Double Moving Horizon Estimation: Linearization by a Nonlinear Transformation

Mohammad Abdollahpouri*[1], Mark Haring[2], Tor Arne Johansen[2], Gergely Takács[1] and Boris Rohal'-Ilkiv[1]

Abstract— **Moving horizon estimation (MHE) is a constrained non-convex optimization problem in principle, which needs to be solved online. One approach to avoid dealing with several local minima is to linearize the nonlinear dynamics. This type of convex approximation usually utilizes the estimated state as a linearization trajectory, providing no guarantees of stability and optimality in general. In this paper, we study the cascade of a linear and linearized observer, which is called double MHE. The first stage makes use of a model transformation, that in the nominal case is globally equivalent to the nonlinear dynamics. Since this approach does not consider the input and output disturbances optimally, the second stage uses the first stage estimates as an external signal for linearizing the nonlinear dynamics to improve the quality of estimation. The overall configuration can be transformed into two quadratic programs. This approach not only avoids solving a non-convex optimization problem, but also reduces the computational complexity significantly compared to the one needed for solving a non-convex problem. This estimation method has been validated in a simulation study, where our approach converged to the global minimum without the need to explicitly solve a non-convex optimization problem.**

I. Introduction

Moving horizon estimation (MHE) khar has emerged as a powerful nonlinear state estimation technique that surpasses typical recursive-based methods [1] such as the extended Kalman filter [2] and the unscented Kalman filter [3]; see also [4], [5], [6] for comparison results. Most model based control designs require a precise estimation method in practice. This challenging problem either utilizes the nonlinear dynamics explicitly under some type of stability or boundedness consideration, or the linearized model is employed in a sub-optimal estimation design; see [7].

The authors gratefully acknowledge the contribution of the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 607957 (TEMPO); the financial contribution of the Slovak Research and Development Agency (APVV) under the contract APVV-14-0399; the support of the Scientific Grant Agency (VEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the contract 1/0144/15; the Research Council of Norway, Statoil, DNV GL and Sintef through the Centers of Excellence funding scheme, project number 223254 - Centre for Autonomous Marine Operations and Systems (NTNU-AMOS); and grant 250725; and the support of EEA Grant no. SK06-II-01-004

*Mohammad Abdollahpouri is currently affiliated with the Chalmers University of Technology in Göteborg, Sweden. mohammad.abdollahpouri@chalmers.se

[1]Institute of Automation, Measurement and Applied Informatics, Faculty of Mechanical Engineering, Slovak University of Technology in Bratislava, 812 31 Bratislava, Slovakia.

[2]Center for Autonomous Marine Operations and Systems (NTNU-AMOS), Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7491 Trondheim, Norway.

The MHE can consider the nonlinear dynamics as well as the constraints explicitly in an optimization framework; for an introduction see [8], [9]. Constraints on the states and parameters can help the estimation method improve the quality of estimates, which can further provide a better control performance. Constraints on the unmeasured input and output disturbances are usually used to model truncated normal densities [10]. The main challenge in the MHE, similar to any nonlinear optimization problem is, to obtain the global solution in a time that is real-time tractable on digital computers.

There are many approaches that approximate the nonlinear dynamic optimization problem by a convex one. Many of these approaches express the dynamics—mostly by linearization—hence obtaining the global solution is difficult without having a reliable linearization path available. Recently Johansen et al. utilized a model transformation that is globally equivalent to the nonlinear dynamics with no input and output disturbances in a multi stage estimation technique [11]. The resulting model is linear time-varying (LTV), where its time-varying parameters depend on a sequence of the calculated/measured input and output. This transformation considers the input and output disturbances sub-optimally, hence the transformed system may not completely reflect the stochastic model. Utilizing this transformed model in an estimation framework that is designed optimally for a class of disturbances, might end up with biased estimates.

This model transformation has been employed in a two stage estimation method called double Kalman filtering (D-KF) [11] and its stability analysis in discrete-time has been studied in [12]. The first stage of the DKF utilizes the transformed LTV model and provides globally convergent estimates. The sub-optimal estimation might not be precise, therefore the second stage linearizes the nonlinear dynamics about the estimates provided by the first stage. The results are expected to have an improved precision compared to a classical extended Kalman filter (EKF). Since EKF uses its own estimates as linearization point, it is probable that its estimates converge to a wrong value if one of the following scenarios happen: if the estimation algorithm is provided with an initial guess that is far from the true initial value, or if the system dynamics changes abruptly so the previous estimates provide an inaccurate linearization path for EKF, or if there exist multiple solutions for a specific series of measurements [5].

This paper introduces the so-called double MHE cascade configuration that is based on the idea presented in

[11]. Although MHE may in many practical cases improve estimation precision, increase robustness to non-Gaussian noise and sensor malfunctions by its window of moving measurements and provide state constraints (e.g. see [6]), this comes with a price of increased computational complexity and implementation difficulties on digital computers having limited computing power. The method proposed in this paper aims not only to preserve the advantages of MHE over Kalman-based methods, but also reduce its computational complexity significantly.

The proposed method consists of two stages of linear MHE—a convexified and a linearized MHE—in a cascade configuration, where the result of the first stage provides a linearization trajectory for the second stage. These two linear MHE problems can be translated to two individual quadratic programs (QP), where they can be easily implemented using conventional QP solvers. Since the first stage utilizes the globally equivalent LTV model in the nominal case, its solution can be different than the global minimum in the presence of input and output disturbances. Therefore, the second stage uses the first stage estimates and improves the estimation performance. Using a typical nonlinear solver for the MHE problem, e.g. a gradient based method, renders the estimation technique dependent on the initial guess. A warm start method for solving a nonlinear MHE is similar in the structure to the internal feedback of EKF, which makes the method sensitive to sudden changes in dynamics or a wrong initial guess. Hence, the performance may be diminished and in some situations results in a sub-optimal solution that is quite far from the global minimum. However, the approach proposed in this paper avoids these weaknesses by solving two consequent QPs, where in the nominal case the first QP is convergent. Therefore, if the disturbances are assumed bounded—even though it is not guaranteed—this approach has an increased chance of obtaining a global solution compared to the case of solving a nonlinear MHE using a sub-optimal solver.

II. Problem Formulation

Nonlinear process dynamics is defined by

$$x_{k+1} = f(x_k, u_k) + w_k, \tag{1a}$$
$$y_k = h(x_k) + v_k, \tag{1b}$$

where $x_k \in \mathbb{X} \subseteq \mathbb{R}^n$, $u_k \in \mathbb{U} \subseteq \mathbb{R}^m$, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^p$ are the state vector, control signal, input and output disturbances, respectively. The nonlinear dynamics is denoted by $f(\cdot,\cdot) : \mathbb{X} \times \mathbb{U} \to \mathbb{X}$ and the nonlinear measurement function is expressed by $h(\cdot) : \mathbb{X} \to \mathbb{R}^p$. Furthermore, $y \in \mathbb{R}^p$ is the bounded observed output. The discrete time index is denoted by k.

Assumption 2.1: The sets $\mathbb{X}$ and $\mathbb{U}$ are compact with $0 \in \mathbb{X}$ and $0 \in \mathbb{U}$.

Assumption 2.2: There exists a map $\psi(\cdot,\cdot) : \mathbb{R}^{pN} \times \mathbb{R}^{m(N-1)} \to \mathbb{R}^n$, which is Lipschitz continuous in all of its arguments and a positive integer N such that the state of (1) for all $k \geq N-1$ can be written as

$$x_k = \psi\big((y_j)_{j=l}^{k}, (u_j)_{j=l}^{k-1}\big),$$

where $l = k - N + 1$.

A. Linear models

This subsection introduces two types of linear models that will be later utilized in the double moving horizon estimation framework. The first model is globally equivalent to the system in (1) with the disturbances removed; their output responses for any input is identical. The second model is the linearized version of (1) about an arbitrary linearization trajectory.

1) LTV model: Using Assumption 2.2, it is possible to transform the nonlinear system (1) into the following LTV model

$$x_{k+1} = f(0, u_k) + F_k x_k + \acute{w}_k, \tag{2a}$$
$$y_k = h(0) + H_k x_k + \acute{v}_k, \tag{2b}$$

where $F_k = F\big((y_j)_{j=l}^{k}, (u_j)_{j=l}^{k-1}\big)$ and $H_k = H\big((y_j)_{j=l}^{k}, (u_j)_{j=l}^{k-1}\big)$ are the time varying parameters that provide an equivalent model for the nonlinear dynamics (1) globally, assuming $w_k = 0$ and $v_k = 0$. The disturbances can be reformulated as

$$\acute{w}_k = w_k + \acute{w}\big((y_j)_{j=l}^{k}, (u_j)_{j=l}^{k-1}, (v_j)_{j=l}^{k}\big),$$
$$\acute{v}_k = \acute{v}\big((y_j)_{j=l}^{k}, (u_j)_{j=l}^{k-1}, (v_j)_{j=l}^{k}\big).$$

For more information on this approach readers are encouraged to consult [11], [12].

Remark 2.1: The main challenge in utilizing the transformed LTV model (2) is the possibility of transforming the characteristics of input and output disturbances. Since most of the optimal estimation methods make some assumptions on the disturbances, those might not hold after the transformation. Therefore, different tuning is necessary if the estimation technique is utilizing the LTV model, since it is quite hard to achieve an analytical expression for the transformed disturbances.

The following example describes how the transformed model can be obtained, and illustrates the effect of this transformation on the disturbances following the discussion in Remark 2.1.

Example 2.1: Let a scalar system with nonlinear dynamics and measurement function in discrete time be described as

$$x_{k+1} = x_k^3 + x_k^2 u_k + w_k,$$
$$y_k = x_k^2 + v_k,$$

where $x_k \in \mathbb{R}^+$ and one can immediately calculate the state as $x_k^2 = y_k - v_k \to x_k = \sqrt{y_k - v_k}$, which can further simplify the nonlinear dynamics as

$$x_{k+1} = (y_k - v_k)x_k + (y_k - v_k)u_k + w_k,$$
$$y_k = x_k\sqrt{y_k - v_k} + v_k.$$

Furthermore, the LTV model can be simplified as

$$x_{k+1} = y_k x_k + y_k u_k + \acute{w}_k := F_k x_k + G_k u_k + \acute{w}_k,$$
$$y_k = \sqrt{y_k} x_k + \acute{v}_k := H_k x_k + \acute{v}_k,$$

where the transformed input disturbance is denoted as $\acute{w}_k = w_k - v_k u_k - v_k\sqrt{y_k - v_k}$ and the transformed output disturbance is $\acute{v}_k = y_k - \sqrt{y_k(y_k - v_k)}$. Note that $\acute{v}_k$ and $\acute{w}_k$ are zero if w_k and v_k are zero.

Definition 1: A LTV system of the form (2) is uniformly completely observable (UCO) if there exist constants $c_1, c_2 \in \mathbb{R}^+$ and a positive integer N such that

$$c_1 I \le \sum_{i=0}^{N-1} \|H_{k+i}\Phi(k+i,k)\|^2 \le c_2 I$$

for all $k \ge 0$, where the transition matrix is defined as

$$\Phi(k+i,k) = F_{k+i-1}F_{k+i-2}\cdots F_k,$$

for $0 \le i \le N-1$, while $\Phi(k,k) = I$.

Assumption 2.3: The LTV system (2) is UCO.

2) Linearized model: The linearization of (1) about an arbitrary state $\bar{x}_k$, results in the following linear system

$$x_{k+1} = f(\bar{x}_k, u_k) + A_k(x_i - \bar{x}_k) + \mathcal{Q}_k + w_k, \tag{3a}$$
$$y_k = h(\bar{x}_k) + C_k(x_k - \bar{x}_k) + \mathcal{R}_k + v_k, \tag{3b}$$

where the matrices A_k and C_k are defined as

$$A_k = \frac{\partial f}{\partial x_k}(\bar{x}_k, u_k),\ C_k = \frac{\partial h}{\partial x_k}(\bar{x}_k)$$

and the higher order terms are denoted by $\mathcal{Q}_k$ and $\mathcal{R}_k$. Their boundedness properties depend directly on the behavior of $\bar{x}_k$. From [12, Prop. 4.2.], denoting $\tilde{x} := \bar{x}_k - x_k$ gives that there exist constants ϵ_q and ϵ_r, such that

$$\|\mathcal{Q}_k\| \le \epsilon_q \|\tilde{x}_k\|^2, \quad \|\mathcal{R}_k\| \le \epsilon_r \|\tilde{x}_k\|^2, \quad \forall k \ge 0.$$

Therefore if $\bar{x}_k$ is provided by an estimator with converging estimation error, the higher order terms in (3) vanish asymptotically.

Assumption 2.4: The linearized system (3) is UCO.

Assumption 2.5: The evolution of the linear/linearized systems in (2) or (3) respects the constraints; i.e. from any initial state and for any possible sequence of input disturbances $(w_j)_{j=0}^{k-1} \in \mathbb{W}^k$, the state satisfies the constraint; $x_k \in \mathbb{X}$.

III. Linear Moving Horizon Estimation

In this paper we formulate the linear and linearized MHE (LMHE) utilizing the linear dynamics in (3). Note that this formulation is also valid for the system in (2). This estimation method is the solution to the following constrained optimal control problem in discrete time

$$\text{LMHE}: \quad \Theta_k^* = \min_{\hat{x}_l, (\xi_j)_{j=l}^{k-1}} \Theta_k\big(\hat{x}_l, (\xi_j)_{j=l}^{k-1}\big) \tag{4a}$$
$$\text{s.t.} \quad \hat{x}_{i+1} = A_i\hat{x}_i + b_i + \xi_i, \quad i = l, \ldots, k-1, \tag{4b}$$
$$y_i = C_i\hat{x}_i + d_i + \nu_i, \quad i = l, \ldots, k, \tag{4c}$$
$$\hat{x}_i \in \mathbb{X}, \quad \xi_i \in \mathbb{W}, \quad i = l, \ldots, k, \tag{4d}$$

where the cost function is defined by

$$\Theta_k\big(\hat{x}_l, (\xi_j)_{j=l}^{k-1}\big) := \Theta_l^* + \|\hat{x}_l - s_l\|_{P_l}^2 + \sum_{i=l}^{k-1}\|\xi_i\|_{Q_i}^2 + \sum_{i=l}^{k}\|\nu_i\|_{R_i}^2$$

where $Q_i = Q_i^{\mathrm{T}} \succ 0$ and $R_i = R_i^{\mathrm{T}} \succ 0$ are the tuning parameters with the appropriate dimension. The arrival cost term is defined by $P_l = P_l^{\mathrm{T}} \succ 0$ and s_l, serves to approximate the effect of the missing data outside the moving horizon; see [8]. Note that b_i and d_i are the time-varying terms that can be easily calculated either from (3) or (2), depending on the model that is used in the LMHE formulation.

The main contribution of this work is to reduce the computational complexity of the estimation method compared to the one needed to solve an NMHE, and to increase the possibility of finding a global minimum. Therefore the linear MHE is transformed to a QP, where compared to utilizing a nonlinear programming solver, the complexity of the problem will decrease significantly. The following reformulation can be used to efficiently implement the method in computing hardware such as embedded micro-controllers, using the existing software tools for solving constrained QP [13], [14]

$$\text{LMHE}: \quad \min_{\mathcal{X}_k} \frac{1}{2}\mathcal{X}_k^{\mathrm{T}}\mathcal{H}_k\mathcal{X}_k + \mathcal{F}_k^{\mathrm{T}}\mathcal{X}_k + \mathcal{E}_k \tag{5a}$$
$$\text{s.t.} \quad \mathcal{X}_k \in \hat{\mathbb{X}}, \tag{5b}$$

where $\hat{\mathbb{X}} = \mathbb{X}^N$, $\mathcal{H}_k \in \mathbb{R}^{nN\times nN}$, $\mathcal{F}_k \in \mathbb{R}^{nN}$ and $\mathcal{E}_k \in \mathbb{R}$ can be obtained as follows

$$\begin{aligned}
\mathcal{X}_k &= \begin{bmatrix}\hat{x}_l^{\mathrm{T}} & \hat{x}_{l+1}^{\mathrm{T}} & \hat{x}_{l+2}^{\mathrm{T}} & \ldots & \hat{x}_{k-1}^{\mathrm{T}} & \hat{x}_k^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}},\\
\mathcal{H}_k(1,1) &= P_l + A_l^{\mathrm{T}}Q_lA_l + C_l^{\mathrm{T}}R_lC_l,\\
\mathcal{H}_k(r+1,r+1) &= \\
&\quad Q_{l+r-1} + A_{l+r}^{\mathrm{T}}Q_{l+r}A_{l+r} + C_{l+r}^{\mathrm{T}}R_{l+r}C_{l+r},\\
\mathcal{H}_k(N,N) &= Q_{k-1} + C_k^{\mathrm{T}}R_kC_k,\\
\mathcal{H}_k(q,q+1) &= -A_{l+q-1}^{\mathrm{T}}Q_{l+q-1},\\
\mathcal{H}_k(q+1,q) &= -Q_{l+q-1}A_{l+q-1},\\
\mathcal{F}_k(1) &= 2A_l^{\mathrm{T}}Q_lb_l - 2C_l^{\mathrm{T}}R_l(y_l - d_l) - 2P_ls_l,\\
\mathcal{F}_k(r+1) &= 2A_{l+r}^{\mathrm{T}}Q_{l+r}b_{l+r} - \\
&\quad 2C_{l+r}^{\mathrm{T}}R_{l+r}(y_{l+r} - d_{l+r}) - 2Q_{l+r-1}b_{l+r-1},\\
\mathcal{F}_k(N) &= -2C_k^{\mathrm{T}}R_k(y_k - d_k) - 2Q_{k-1}b_{k-1},\\
\mathcal{E}_k &= \|s_l\|_{P_l}^2 + \sum_{i=l}^{k-1}\|b_i\|_{Q_i}^2 + \sum_{i=l}^{k}\|y_i - d_i\|_{R_i}^2
\end{aligned}$$

for all $r = 1, 2, \ldots, N-2$ and $q = 1, 2, \ldots, N-1$. It should be noted that $\mathcal{H}_k$ is a block matrix and $\mathcal{H}_k(r,c) \in \mathbb{R}^{n\times n}$ is the sub-matrix associated with the $(rn-n+1)^{\text{th}}$ to rn^{th} row and $(cn-n+1)^{\text{th}}$ to cn^{th} column of the matrix $\mathcal{H}_k$. Furthermore, $\mathcal{F}_k(j) \in \mathbb{R}^n$ denotes j^{th} block of the length n of the vector $\mathcal{F}_k$ for $1 \le j \le N$.

IV. Double MHE

We propose a cascaded linear and linearized MHE, that we shall call double MHE (DMHE), summarized in Fig. 1. In the first stage, the linear MHE formulation in (4) utilizes the system model in (2). As it has been previously mentioned, this model is nominally and globally equivalent to the nonlinear dynamics, hence in the absence of disturbances provides a global minimum to the nonlinear optimization problem of MHE. Let us refer to this problem as the convexified MHE (CMHE). Since in reality, disturbances influence the system dynamics, the result of CMHE might be different from the global minimum, therefore the second stage of DMHE will utilize the result of CMHE to improve the quality of estimation. This stage (LMHE) makes use of the linearized model (3), considering the linearization trajectory provided by CMHE. This stage will also be initialized by the CMHE estimates, so the possibility of obtaining a global minimum shall increase. The details of this estimation method is given in Algorithm 1.

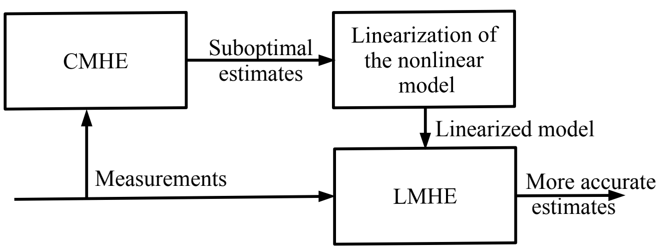


Fig. 1. Schematic overview of the proposed double MHE

Algorithm 1 Implementation of DMHE

Initialization: $\S_0$

Input: $(y_j)_{j=l}^{k}$, $(u_j)_{j=l}^{k}$

Update the LTV model (2)

Update the parameters in (5):

$b_i \leftarrow (\bar{x}_i, u_i) - A_i(\bar{x}_i)$

$d_i \leftarrow h(\bar{x}_i) - C_i\bar{x}_i$

Solve (5) with the updated parameters and obtain $\bar{x}_k$

Update the linearized model (3)

Update the parameters in (5):

$b_i \leftarrow f(0, u_i)$

$d_i \leftarrow h(0)$

$A_i \leftarrow F_i$

$C_i \rightarrow H_i$

Solve (5) with the updated parameters and update $\hat{x}_k$

Output: $\hat{x}_k$

To illustrate the method, let us assume that the MHE problem is non-convex with one suboptimal solution and one global minimum as shown in Fig. 2. The solution of CMHE is approximately $x = 0.8$, which is in the neighborhood of the global minimum; $x = 1$. Utilizing the solution of CMHE for linearization makes the DMHE method to converge to the global solution as the arrows in Fig. 2 demonstrate the effect of the two stage estimation approach in DMHE.

V. Nominal stability analysis

Under nominal conditions—without considering the input and output disturbances—the first stage is globally equivalent with the nonlinear optimization problem. Since the linear

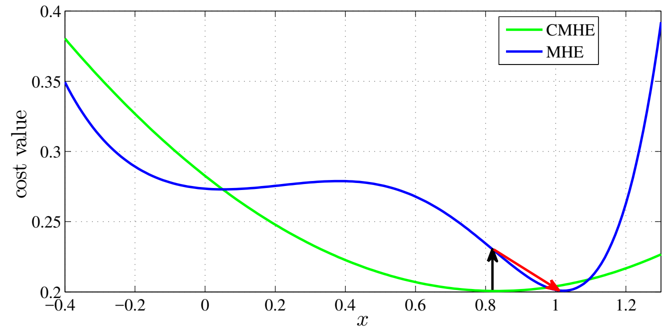


Fig. 2. Comparison of convexified and non-convex MHE cost functions; black and red arrows now represent the first and second stages of the DMHE, which gets an improved solution from its initial guess.

MHE is transformed to a QP, it is globally convergent. In the second stage the main difference is due to the terms b_i and d_i in (4), which depend on the higher order terms, $\mathcal{Q}_k$ and $\mathcal{R}_k$. In this case, stability analysis would be possible for the MHE that employs the linear time-varying model, if $\bar{x}_k \rightarrow x_k$ in the first stage.

In this paper, we are not analyzing the stability of this cascade configuration. Readers are encouraged to see [12] for a detailed stability analysis of a cascade design of two uniformly globally asymptotically stable Kalman filters, followed by the analogy that was proposed in [15].

VI. Numerical example

Consider the following scalar nonlinear system

$$x_{k+1} = -20T_s x_k^3 + 10T_s x_k^2 + x_k + T_s, \tag{6a}$$

$$y_k = x_k + v_k, \tag{6b}$$

where the sampling time $T_s = 0.01$ s, and the additive measurement noise is assumed to be uniformly distributed with a variance of 0.1. Compared to a normally distributed disturbance, a uniform noise has a non-zero mean, and is considered as a biased output disturbance. This bias in the disturbance deteriorate the validity of the transformed model (2). This issue can be observed in the biased estimates of the CMHE. To overcome this, the DMHE configuration can provide better estimates by linearizing the nonlinear dynamics about the biased solution and result in a smaller estimation error; see Fig. 3(a).

The LTV model that will be utilized in CMHE is as follows

$$\begin{aligned} x_{k+1} = &\\ &(-20T_s(y_k - v_k)^2 + 10T_s(y_k - v_k) + 1)x_k + T_s u_k,\\ &= (-20T_s y_k^2 + 10T_s y_k + 1)x_k + T_s u_k + \acute{w}_k, \end{aligned}$$

where u_k is selected to have three constant values during each second, $\acute{w}_k$ as the new input disturbance summarizes the model transformation effect on disturbances. Since v_k is assumed to be uniformly distributed noise with mean value of 0.05 and its transformation in the dynamics is expressed by $\acute{w}_k$, it might lead to biased estimation for the CMHE. To improve the quality of estimation, the second stage employs the linearized model that is expressed by

$$x_{k+1} = (-20T_s\bar{x}_k^3 + 10T_s\bar{x}_k^2 + \bar{x}_k) + (-60T_s\bar{x}_k^2 + 20T_s\bar{x}_k + 1)(x_k - \bar{x}_k) + T_s u_k.$$

For both optimization problems the horizon is $N = 10$. The tuning parameters (see the optimization problem in (4)) for the first stage are chosen as $Q = 10^{-2}$, $R = 1$ and $P = 0.01$. The second stage uses the following parameters for its tuning: $Q = 10^{-4}$, $R = 10^{-5}$ and $P = 0.1$. Note the difference in selecting Q for CMHE and LMHE. Even though the nonlinear and linearized dynamics are assumed to have no input disturbance, the associated tuning variable can not be zero. This tuning parameter is responsible for incorporating the state update in the optimization problem. The biased measurement is not only effective in the measurement error part of the CMHE cost function, it has some influence on the state update part as well. Hence R is chosen smaller for LMHE than for the one in CMHE to overcome the negative influence of the measurement bias in the cost function of LMHE problem. The NMHE problem uses the same tuning parameter as in LMHE, since the stochastic model of both linearized and nonlinear model is similar. It should be declared that the linear and linearized MHE problems were solved using the "quadprog" MATLAB function, while the nonlinear MHE problem was solved by "fmincon" MATLAB function, and all the simulations were executed on an Ubuntu 16.04 desktop computer (Intel i7, 2.3 GHz, 16 GB RAM).

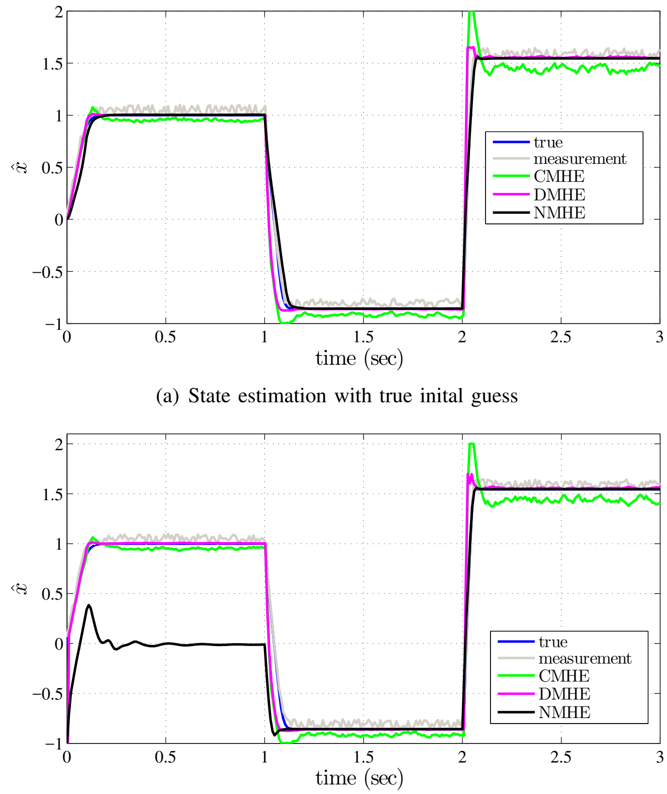
The main issue with a typical NMHE utilizing its own estimates for linearization and initialization is the negative effect of a poor initial guess as demonstrated in Fig. 3(b). Note that since the initialization ($\hat{x}_0 = -1$) is slightly different from the true value ($x_0 = 0$), the estimated state for the NMHE converged to the sub-optimal solution in the beginning of the test; c.f. Fig. 2, where the local minimum is located at around 0. On the other hand, it is illustrated that if the initialization is correct, all the methods have the same chance of getting to the correct estimates; see Fig. 3(a). Although the results of CMHE is biased because of the assumed colored measurement noise, the DMHE approach is able to bring the solution to the correct value. The interesting observation here is that the DMHE method, no matter how the initialization is selected, can converge to a small neighborhood of the global minimum. See Table I for a comparison study on the estimation performance defined as the sum of the absolute value of the estimation error and the execution time. Furthermore, the timing analysis demonstrates the advantage of our method compared to utilizing a nonlinear solver. In these simulations, the following constraint have been considered to improve the estimation performance

$$-1 \le \hat{x}_i \le 2, \quad i = l, \dots, k,$$

where this choice of constraint is derived from our knowledge of the steady-state value of this stable system.

TABLE I
PERFORMANCE COMPARISON FOR THE NUMERICAL EXAMPLE

estimation method	estimation error with true $\hat{x}_0$	estimation error with wrong $\hat{x}_0$	execution time (s)
NMHE	9.02	107.58	1.3
CMHE	28.49	31.25	0.009
DMHE	9.33	11.59	0.02



(a) State estimation with true inital guess

(b) State estimation with wrong inital guess

Fig. 3. Comparison of different linear MHE methods.

VII. CONCLUSION

A cascade configuration of two linear MHE problems has been proposed. In the first stage, a nominally globally equivalent LTV model has been employed that is not optimally designed for systems with input and output disturbances. Although the first stage MHE problem becomes convex, its solution might be different from the global minimum due to the effect of disturbances that are sub-optimally accounted for. Even though global convergency is not guaranteed in the presence of disturbances, the second stage utilizes the result of the first stage as linearization trajectory to improve its solution. The stability analysis of the cascaded linear MHEs are proposed as a future work. A numerical example provides some illustrations on the improved estimation results obtained by the double MHE method compared to nonlinear MHE. Not only is the increased performance appealing, but the reduced computational complexity compared to the nonlinear MHE is interesting for implementation on embedded hardware.

REFERENCES

[1] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[2] A. Gelb, *Applied optimal estimation*. MIT press, 1974.

[3] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000, pp. 153–158.

[4] S. C. Patwardhan, S. Narasimhan, P. Jagadeesan, B. Gopaluni, and S. L. Shah, "Nonlinear bayesian state estimation: A review of recent developments," *Control Engineering Practice*, vol. 20, no. 10, pp. 933–953, 2012.

[5] E. L. Haseltine and J. B. Rawlings, "Critical evaluation of extended Kalman filtering and moving-horizon estimation," *Industrial & engineering chemistry research*, vol. 44, no. 8, pp. 2451–2460, 2005.

[6] M. Abdollahpouri, G. Takács, and B. Rohal'-Ilkiv, "Real-time moving horizon estimation for a vibrating active cantilever," *Mechanical Systems and Signal Processing*, vol. 86, pp. 1–15, 2017.

[7] D. Simon, *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. Wiley–Interscience, 2006.

[8] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, vol. 48, pp. 246–258, 2003.

[9] F. Allgöwer, T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright, "Nonlinear predictive control and moving horizon estimation-an introductory overview," in *Advances in control*. Springer, 1999, pp. 391–449.

[10] D. G. Robertson, J. H. Lee, and J. B. Rawlings, "A moving horizon-based approach for least-squares estimation," *AIChE Journal*, vol. 42, no. 8, pp. 2209–2224, 1996.

[11] T. A. Johansen and T. I. Fossen, "Nonlinear filtering with eXogenous Kalman filter and double Kalman filter," in *European Control Conference*, 2016, pp. 1722–1727.

[12] M. Abdollahpouri, M. Haring, T. A. Johansen, G. Takács, and B. Rohal'-Ilkiv, "Nonlinear state and parameter estimation using discrete-time double Kalman filter," in *The 20th World Congress of the IFAC*, vol. 50, no. 1, 2017, pp. 11 632 – 11 638.

[13] H. J. Ferreau, H. G. Bock, and M. Diehl, "An online active set strategy to overcome the limitations of explicit MPC," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 8, pp. 816–830, 2008.

[14] D. K. M. Kufoalor, B. Binder, H. J. Ferreau, L. Imsland, T. A. Johansen, and M. Diehl, "Automatic deployment of industrial embedded model predictive control using qpOASES," in *European Control Conference*. IEEE, 2015, pp. 2601–2608.

[15] T. A. Johansen and T. I. Fossen, "The eXogenous Kalman filter (XKF)," *International Journal of Control*, pp. 1–7, 2016.