available on-line at March 3rd in: http://www.eeci-institute.eu/pdf/M7-Textes/19_03_08/L9_distributed.pdf

N. Haverbeke, T. Van Herpe, M. Diehl, G. Van den Berghe, and B. de Moor (2008). 'Nonlinear Model Predictive Control with Moving Horizon State and Disturbance Estimation- Application to the Normalization of Blood Glucose in the Critically Ill'. Proc. of the 17th World IFAC Congress, Seoul, Korea.

J. D. Hedengren, and K.V. Allsford (2007). Moving Horizon Estimation and Control for an Industrial Gas Phase Polymerization Reactor. *Proc. of the ACC,* New York City, USA.

S. S. Jang, B. Joseph, and H. Mukai (1986). Comparison of two Approaches to Online Parameter and State Estimation of Nonlinear Systems. *Ind. Eng. Chem. Process Des. Dev.* **25,** 809-814.

J. B. Jørgensen, J. B. Rawlings, and S. B. Jørgensen (2004). Numerical Methods for Large Scale Moving Horizon Estimation and Control. *DYCOPS 7, 7th International Symposium on Dynamics and Control of Process Systems,* Massachussets, USA.

S. J. Julier, and J. K. Uhlmann (2004). Uncented Filtering and Nonlinear Estimation. *Proc. of IEEE.* **93**(3), 401-422.

W. Kang (2006). Moving Horizon Numerical Observers of Nonlinear Control Systems. *IEEE Trans. on Auto. Cont.* **51**(2), 344-349.

S. K. Katti (1982). Comments on: Decentralization, Stabilization, and Estimation of Large-Scale Linear Systems. *IEEE Trans. on Auto. Cont.* **27**(3) 747-748.

M. Kawohl, T. Heine, and R. King (2007). Model Based Estimation and Optimal Control of fed-batch Fermantation Processes for the Production of antibiotics. *Chemical Engineering and Processing,* 46, 1223-1241.

T. Kraus, P. Külh, L. Wirsching, H. G. Bock, and M. Diehl (2006). A Moving Horizon State Estimation algorithm applied to the Tennessee Eastman Benchmark Process. *IEEE Intern. Conf. on Multisensor Fusion and Integration for Intelligent Systems.* Heidelberg, Germany.

A. J Krener, and A. Isidori (1983). Linearization by Output Injection and Nonlinear Observers. *Syst. Cont. Letters.* **3**(47).

A. Küpper, M. Diehl, J. P. Schlöder, H. G- Bock, and S. Engell (2008). Efficient Moving Horizon State and Parameter Estimation for SMB Processes. Submitted for publication in *J. Process Contr*ol.

K.V. Ling and K. W. Ling (1999). Receding Horizon Recursive State Estimation. *IEEE Trans. on Auto.Cont.* **44**(9), 1750-1753.

M. L. Luyben, B. D. Tyreus, and W. L. Luyben (1997). Plantwide Control Design Procedure. *AIChE J.*, **43,** 3161-3174.

D. Q. Mayne, and H. Michalska (1992). Moving Horizon Observer-Based Control. *Proc. of the 31 st Conf. on Decis. and Cont*, Tucson, Arizona.

S. S. Mahil, and S. Bommaraju (1992). Decentralized Estimators for a Large Scale Dynamic System. *IEEE Inter. Conf. on Systems, Man and Cybernetics,* Chicago, USA.

H. Michalska, and D. Q. Mayne (1995). Moving Horizon Observers and Observer-Based Control. *IEEE Trans. on Auto. Cont.* **40**(6), 995-1006.

A. Mjaavatten, and B. A. Foss (1997). A Modular System for Estimation and Diagnosis. *Computers Chem. Eng.* **21**(11), 1203-1218.

A. G. O. Mutambara (1998). *Decentralized Estimation and Control for Multisensor Systems.* BocaRaton, FL: CRC Press.

C. C. Qu, and J. Hahn (2008). Computation of Arrival Cost for Moving Horizon Estimation via Unscented Kalman Filter. *J. Process Control.*

T. Raïssi, N. Ramdani, and Y. Candau (2005). Bounded Error Moving Horizon State Estimation for Nonlinear continuous-time Systems: Application to a Bioproccess System. *J. Process. Contr,* **15,** 537-545.

C.V. Rao (2000). Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems. Doctoral Dissertation, University of Wisconsin-Madison, USA.

C. V. Rao and J. B. Rawlings (2000). Nonlinear Moving Horizon State Estimation. In Progress in Systems and Control Theory, Vol 26 (pp: 45-69). Birkhäuser Verlag.

C. V. Rao, J. B. Rawlings, and J. H. Lee (2001). Constrained Linear State Estimation -A Moving Horizon Approach. *Automatica.* **37**, 1619-1628.

C. V. Rao, J. B. Rawlings, and D. Q. Mayne (2003). Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. *IEEE Trans. on Auto. Cont.* **48**(2), 246-257.

C. W. Sanders, E. C. Tacker, and T. D. Linton (1974). A New Class of Decentralized Filters for Interconnected Systems. *IEEE Trans. on Auto. Cont.* **AC-19**, 259-262.

C.W. Sanders, E.C. Tacker, T. D. Linton, and R.Y.-S-Ling (1978). Specific Structures for Large-Scale State Estimation Algorithms Having Information Exchange. *IEEE Trans. on Auto. Cont.* **23**(2), 255-261.

S. Samar, and D. Gorinevsky (2006). Model Predictive Estimation of Evolving Faults. *Proc. of the ACC,* Minneapolis, Minnesota USA.

D.D.Siljak, and M.B. Vukcevic (1976). Decentralization, Stabilization, and Estimation of Large-Scale Linear Systems. *IEEE Trans. on Auto. Cont.* 363-366.

R. Vadigepalli, and F. J. Doyle III (2003). Structural Analysis of Large-Scale Systems for Distributed State Estimation and Control Applications. *Cont. Eng. Practice.* **11**, 895-905.

R. Vadigepalli, and F. J. Doyle III (2003). A Distributed State Estimation and Control Algorithm for Plantwide Processes. *IEEE Trans. on Cont. Sys. Tech.* **11**(1), 119-127.

A. Wätcher, and L. T. Biegler (2006). On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Program.* **106** (1), 25-57.

V. M. Zabala, C. D. Laird, and L. T. Biegler (2007). A fast Computational Framework for Large-Scale Moving Horizon Estimation. *Proceedings of the 8th International Symposium on Dynamics and Control of Process Systems*, Cancun, Mexico.

V. Zabala, C. D. Laird, and L. T. Biegler (2008). A fast Moving Horizon Estimation Algorithm Based on Nonlinear Programming Sensivity. Submitted for publication in *J. Process Control.*