



Brief paper

Minimum variance constrained estimator[☆]Prabhat K. Mishra^{*}, Girish Chowdhary, Prashant G. Mehta

Coordinated Science Laboratory, University of Illinois at Urbana Champaign (UIUC), USA

ARTICLE INFO

Article history:

Received 14 January 2021
 Received in revised form 27 August 2021
 Accepted 24 October 2021
 Available online 6 January 2022

Keywords:

Constrained estimation
 MHE
 Kalman filter
 Minimum variance duality

ABSTRACT

This paper is concerned with the problem of state estimation for discrete-time linear systems in the presence of additional (equality or inequality) constraints on the state (or estimate). By use of the minimum variance duality, the estimation problem is converted into an optimal control problem. Two algorithmic solutions are described: the full information estimator (FIE) and the moving horizon estimator (MHE). The main result is to show that the proposed estimator is stable in the sense of an observer. The proposed algorithm is distinct from the standard algorithm for constrained state estimation based upon the use of the minimum energy duality. The two are compared numerically on the benchmark batch reactor process model.

© 2021 Elsevier Ltd. All rights reserved.

1. Introduction

In many practical estimation problems arising in control applications, there are invariably additional constraints on the state process (Goodwin, Seron, & De, 2006). In such applications, Kalman filter (KF) may yield sub-optimal estimates that violate the constraints. It is notable also that the KF is derived under the assumption of (unbounded) Gaussian noise, which is also unrealistic in the constrained settings of the problem. In particular, in the presence of unbounded noise, local stability results are not applicable and global stability results are very conservative due to actuator saturation (Chatterjee, Ramponi, Hokayem, & Lygeros, 2012; Mishra, Chatterjee, & Quevedo, 2017). Although clever modifications in KF are still possible (Yang & Blasch, 2009), the stability and optimality properties of such modifications require further investigation (Simon, 2010). For these reasons, constrained estimation is a problem of paramount practical importance; c.f., (Goodwin et al., 2006) for a book length treatment.

A popular strategy for constrained estimation is based on the use of duality between estimation and optimal control. A practical advantage of converting a constrained estimation problem into a constrained optimal control problem is that model predictive control (MPC) methods, algorithms, and softwares can readily be

applied to obtain a solution. The resulting estimation algorithms are referred to as the *full information estimator* (FIE), when all the observations are used, and *moving horizon estimator* (MHE), when a moving window of most recent past observations are used. Practically, a MHE algorithm is preferred because the number of decision variables in the optimization problem do not increase as more observations are collected.

In linear settings of the problem, there are in general two types of duality: the minimum energy (or maximum likelihood) duality and the minimum variance duality. Refer to Pavn and Wets (1982), Pearson (1966) and Todorov (2008) for more discussion on duality. For the construction of estimators, the minimum energy duality is by far the more popular technique with contributions in Alessandri, Baglietto, and Battistelli (2010), Alessandri and Gaggero (2020), Brembeck (2019), Copp and Hespanha (2017), Farina, Ferrari-Trecate, and Scattolini (2010), Flayac (2019) and Schneider, Hannemann-Tamás, and Marquardt (2015) and numerical algorithms in Haverbeke (2011) and Morabito, Kögel, Bullinger, Pannocchia, and Findeisen (2015). Although minimum variance control has attracted much attention (Bakolas, 2018a, 2018b; Makkapati, Rajpurohit, Okamoto, & Tsiotras, 2020), and these recent papers provide motivation also for our work, the use of minimum variance duality for constrained estimation has received comparatively less attention.

State estimation problem for linear systems with equality constraints is considered in Ko and Bitmead (2007) and Teixeira, Chandrasekar, Torres, Aguirre, and Bernstein (2007), and with inequality constraints in Liew (1976). Since the number of decision variables in the underlying optimization program increases as more measurements are collected, an MHE algorithm is proposed in the early work of Jazwinski (1968) in the absence of constraints. This algorithm is extended in Muske, Rawlings,

[☆] This work was supported in part by Navy, USA N00014-19-1-2373 and National Science Foundation, USA 1739874. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Ali Zemouche under the direction of Editor Torsten Söderström.

^{*} Corresponding author.

E-mail addresses: pmishra@illinois.edu (P.K. Mishra), girishc@illinois.edu (G. Chowdhary), mehtapg@illinois.edu (P.G. Mehta).

and Lee (1993) to incorporate constraints. The stability properties of this constrained MHE algorithm are studied rigorously in Rao, Rawlings, and Lee (2001). Enhancements of these basic algorithms have been considered both in deterministic (observer design) (Alessandri, Baglietto, & Battistelli, 2003; Gharbi & Ebenbauer, 2018; Sui & Johansen, 2014; Sui, Johansen, & Feng, 2010) and stochastic (filter design) (Jazwinski, 1968; Kwon, Kim, & Han, 2002; Muske et al., 1993; Rao et al., 2001) settings of the problem. In stochastic settings, the MHE optimal control problem is still deterministic but statistical information about uncertainties and prior are used to design the (maximum likelihood-type) objective function. More recent extensions include a game theoretic formulation in Garcia T., Marquez-Ruiz, Botero C., and Angulo (2018). It is noted that Garcia T. et al. (2018), Jazwinski (1968), Kwon et al. (2002), Muske et al. (1993) and Rao et al. (2001) are based on the use of the minimum energy duality. We refer readers to Rao (2000, Appendix B) for a quick review of duality and to Mortensen (1968) for minimum energy duality in particular.

In this paper, an alternate form of duality, viz., the minimum variance duality is employed to transform the minimum variance estimation problem into a deterministic optimal control problem. The state estimate is constructed as a linear function of past measurements. Without constraints, the optimal estimate is equivalent to a Kalman filter. Both the FIE and the MHE are described for the unconstrained case, together with expression for choosing the terminal cost in the MHE.

The main focus of this paper is on the modification of these (unconstrained) FIE and MHE algorithms in the presence of constraints. In particular, a certain approximate expression for the terminal cost is introduced for the constrained MHE. The main result of this paper is to establish sufficient conditions to obtain stability (in the sense of an observer) for the constrained FIE and MHE algorithms. Furthermore, we also establish a certain type of stochastic stability by showing that the variance of the constrained FIE converges under certain technical conditions.

Although estimators based on minimum variance duality are less well studied (Kim, Mehta, & Meyn, 2019), some closely related estimators have appeared in Darouach and Zasadzinski (1997), Darouach, Zasadzinski, and Boutayeb (2003), Kwon, Han, Kwon, and Kwon (2007), Kwon et al. (2002), Kwon, Kim, and Park (1999) and Zhao, Shmaliy, Huang, and Liu (2015). In contrast to our paper, these prior works do not incorporate equality or inequality constraint on state (or estimate) in the estimator design. The original contributions of our paper are as follows:

- Based on minimum variance duality, a MHE is presented in (12) and its equivalence with FIE is shown in Lemma 2. This contribution is different from Kwon et al. (2002) and Kwon et al. (2007) in the sense that unbiasedness constraints are not required. The proposed estimator (13) is equivalent to KF.
- Constrained FIE and MHE algorithms are presented in (16) and (19), respectively. Apart from the fact that these algorithms are distinct from Rao et al. (2001), our minimum variance-based approach has certain technical advantages.
- Although the notion of stability is borrowed from Rao et al. (2001), Theorems 1 and 2 are first such results on stability of constrained minimum variance estimators.
- Under certain technical conditions, the variance of constrained FIE is shown to converge in Theorem 3.

The remainder of this paper is organized as follows: The problem statement appears in Section 2 followed by a description of the minimum variance duality for the construction of the unconstrained estimators, both FIE and MHE, in Section 3.1. These are extended to the constrained case in Section 3.2. The main results on stability of the constrained FIE and MHE appear in Section 4. The algorithms are illustrated with the aid of some

numerical experiments in Section 5. The paper closes with some conclusions and directions for future research in Section 6. All the proofs appear as part of the two appendices, Appendices A and B, for the unconstrained and the constrained cases, respectively.

Let $\mathbb{R}$, $\mathbb{N}_0$, $\mathbb{Z}_+$ denote the set of real numbers, the non-negative integers and the positive integers, respectively. We use the symbols $\mathbf{0}$ and I to denote zero matrix and identity matrix, respectively, of appropriate dimensions. For any vector or matrix sequence $(M_n)_{n \in \mathbb{N}_0} \in \mathbb{R}^{r \times m}$, $r, m \in \mathbb{Z}_+$, let $M_{n:k} \in \mathbb{R}^{r \times m}$ denote the matrix $[M_n^T \ M_{n+1}^T \ \dots \ M_{n+k-1}^T]^T$, $k \in \mathbb{Z}_+$. Let $\lambda_{\max}(M)$ denote the largest eigenvalue value of M , $\lambda_{\min}(M)$ its smallest eigenvalue, $M^\dagger$ its Moore–Penrose pseudo inverse and $\text{tr}(M)$ its trace. The Euclidean norm of a vector A is denoted by $|A|$. The Frobenius norm of a matrix A is denoted by $|A|_F$. A t step reachability matrix of a matrix pair (A, B) is given by $R_t(A, B) := [A^{t-1}B \ \dots \ AB \ B]$.

2. Problem statement

Consider a linear discrete-time system

$$\begin{aligned} x_{t+1} &= Ax_t + w_t, \\ y_t &= Cx_t + \varsigma_t, \end{aligned} \quad (1)$$

where $x_t \in \bar{\mathcal{X}} \subset \mathbb{R}^d$, $y_t \in \mathbb{R}^q$ are state and measurement of the system at time t , respectively. The system matrix $A \neq \mathbf{0}$. The additive process noise w_t and the measurement noise ς_t are mean zero, mutually independent and identically distributed random vectors with variance Q and R , respectively. The initial state of the system x_0 is a random vector with mean $\hat{x}_0^-$ and variance Σ_0^- , and is independent of the process noise and the measurement noise.

The minimum variance estimation problem is to compute $\hat{x}_t$ at time t such that the variance of error $x_t - \hat{x}_t$ is minimized over some class of admissible estimators. In this paper, the admissible estimators are assumed to be linear deterministic functions of available measurements. It is also assumed that some additional insight into the states (or estimates) is given in terms of equality and inequality constraints such that the estimated states belong to a convex set $\mathcal{X} \supseteq \bar{\mathcal{X}}$, i. e. $\hat{x}_t \in \mathcal{X}$ for all t . We make the following assumption:

Assumption 1. The set $\mathcal{X}$ is positively invariant under the nominal dynamics, i. e. $Ax \in \mathcal{X}$ for every $x \in \mathcal{X}$.

The above assumption is meaningful. Suppose $A\hat{x}_t \notin \mathcal{X}$ for some $\hat{x}_t \in \mathcal{X}$ then there is a non-zero probability that $x_{t+1} \notin \bar{\mathcal{X}}$ for random w_t , e.g., when $\hat{x}_t = x_t$ and a bounded disturbance set with known bounds is not safely prescribed. The optimization problem is as follows:

$$\min_{\hat{x}_t \in \mathcal{X}} \mathbb{E} \left[|x_t - \hat{x}_t|^2 \right]. \quad (2)$$

The solution approach is based on duality between estimation and control. In the following section, we begin by presenting an unconstrained estimator which is useful for the development of a constrained estimator in Section 3.2.

3. Minimum variance estimators

3.1. Unconstrained estimator

In this section, we assume $\mathcal{X} = \mathbb{R}^d$, i.e., the constraints are not present. We are interested in an estimator linearly parameterized in the innovation terms as follows:

$$\hat{x}_t = A^t \hat{x}_0^- - \sum_{i=0}^t \alpha_i^\top (y_{t-i} - CA^{t-i} \hat{x}_0^-), \quad (3)$$

in which weights $\alpha_i \in \mathbb{R}^{q \times d}$ are the decision variables for the optimization problem (2). In order to convert the minimum variance estimation objective into an optimal control problem, a dual process (in forward time) is introduced:

$$\begin{aligned} z_{i+1} &= A^\top z_i + C^\top \alpha_{i+1}; \quad i = 0, \dots, t-1, \\ z_0 &= I + C^\top \alpha_0, \end{aligned} \quad (4)$$

where $z_i \in \mathbb{R}^{d \times d}$ is a matrix valued dual state and $\alpha_i \in \mathbb{R}^{q \times d}$ is control signal for the dual process. From (4) we have

$$z_t^\top = A^t + \sum_{i=0}^t \alpha_i^\top C A^{t-i}. \quad (5)$$

By substituting (5) into (3), we get the following expression:

$$\hat{x}_t = z_t^\top \hat{x}_0^- - \sum_{i=0}^t \alpha_i^\top y_{t-i}. \quad (6)$$

A slight modification of the standard result on minimum variance duality (Åström, 1970, Page 238),¹ in which only the past measurements are used to design an estimator, i. e. $\alpha_t = \mathbf{0}$, is required to include the current measurement. Let $\ell_i := z_i^\top Q z_i + \alpha_i^\top R \alpha_i$, $\Gamma_0(z_t) := z_t^\top \Sigma_0^- z_t$ and

$$S_t(\alpha_{0:t+1}) := \alpha_t^\top R \alpha_t + \sum_{i=0}^{t-1} \ell_i. \quad (7)$$

The estimate (3) takes into account all measurements available at time t . Therefore, the corresponding estimator is called *full information estimator* (FIE). Using the dual process (4), the FIE optimal control problem is expressed as follows:

$$\text{FIE: } \begin{cases} \text{minimize} & \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})) \\ \text{subject to} & \text{dual dynamics (4)}. \end{cases} \quad (8)$$

FIE (8) is solved at each time $t = 0, 1, \dots$. The resulting optimal solution is denoted as $\alpha_{0:t+1|t}$, where $\alpha_{k|t}$ is the optimal weight α_k computed at time t . Set

$$\Sigma_t := \Gamma_0(z_{t|t}) + S_t(\alpha_{0:t+1|t}), \quad (9)$$

where $S_t(\alpha_{0:t+1|t})$ is the optimal value of $S_t(\alpha_{0:t+1})$ obtained by solving FIE (8). Then the optimal value of the objective function in (8) is $\text{tr}(\Sigma_t)$. The estimate $\hat{x}_{t|t}$ at time t is obtained by substituting the optimal values $\alpha_{0:t+1|t}$ and $z_{t|t}$ in (6). In the remainder of the manuscript, we will use $\hat{x}_t$ to denote the estimate obtained by substituting the optimizers in (6). We have the following Lemma to show the equivalence of FIE (2) and (8) whenever $\mathcal{X} = \mathbb{R}^d$.

Lemma 1. Consider the system (1) and the dual process (4). If $\hat{x}_t$ is given by (6) then

$$\mathbb{E} \left[|x_t - \hat{x}_t|^2 \right] = \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})).$$

Remark 1. The dual process is typically considered backward in time. However, because the optimal control problem is deterministic, a forward time dual process may equivalently be considered simply by renaming the indices. This is done here to yield the standard form of an optimal control problem where the time arrow is forward.

We present a finite horizon approximation of FIE (8), which we refer to as *moving horizon estimator* (MHE). For this purpose, define

$$\Sigma_t^- := A \Sigma_{t-1} A^\top + Q \text{ and } \hat{x}_t^- := A \hat{x}_{t-1}. \quad (10)$$

¹ See Åström (1970, Exercise 1, Page 240) in which invertibility of the system matrix A is assumed to define a dual process.

Fix $N \in \mathbb{N}_0$ and for $t \geq N+1$ define

$$\Gamma_{t-N}(z_N) := z_N^\top \Sigma_{t-N}^- z_N. \quad (11)$$

The unconstrained MHE is as follows:

$$\text{MHE: } \begin{cases} \text{minimize} & \text{tr}(\Gamma_{t-N}(z_N) + S_N(\alpha_{0:N+1})) \\ \text{subject to} & \text{dual dynamics (4)}. \end{cases} \quad (12)$$

For $t \leq N$, set $\Gamma_{t-N} = \Gamma_0$, $S_N = S_t$, which is identical to solving the FIE problem (8). For $t \geq N+1$, the MHE problem utilizes the most recent $N+1$ measurements together with the previously computed Σ_{t-N-1} to obtain Σ_{t-N}^- . The resulting estimator and the error covariance matrix are

$$\hat{x}_t = z_{N|t}^\top \hat{x}_{t-N}^- - \sum_{i=0}^N \alpha_{i|t}^\top y_{t-i}, \quad (13)$$

$$\Sigma_t = \text{tr}(\Gamma_{t-N}(z_{N|t}) + S_N(\alpha_{0:N+1|t})), \quad (14)$$

where $\alpha_{i|t}$ for $i = 0, \dots, N$, and $z_{N|t}$ are obtained by solving MHE (12) at time t . It is straightforward to show that, when $N = 0$, MHE (12) is the KF. A direct implication of dynamic programming is the following result:

Lemma 2. If $R > 0$ then FIE (8) is equivalent to MHE (12) and the estimate (6) is equal to the estimate (13).

Proofs of Lemmas 1 and 2 are given in Appendix A.

3.2. Constrained estimator

If the matrix pair (A, C) is observable then there exists an integer $n \leq d \in \mathbb{Z}_+$ such that $\text{rank}(R_n(A^\top, C^\top)) = d$. The smallest such n is referred to as the observability index of (A, C) . Our construction of the constrained FIE depends on n . In particular, we augment the FIE (8) with the following additional constraints:

$$z_{t-j}^\top \hat{x}_0^- - \sum_{i=0}^{t-j} \alpha_i^\top y_{t-j-i} \in \mathcal{X}, \quad (15)$$

where $j = 0$ for $t \leq n$ and $j = 0, \dots, t-n$, for $t \geq n+1$. Note that the left hand side of the constraint is same as $\hat{x}_{t-j}$ according to (6). Although we are interested in this constraint only with $j = 0$, inclusion of the intermediate constraints, for $j = 1, \dots, t-n$, helps to ensure some properties. Additional details on this appear in the next section. The constrained FIE problem is formally defined as follows:

$$\text{CFIE: } \begin{cases} \text{minimize} & \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})) \\ \text{subject to} & \text{dual dynamics (4),} \\ & \text{constraints (15)}. \end{cases} \quad (16)$$

The solution of the CFIE (16) is used to construct the constrained full information estimate by using the right hand side of (6). It is denoted $\hat{x}_t^{\text{cf}}$ to distinguish it from unconstrained estimate $\hat{x}_t$ obtained by solving (8) or (12). In particular,

$$\hat{x}_t^{\text{cf}} := z_{t|t}^\top \hat{x}_0^- - \sum_{i=0}^t \alpha_{i|t}^\top y_{t-i}, \quad (17)$$

$$\Sigma_t^{\text{cf}} := \Gamma_0(z_{t|t}) + S_t(\alpha_{0:t+1|t}),$$

where $z_{t|t}$ and $\alpha_{0:t+1|t}$ are obtained by solving (16).

Remark 2 (Feasibility and Convexity). If $\hat{x}_0^- \in \mathcal{X}$ then the optimal control problem (16) is feasible for all t because $\alpha_{0:t+1} = \mathbf{0}$ satisfies (15). The left hand side of (15) is affine in decision variables $\alpha_{0:t+1}$ and the set $\mathcal{X}$ is convex. The set of decision

variables $\alpha_{0:t+1}$ in (15) is convex due to the fact that the inverse image of a convex set under an affine function is convex (Boyd & Vandenberghe, 2004, Page 38) and the intersection of convex sets is convex.

Remark 3. The right hand side of (6) is linear in the past measurements. The justification comes from the unconstrained linear Gaussian case where such a structure is sufficient to obtain the minimum variance estimator. In the presence of constraints and non-Gaussian noise, an optimal estimate may not be linear in the past measurements. It is noted that the assumed structure is also nonlinear because of the dependence of $\alpha_{0:t+1}$ on $y_{0:t+1}$ via constraint (15).

In the presence of constraints, the design of an MHE algorithm, that is provably equivalent to the FIE algorithm, is challenging because of the difficulty in approximating the terminal cost. Therefore, approximation of the terminal cost (which is also referred to as *arrival cost* in the standard MHE literature) is necessary. The goal is to approximate the FIE as closely as possible while maintaining computational tractability and guaranteeing stability.

Similar to CFIE (16), constrained MHE can also be defined by adding extra constraints to the unconstrained MHE (12). The constrained MHE estimator is denoted as $\hat{x}_t^{\text{cm}}$, where the superscript cm is used to reflect the fact that this estimate at time t may be different from the unconstrained estimate $\hat{x}_t$ and the CFIE estimate $\hat{x}_t^{\text{cf}}$. Similarly, the corresponding error covariance matrix is denoted by Σ_t^{cm} to distinguish it from (14).

We need to define priors $\Sigma_{t-N}^{\text{cm-}}$ and $\hat{x}_{t-N}^{\text{cm-}}$ to compute the terminal cost of the constrained MHE and its estimate as we did in (11) and (13), respectively, for the unconstrained case. One possible choice is to use Σ_{t-N}^- and $\hat{x}_{t-N}^-$ obtained from the unconstrained case by using (10) and (12), which is same as running a KF in parallel. The standard MHE (Rao et al., 2001) follows this approach. Other MHE approaches like Sui and Johansen (2014) also use priors from the unconstrained case. Since our approach not only gives an estimated state which satisfies constraints but also an error covariance matrix, it is reasonable to replace Σ_{t-1} in (10) by Σ_{t-1}^{cm} to get $\Sigma_{t-1}^{\text{cm-}}$ and $\hat{x}_{t-1}$ by $\hat{x}_{t-1}^{\text{cm}}$ to get $\hat{x}_{t-1}^{\text{cm-}}$. This choice is intuitive because the pair $(\hat{x}_{t-1}^{\text{cm-}}, \Sigma_{t-1}^{\text{cm-}})$ represents our prior knowledge about the pair $(\hat{x}_t^{\text{cm}}, \Sigma_t^{\text{cm}})$ in the presence of constraints. More precisely,

$$\Sigma_t^{\text{cm-}} := A \Sigma_{t-1}^{\text{cm}} A^\top + Q \text{ and } \hat{x}_t^{\text{cm-}} := A \hat{x}_{t-1}^{\text{cm}}. \quad (18)$$

The constrained MHE problem is formally written as follows:

$$\text{CMHE: } \begin{cases} \text{minimize} & \text{tr}(\Gamma_{t-N}^{\text{cm}}(z_N) + S_N(\alpha_{0:N+1})) \\ \text{subject to} & \text{dual dynamics (4), and} \\ & z_N^\top \hat{x}_{t-N}^{\text{cm-}} - \sum_{i=0}^N \alpha_i^\top y_{t-i} \in \mathcal{X}, \end{cases} \quad (19)$$

where $\Gamma_{t-N}^{\text{cm}}(z_N) := z_N^\top \Sigma_{t-N}^{\text{cm-}} z_N$ for $t \geq N+1$. Similar to MHE (12) for $t \leq N$, we set $\Gamma_{t-N}^{\text{cm}} = \Gamma_0$, $S_N = S_t$ and similarly modify constraint by taking all $t+1$ measurements. Alternatively, we can run CFIE (16) for $t \leq N$. Further, the estimate (6) and corresponding covariance matrix can be written as

$$\hat{x}_t^{\text{cm}} := z_{N|t}^\top \hat{x}_{t-N}^{\text{cm-}} - \sum_{i=0}^N \alpha_{i|t}^\top y_{t-i}, \quad (20)$$

$$\Sigma_t^{\text{cm}} := \Gamma_{t-N}^{\text{cm}}(z_{N|t}) + S_N(\alpha_{0:N+1|t}),$$

where $\alpha_{i|t}$ for $i = 0, \dots, N$, and $z_{N|t}$, are obtained by solving CMHE problem (19) at time t .

4. Main results

In this section, stability of the proposed constrained estimators is presented by using the notion of stability introduced in Rao et al. (2001). Recall that the classical notion of stability of an observer is obtained by modifying the definition of the stability of a regulator. In an analogous manner, the definition of the stability of a constrained regulator, which is given in Keerthi and Gilbert (1988, §2), is modified in Rao et al. (2001) to introduce the following definition:

Definition 1 (Keerthi & Gilbert, 1988; Rao et al., 2001). The estimator is a stable observer for the system

$$x_{t+1} = Ax_t; \quad y_t = Cx_t; \quad x_t \in \mathcal{X}, \quad (21)$$

if for any $\varepsilon > 0$, there exists $\delta > 0$ and $T \in \mathbb{Z}_+$ such that if $\hat{x}_0^- \in \mathcal{X}$ and $|x_0 - \hat{x}_0^-| \leq \delta$ then $|\hat{x}_t - A^t x_0| \leq \varepsilon$ for all $t \geq T$. If in addition, $\hat{x}_t \rightarrow A^t x_0$ as $t \rightarrow \infty$ then the estimator is called asymptotically stable observer for the system (21).

Our approach has a minor advantage over (Rao et al., 2001) in the sense that a key assumption is relaxed. In particular, we do not assume any upper bound on cost a priori but it comes naturally from the observability of the system. For the stability of CFIE we need one of the following two conditions to hold:

(C1) $Q - \Sigma_0^- \geq 0$.

(C2) There exists some $K_t \in \mathbb{R}^{q \times d}$ at each time $t \geq n+1$ such that $\alpha_t = K_t z_{t-1|t-1}$ satisfies (15) for $j=0$ and the following stability criterion with $\tilde{A}_t = A^\top + C^\top K_t$:

$$\tilde{A}_t^\top \Sigma_0^- \tilde{A}_t - \Sigma_0^- \leq -(K_t^\top R K_t + Q). \quad (22)$$

The main stability result for the CFIE is as follows:

Theorem 1. Suppose Assumption 1 holds, $|\hat{x}_0^- - x_0| < \infty$, $\Sigma_0^- > 0$, (A, C) is observable, and one of the two conditions, either (C1) or (C2), is satisfied. Then CFIE is an asymptotically stable observer for the system (21).

Remark 4. It is easily verified that the conditions (C1) and (C2) cannot simultaneously hold unless $A = 0$, which, because the matrix pair (A, C) is observable, represents a trivially false case when C is not a full column rank matrix. Let, if possible, (C1) and (C2) hold simultaneously then (C2) gives

$$0 \leq \tilde{A}_t^\top \Sigma_0^- \tilde{A}_t \leq -(K_t^\top R K_t + Q - \Sigma_0^-) \leq 0, \quad (23)$$

which implies $\tilde{A}_t^\top \Sigma_0^- \tilde{A}_t = K_t^\top R K_t + Q - \Sigma_0^- = 0$. Therefore, $\tilde{A}_t = 0$ because $\Sigma_0^- > 0$ and $K_t^\top R K_t + Q = \Sigma_0^-$, which results in $Q \leq \Sigma_0^-$ and due to (C1) we get $Q = \Sigma_0^-$. By substituting $Q = \Sigma_0^-$ in (23), we get $K_t^\top R K_t = 0$, which results in $K_t = 0$ because $R > 0$. Since $\tilde{A}_t = 0$ due to (23), the substitution of $K_t = 0$ shows that $A = 0$.

We have the following result on stability of CMHE:

Theorem 2. Suppose Assumption 1 holds, $\Sigma_0^- > 0$, $R > 0$ and (A, C) is observable then for $N \geq n$, CMHE is stable observer for the system (21). If, in addition, $Q > 0$, $|\hat{x}_0^- - x_0| < \infty$, then CMHE is asymptotically stable observer for the system (21).

In Theorems 1 and 2, we proved stability of the proposed estimators in the sense of an observer. Since the cost function represents variance in the proposed approach, we get its convergence for the system (1) also under the following assumption:

Assumption 2. There exist $\alpha_0 \in \mathbb{R}^{q \times d}$, and a sequence of matrices $(K_i)_{i \in \mathbb{Z}_+}$ such that $\alpha_{i+1} = K_i z_i$ and α_0 satisfy (15). There exist $\lambda_0 > 0$, $\lambda_i < 1$ for $i \in \mathbb{Z}_+$ such that

$$(C3) (I + C^T \alpha_0)^T Q (I + C^T \alpha_0) + \alpha_0^T R \alpha_0 \leq \lambda_0 Q$$

$$(C4) (A^T + C^T K_i)^T Q (A^T + C^T K_i) + K_i^T R K_i \leq \lambda_i Q \text{ for } i \in \mathbb{Z}_+$$

The above assumption gives a sufficient condition for the feasibility of (16) and the existence of a stabilizing controller for the dual process (4). Notice that (16) is feasible due to Remark 2. The above assumption helps us to get an upper bound of the cost in (16). We have the following result:

Theorem 3. If $\hat{x}_0^- \in \mathcal{X}$, (A, C) is observable and for all $t \geq n + 1$ either (C1) with Assumption 2 holds or (C2) is satisfied, then there exists $s' \geq 0$ such that

$$\mathbb{E} \left[|x_t - \hat{x}_t^{cf}|^2 \right] \rightarrow s'. \quad (24)$$

Proofs of Theorems 1–3 are given in Appendix B.

5. Numerical experiments

For numerical experiments, we consider the benchmark model of a well-mixed, constant volume, isothermal batch reactor. This model has previously been considered in Haseltine and Rawlings (2005) and Sui and Johansen (2014). The system dynamics is given by (1), where

$$A = \begin{bmatrix} 0.8831 & 0.0078 & 0.0022 \\ 0.1150 & 0.9563 & 0.0028 \\ 0.1178 & 0.0102 & 0.9954 \end{bmatrix}, \quad C = [32.84 \quad 32.84 \quad 32.84]$$

The observability index of (A, C) is 3. The additive process and measurement noise are both assumed to be Gaussian with zero means, and variances, $(0.01)^2 I$ and $(0.25)^2$, respectively. The mean of the initial prior is $\hat{x}_0^- = [1 \quad 1 \quad 4]^T$. Since the states represent concentration of chemicals in the batch reactor process, these cannot be negative. Therefore, the estimated states are constrained to lie in the set $\mathcal{X} := \{x \in \mathbb{R}^d \mid x \geq 0\}$.

Experiment 1. In the first experiment, we assume that initial state is also Gaussian with prior mean $\hat{x}_0^-$ and prior variance $\Sigma_0^- = I$. This is evident that simulated state of the system can be negative due to the presence of Gaussian noises in simulation but we consider this example for a fair comparison with minimum energy MHE (MEMHE) (Rao et al., 2001).

We demonstrate a comparison between MEMHE and our proposed approach CMHE in Fig. 1. MEMHE is simulated by using nmhe object of freely available MATLAB based software package mpctools (Risbeck & Rawlings, 2016), which is based on CasAdi (Andersson, Gillis, Horn, Rawlings, & Diehl, 2019) and solver Ipopt (Wächter & Biegler, 2006). For CMHE, we use MATLAB-based software package YALMIP (Löfberg, 2004) and a solver SDPT3-4.0 (Toh, Todd, & Tütüncü, 2012) to solve the underlying optimization programs. We chose the optimization horizon $N = 4$ for both approaches and simulated for $N_s = 1000$ sample paths. The empirical mean squared error e_t for both approaches is computed by the following formula:

$$e_t = \frac{1}{N_s} \sum_{i=1}^{N_s} |x_t^i - \hat{x}_t^i|^2, \quad (25)$$

where x_t^i and $\hat{x}_t^i$ denote the simulated and estimated states, respectively, at time t in the i th path.

Fig. 1 depicts that empirical mean squared error in our approach is smaller than that in MEMHE. Interestingly, at $t = 0$ both approaches have almost same e_t but in our approach it immediately drops by approximately one unit and keeps monotonically decreasing after then. However, in case of MEMHE a slight increase is observed at $t = 2$ and after that it monotonically decreases but always remains higher than that of our approach.

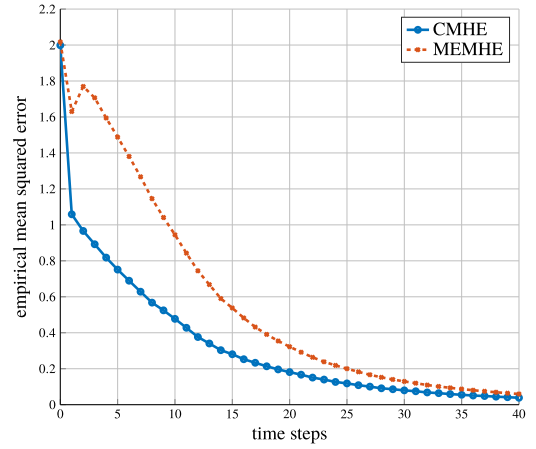


Fig. 1. The empirical mean squared error for 1000 sample paths is smaller in our proposed approach than that in standard MHE when initial state has Gaussian distribution.

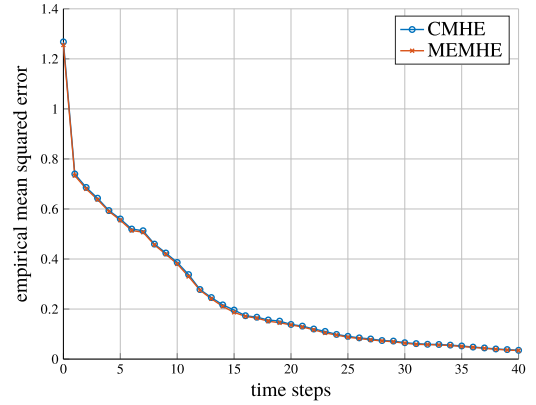


Fig. 2. The empirical mean squared error for 100 sample paths is almost same in our proposed approach and standard MHE when initial state has uniform distribution.

Experiment 2. In this experiment, we consider initial state to be uniformly distributed between $[0, 2\hat{x}_0^-]$. Rest of the simulation data is same as in Experiment 1. We simulate for $N_s = 100$ sample paths and compare between our proposed approach CMHE and standard MEMHE in Fig. 2. The empirical mean squared error e_t is computed according to (25). Fig. 2 depicts that both approaches have almost the same empirical mean squared error for 100 sample paths.

Experiment 3. In this experiment, we choose optimization horizon $N = 3$ and simulate only for one sample path. Rest of the simulation data is same as in Experiment 2. We compare the norm of estimate and cost by using CMHE and CFIE in Fig. 3. Both approaches give almost same estimate and incur almost same cost even though the optimization problem of CFIE has intermediate constraints, which are absent in CMHE.

6. Conclusions and directions for future research

In this paper, the minimum variance duality is used to convert the minimum variance estimation problem into a deterministic optimal control problem. The main contribution is the specification and the stability analysis of the FIE and MHE algorithms in the presence of state constraints. The proposed algorithms are distinct from and possess several useful features compared to

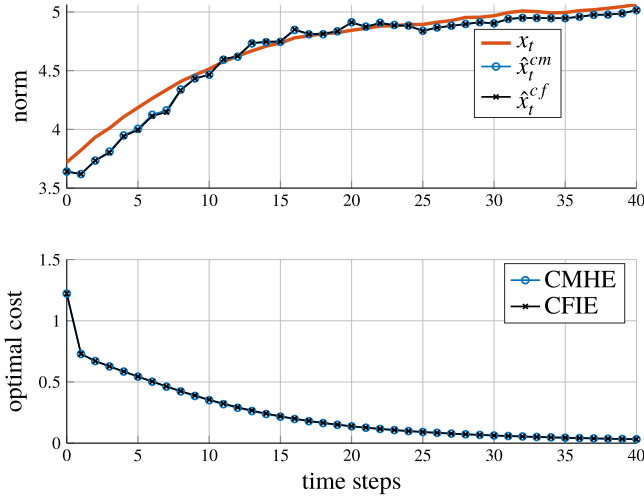


Fig. 3. Norm of estimate and optimal cost in both CMHE and CFIE are almost same.

the standard MHE algorithms based on the use of the minimum energy duality. In particular, there is no need to run a KF in parallel to approximate the terminal cost for the MHE. Both the constrained FIE and MHE algorithms are stable in the sense of an observer. Moreover, stochastic stability of constrained FIE is also established.

This work opens up several avenues for future research: Some ideas of linear model predictive control with time varying terminal cost and constraints (Pluymers, Roobrouck, Buijs, Suykens, & De Moor, 2005), and approximate dynamic programming methods with accumulating constraints (Bertsekas, 2005) may be useful for the further study of the constrained MHE. Several interesting extensions of the proposed approach may be possible including control design (Copp & Hespanha, 2017), systems with intermittent observations (Mishra, Chatterjee, & Quevedo, 2020), distributed architecture (Farina et al., 2010), the problem of unknown prior (Kong et al., 2020; Kwon et al., 2002) and inclusion of pre-estimating observer (Alessandri et al., 2003; Kong & Sukkarieh, 2018; Kwon et al., 2002).

Acknowledgment

The first author is thankful to Jin W. Kim for suggesting a useful reference.

Appendix A. Proofs of Section 3.1

Proof of Lemma 1. Since $z_0 - C^\top \alpha_0 = I$, we have

$$x_t = (z_0 - C^\top \alpha_0)^\top x_t = z_0^\top x_t - \alpha_0^\top C x_t. \quad (\text{A.1})$$

By using the system dynamics (1) and the dual dynamics (4), we get

$$\begin{aligned} z_i^\top x_{t-i} &= z_i^\top (A x_{t-i-1} + w_{t-i-1}) \\ z_{i+1}^\top x_{t-i-1} &= z_i^\top A x_{t-i-1} + \alpha_{i+1}^\top C x_{t-i-1}. \end{aligned} \quad (\text{A.2})$$

We substitute (A.2) in the expression of $z_0^\top x_t$ as follows:

$$\begin{aligned} z_0^\top x_t &= \sum_{i=0}^{t-1} (z_i^\top x_{t-i} - z_{i+1}^\top x_{t-i-1}) + z_t^\top x_0 \\ z_0^\top x_t &= \sum_{i=0}^{t-1} (z_i^\top w_{t-i-1} - \alpha_{i+1}^\top C x_{t-i-1}) + z_t^\top x_0. \end{aligned}$$

We further substitute $z_0^\top x_t$ in (A.1) to get

$$\begin{aligned} x_t &= \sum_{i=0}^{t-1} z_i^\top w_{t-i-1} - \sum_{i=0}^t \alpha_i^\top C x_{t-i} + z_t^\top x_0 \\ x_t &= \sum_{i=0}^{t-1} z_i^\top w_{t-i-1} - \sum_{i=0}^t \alpha_i^\top (y_{t-i} - \zeta_{t-i}) + z_t^\top x_0. \end{aligned}$$

Further, we consider the estimate (6) and compute $\mathbb{E}[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^\top]$ as follows:

$$\begin{aligned} x_t - \hat{x}_t &= z_t^\top (x_0 - \hat{x}_0^-) + \sum_{i=0}^{t-1} z_i^\top w_{t-i-1} + \sum_{i=0}^t \alpha_i^\top \zeta_{t-i} \\ \mathbb{E}[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^\top] &= z_t^\top \Sigma_0^- z_t + \sum_{i=0}^{t-1} z_i^\top Q z_i + \sum_{i=0}^t \alpha_i^\top R \alpha_i \\ &= z_t^\top \Sigma_0^- z_t + \alpha_t^\top R \alpha_t + \sum_{i=0}^{t-1} \ell_i = \Gamma_0(z_t) + S_t(\alpha_{0:t+1}), \end{aligned}$$

since the process noise, measurement noise and initial states are mutually independent. Therefore, $\mathbb{E}[\|x_t - \hat{x}_t\|^2] = \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1}))$, where z_t is obtained by (4) and $\hat{x}_t$ is given by (6). $\square$

Proof of Lemma 2. At $t = 0$, we compute

$$\begin{aligned} \Gamma_0(z_0) + S_0(\alpha_0) &= z_0^\top \Sigma_0^- z_0 + \alpha_0^\top R \alpha_0 \\ &= \Sigma_0^- + \alpha_0^\top (R + C \Sigma_0^- C^\top) \alpha_0 + \Sigma_0^- C^\top \alpha_0 + \alpha_0^\top C \Sigma_0^-. \end{aligned}$$

Since $\alpha_{0|0} = \arg \min \text{tr}(\Gamma_0(z_0) + S_0(\alpha_0)) = -(C \Sigma_0^- C^\top + R)^{-1} C \Sigma_0^-$, due to our convention (9) we obtain $\Sigma_0 =$

$$\Gamma_0(z_{0|0}) + S_0(\alpha_{0|0}) = \Sigma_0^- - \Sigma_0^- C^\top (C \Sigma_0^- C^\top + R)^{-1} C \Sigma_0^-. \quad (\text{A.3})$$

The FIE cost can be written as

$$\begin{aligned} \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})) &= \text{tr} \left(\Gamma_0(z_t) + \alpha_t^\top R \alpha_t + \sum_{i=0}^{t-1} \ell_i \right) \\ &= \text{tr} (z_t^\top \Sigma_0^- z_t + \alpha_t^\top R \alpha_t + z_{t-1}^\top Q z_{t-1} + S_{t-1}(\alpha_{0:t})) \end{aligned}$$

We substitute $z_t = A^\top z_{t-1} + C^\top \alpha_t$ in the above expression and the minimizer $\alpha_{t|t}^\top = -z_{t-1}^\top A^\top \Sigma_0^- C^\top (C \Sigma_0^- C^\top + R)^{-1}$. Further, by substituting Σ_0 from (A.3), we get

$$\begin{aligned} \text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})) &= \text{tr} (z_{t-1}^\top (A \Sigma_0 A^\top + Q) z_{t-1} + S_{t-1}(\alpha_{0:t})) \\ &= \text{tr} (z_{t-1}^\top \Sigma_1^- z_{t-1} + S_{t-1}(\alpha_{0:t})), \end{aligned} \quad (\text{A.4})$$

where the last equality is due to our definition (10). Therefore, $\Gamma_1(\cdot)$ can be written as $\Gamma_1(z_{t-1}) = z_{t-1}^\top \Sigma_1^- z_{t-1}$. The above expression of cost (A.4) at time $t = 1$ gives $\Sigma_1 = z_{0|1}^\top \Sigma_1^- z_{0|1} + S_0(\alpha_{0|1})$, where $\alpha_{0|1} = -(C \Sigma_1^- C^\top + R)^{-1} C \Sigma_1^-$. By repeating the above process $t - N$ times, we obtain

$$\text{tr}(\Gamma_0(z_t) + S_t(\alpha_{0:t+1})) = \text{tr} (z_N^\top \Sigma_{t-N}^- z_N + S_N(\alpha_{0:N+1})),$$

and therefore, we can define $\Gamma_{t-N}(z_N) = z_N^\top \Sigma_{t-N}^- z_N$. Now for $t \geq N > 0$, we consider the expression of $\hat{x}_t$:

$$\hat{x}_t = z_{t|t}^\top \hat{x}_0^- - \alpha_{t|t}^\top y_0 - \sum_{i=0}^{t-1} \alpha_{i|t}^\top y_{t-i}, \quad \text{where}$$

$$z_{t|t}^\top \hat{x}_0^- - \alpha_{t|t}^\top y_0 = z_{t-1|t}^\top A \hat{x}_0^- - \alpha_{t|t}^\top (y_0 - C \hat{x}_0^-).$$

By substituting $\alpha_{t|t}$ in the above expression, we get $z_{t|t}^\top \hat{x}_0^- - \alpha_{t|t}^\top y_0 = z_{t-1|t}^\top (A \hat{x}_0^- + A \Sigma_0^- C^\top (C \Sigma_0^- C^\top + R)^{-1} (y_0 - C \hat{x}_0^-)) = z_{t-1|t}^\top A \hat{x}_0^-$, which implies $\hat{x}_t = z_{t-1|t}^\top \hat{x}_1^- - \sum_{i=0}^{t-1} \alpha_{i|t}^\top y_{t-i}$, where $\hat{x}_1^- = A \hat{x}_0^-$. At $t = 1$, we can compute $\hat{x}_1^-$ from the above expression. By

repeating the above process $t - N$ times we obtain the desired expression (13). $\square$

Appendix B. Proofs of Section 3.2

Lemma 3. If (C1) holds then $\text{tr}(\Sigma_t^{\text{cf}}) \geq \text{tr}(\Sigma_{t-1}^{\text{cf}})$ for all $t \geq n + 1$.

Proof. Let us define $s_t^* := \text{tr}(\Sigma_t^{\text{cf}})$ for notational simplicity. The optimal cost at time t by substituting $S_t(\alpha_{0:t+1|t}) = \alpha_{t|t}^\top R \alpha_{t|t} + \sum_{i=0}^{t-1} \ell_{i|t}$ in (17) is given by

$$s_t^* = \text{tr} \left(\Gamma_0(z_{t|t}) + \alpha_{t|t}^\top R \alpha_{t|t} + \sum_{i=0}^{t-1} \ell_{i|t} \right), \quad (\text{B.1})$$

where $\ell_{i|t} := z_{i|t}^\top Q z_{i|t} + \alpha_{i|t}^\top R \alpha_{i|t}$. We can observe for all $t \geq n + 1$ that the constraints (15) at time $t - 1$ are same at time t for $j = 0, \dots, t - n - 1$. Therefore, $\alpha_{0:t|t}$, the first t number of decision variables computed at time t , is a feasible control sequence at time $t - 1$. Due to the optimality of $\alpha_{0:t|t-1}$ at time $t - 1$, we get the following inequality:

$$\begin{aligned} s_{t-1}^* &\leq \text{tr} \left(\Gamma_0(z_{t-1|t}) + \alpha_{t-1|t}^\top R \alpha_{t-1|t} + \sum_{i=0}^{t-2} \ell_{i|t} \right) \\ &= \text{tr} \left(z_{t-1|t}^\top (\Sigma_0^- - Q) z_{t-1|t} + \sum_{i=0}^{t-1} \ell_{i|t} \right) \\ &= \text{tr} (z_{t-1|t}^\top (\Sigma_0^- - Q) z_{t-1|t} - \Gamma_0(z_{t|t}) - \alpha_{t|t}^\top R \alpha_{t|t}) + s_t^*, \end{aligned}$$

where the last equality is obtained by substituting (B.1). Since $Q - \Sigma_0^- \geq 0$, for all $t \geq n + 1$ we get

$$s_t^* - s_{t-1}^* \geq \text{tr} (\Gamma_0(z_{t|t}) + \alpha_{t|t}^\top R \alpha_{t|t} + z_{t-1|t}^\top (Q - \Sigma_0^-) z_{t-1|t}) \geq 0. \quad \square$$

Lemma 4. If (C2) holds, then $\text{tr}(\Sigma_t^{\text{cf}}) \leq \text{tr}(\Sigma_{t-1}^{\text{cf}})$ for all $t \geq n + 1$.

Proof. We can observe that $\alpha_{0:t|t-1}$ satisfies (15) at time t for $j = 1, \dots, t - n$. We assumed that $\alpha_t = K_t z_{t-1|t-1}$ satisfies (15) for $j = 0$. Therefore, the control sequence $\alpha_{0:t|t-1}$ along with $\alpha_t = K_t z_{t-1|t-1}$ is a feasible control sequence at time t . We compute z_t by substituting $\alpha_{0:t|t-1}$ and α_t in (4), which gives us $z_t = A^\top z_{t-1|t-1} + C^\top K_t z_{t-1|t-1} = \tilde{A} z_{t-1|t-1}$. Now we recall the expression of the optimal cost $s_t^* := \text{tr}(\Sigma_t^{\text{cf}})$ from (B.1). The optimality of $\alpha_{0:t+1|t}$ in the presence of stability criterion (22) gives us

$$\begin{aligned} s_t^* &\leq \text{tr} \left(\Gamma_0(\tilde{A} z_{t-1|t-1}) + z_{t-1|t-1}^\top K_t^\top R K_t z_{t-1|t-1} + \sum_{i=0}^{t-1} \ell_{i|t-1} \right) \\ &= \text{tr} \left(z_{t-1|t-1}^\top (\tilde{A}^\top \Sigma_0^- \tilde{A} + K_t^\top R K_t + Q) z_{t-1|t-1} \right) \\ &\quad + \text{tr} \left(\alpha_{t-1|t-1}^\top R \alpha_{t-1|t-1} + \sum_{i=0}^{t-2} \ell_{i|t-1} \right) \\ &\leq \text{tr} \left(\Gamma_0(z_{t-1|t-1}) + \alpha_{t-1|t-1}^\top R \alpha_{t-1|t-1} + \sum_{i=0}^{t-2} \ell_{i|t-1} \right) = s_{t-1}^*. \quad \square \end{aligned}$$

Lemma 5. If Assumption 1 holds and the matrix pair (A, C) is observable then there exists $s > 0$ such that for the system (21),

$$\begin{aligned} \text{tr}(\Sigma_t^{\text{cf}}) &\leq s \text{ for all } t \text{ and} \\ \text{tr}(\Sigma_t^{\text{cm}}) &\leq s \text{ for all } N \geq n \text{ for all } t. \end{aligned}$$

Proof. Let us consider the expression of z_t at $t = n$ from (5). We can write it in compact form: $z_n = A^\top (I + C^\top \alpha_0) + R_n (A^\top, C^\top) \alpha_{1:n}$.

If we substitute

$$\alpha_{1:n} = -R_n (A^\top, C^\top)^\dagger (A^\top) (I + C^\top \alpha_0) \quad (\text{B.2})$$

in the above expression for some $\alpha_0 \in \mathbb{R}^{q \times d}$, we get $z_n = 0$. Now we consider the estimator (3) and the nominal system (21). By substituting $y_i = C A^i x_0$ and $x_t = A^t x_0$ for the system (21) in (3), we get

$$\begin{aligned} \hat{x}_t &= A^t (\hat{x}_0^- - x_0) + A^t x_0 + \sum_{i=0}^t \alpha_i^\top C A^{t-i} (\hat{x}_0^- - x_0) \\ &= x_t + \left(A^t + \sum_{i=0}^t \alpha_i^\top C A^{t-i} \right) (\hat{x}_0^- - x_0) \\ &= x_t + z_t^\top (\hat{x}_0^- - x_0), \end{aligned} \quad (\text{B.3})$$

where the last equality is due to (5). If we substitute $\alpha_{1:n}$ from (B.2) in the above expression at $t = n$, we get $\hat{x}_n = x_n \in \mathcal{X} \subseteq \mathcal{X}$ because $z_n = 0$ under (B.2). Therefore, (B.2) is feasible for (16) at $t = n$. Let us define

$$s^0 := \text{tr}(\Gamma_0(z_n) + S_n(\alpha_{0:n+1})), \quad (\text{B.4})$$

where z_n and S_n are obtained by applying the given policy (B.2).

For all $t \geq n$, define $\beta_{0:n+1} = \alpha_{0:n+1}$ and $\beta_i = 0$ for $i > n$. Under the policy $\beta_{0:t+1}$, we have $z_t = 0$ and therefore $\hat{x}_t = x_t$ for all $t \geq n$; this policy is feasible. Since $\text{tr}(\Gamma_0(z_t) + S_t(\beta_{0:t+1})) = s$, optimality of $\alpha_{0:t+1|t}$ gives $\text{tr}(\Sigma_t^{\text{cf}}) \leq s^0$ for all $t \geq n$. For each $t \leq n - 1$, $\text{tr}(\Sigma_t^{\text{cf}}) \leq \text{tr} \left(A^t \Sigma_0^- A^{t^\top} + \sum_{i=0}^{t-1} A^i Q A^{i^\top} \right)$ is bounded, where the inequality holds due to optimality of $\text{tr}(\Sigma_t^{\text{cf}})$ and feasibility of $\alpha_{0:t+1} = 0$. Defining $s := \max\{\text{tr}(\Sigma_0^{\text{cf}}), \text{tr}(\Sigma_1^{\text{cf}}), \dots, \text{tr}(\Sigma_{n-1}^{\text{cf}}), s^0\}$, we get the first part of the result. Similarly, we can observe that $\beta_{0:N+1}$ is feasible for (19) for all $N \geq n$ and $\Sigma_t^{\text{cm}} = \Sigma_t^{\text{cf}}$ for $t \leq N$. $\square$

Proof of Theorem 1. For any $t \geq 0$, the optimal cost $\text{tr}(\Sigma_t^{\text{cf}}) \leq s$ due to Lemma 5. Therefore, Snyder's (1973, Lemma 6) gives us the bound $\lambda_{\min}(\Sigma_0^-) \text{tr}(z_{t|t} z_{t|t}^\top) \leq \text{tr}(z_{t|t}^\top \Sigma_0^- z_{t|t}) \leq \text{tr}(\Sigma_t^{\text{cf}}) \leq s$, which further implies

$$\text{tr}(z_{t|t} z_{t|t}^\top) \leq \frac{s}{\lambda_{\min}(\Sigma_0^-)}. \quad (\text{B.5})$$

Set $|\hat{x}_0^- - x_0| < \delta$ and consider $|\hat{x}_t^{\text{cf}} - x_t|^2$. Since from (B.3) $\hat{x}_t^{\text{cf}} - x_t = z_{t|t}^\top (\hat{x}_0^- - x_0)$, by using the bound (B.5) we get

$$\begin{aligned} |\hat{x}_t^{\text{cf}} - x_t|^2 &= (\hat{x}_0^- - x_0)^\top z_{t|t} z_{t|t}^\top (\hat{x}_0^- - x_0) \leq \lambda_{\max}(z_{t|t} z_{t|t}^\top) |\hat{x}_0^- - x_0|^2 \\ &\leq \text{tr}(z_{t|t} z_{t|t}^\top) |\hat{x}_0^- - x_0|^2 \leq \frac{s}{\lambda_{\min}(\Sigma_0^-)} \delta^2 =: \varepsilon^2. \end{aligned} \quad (\text{B.6})$$

Therefore, for a given $\varepsilon > 0$, we can choose $\delta = \sqrt{\frac{\lambda_{\min}(\Sigma_0^-)}{s}} \varepsilon$ which results in $|\hat{x}_t^{\text{cf}} - x_t| \leq \varepsilon$ when $|x_0 - \hat{x}_0^-| \leq \delta$ for all $t \geq 0$. In order to prove convergence of $\hat{x}_t^{\text{cf}}$ to x_t for the system (21), we first consider the case when $Q - \Sigma_0^- \geq 0$. For all $t \geq n + 1$, $\text{tr}(\Sigma_t^{\text{cf}})$ is a monotonically increasing sequence due to Lemma 3 and it is bounded above due to Lemma 5. Therefore, it is convergent. From Lemma 4, $\text{tr}(\Sigma_t^{\text{cf}}) - \text{tr}(\Sigma_{t-1}^{\text{cf}}) \rightarrow 0$, which implies $\text{tr}(z_{t|t} z_{t|t}^\top) \rightarrow 0$ because $\Sigma_0^- > 0$. Then (B.6) immediately confirms that $|\hat{x}_t^{\text{cf}} - x_t| \rightarrow 0$ as $t \rightarrow \infty$. Now, we consider the second case when the stabilizing condition (22) of Lemma 4 is satisfied ((C2) holds). In this case, $\text{tr}(\Sigma_t^{\text{cf}})$ is a monotonically decreasing sequence which is bounded below. Similar to the first case, the convergence of $\text{tr}(\Sigma_t^{\text{cf}})$ implies $\text{tr}(z_{t|t} z_{t|t}^\top) \rightarrow 0$, which further implies $|\hat{x}_t^{\text{cf}} - x_t| \rightarrow 0$. $\square$

Proof of Theorem 2. Let us consider the expression of Σ_t^{cm} from (20), $\Sigma_t^{\text{cm}} = z_{N|t}^\top \Sigma_{t-N}^{\text{cm}} z_{N|t} + \alpha_{N|t}^\top R \alpha_{N|t} + \sum_{i=0}^{N-1} \ell_{i|t}$, where

$\ell_{i|t} = z_{i|t}^\top Q z_{i|t} + \alpha_{i|t}^\top R \alpha_{i|t}$. By substituting the expression of $\Sigma_{t-N}^{\text{cm}-}$ from (18), we get $\Sigma_t^{\text{cm}} = z_{N|t}^\top (A \Sigma_{t-(N+1)}^{\text{cm}} A^\top + Q) z_{N|t} + \alpha_{N|t}^\top R \alpha_{N|t} + \sum_{i=0}^{N-1} \ell_{i|t} = \sum_{i=0}^N \ell_{i|t} + z_{N|t}^\top A \Sigma_{t-(N+1)}^{\text{cm}} A^\top z_{N|t}$. Let us define $\gamma_{t,j} := A^\top z_{N|t-(j-1)(N+1)} \gamma_{t,j-1}$ with $\gamma_{t,0} = I$. Therefore,

$$\Sigma_t^{\text{cm}} = \sum_{i=0}^N \ell_{i|t} + \gamma_{t,1}^\top \Sigma_{t-(N+1)}^{\text{cm}} \gamma_{t,1}. \quad (\text{B.7})$$

For any $t = k(N+1) + r$, where $k \in \mathbb{Z}_+$ and $r \in \{0, \dots, N\}$, define $V_j = \sum_{i=0}^N \ell_{i|t-j(N+1)}$, by recursively solving (B.7) we get:

$$\begin{aligned} \Sigma_t^{\text{cm}} &= \sum_{j=0}^{k-1} \gamma_{t,j}^\top V_j \gamma_{t,j} + \gamma_{t,k}^\top \Sigma_r^{\text{cm}} \gamma_{t,k} \\ &= \sum_{j=0}^{k-1} \gamma_{t,j}^\top V_j \gamma_{t,j} + \gamma_{t,k}^\top \left(z_{r|r}^\top \Sigma_0^- z_{r|r} + \alpha_{r|r}^\top R \alpha_{r|r} + \sum_{i=0}^{r-1} \ell_{i|r} \right) \gamma_{t,k}. \end{aligned} \quad (\text{B.8})$$

Since $s \geq \text{tr}(\Sigma_t^{\text{cm}})$ for $t = k(N+1) + r$ due to Lemma 5, after ignoring some non-negative terms, we get $s \geq \text{tr}(\Sigma_t^{\text{cm}}) \geq \text{tr}(\gamma_{t,k} z_{r|r}^\top \Sigma_0^- z_{r|r} \gamma_{t,k}) \geq \text{tr}(\gamma_{t,k} z_{r|r} \gamma_{t,k}^\top z_{r|r}^\top \Sigma_0^-)$. Therefore,

$$\text{tr}(\gamma_{t,k} z_{r|r} \gamma_{t,k}^\top z_{r|r}^\top \Sigma_0^-) \leq \frac{s}{\lambda_{\min}(\Sigma_0^-)}. \quad (\text{B.9})$$

Now, we consider the expression of estimator for CMHE for the system (21) and substitute the expression of $\hat{x}_{t-N}^{\text{cm}-}$ according to our definition (18). For $t = k(N+1) + r$, similar to (B.3), we consider $\hat{x}_t^{\text{cm}} - x_t = z_{N|t}^\top (\hat{x}_{t-N}^{\text{cm}-} - x_{t-N}) = z_{N|t}^\top A (\hat{x}_{t-(N+1)}^{\text{cm}} - x_{t-(N+1)}) = \gamma_{t,k}^\top (\hat{x}_r^{\text{cm}} - x_r) = \gamma_{t,k}^\top z_{r|r}^\top (\hat{x}_0^- - x_0)$. Therefore,

$$\begin{aligned} |\hat{x}_t^{\text{cm}} - x_t|^2 &= |\gamma_{t,k}^\top z_{r|r}^\top (\hat{x}_0^- - x_0)|^2 \leq \lambda_{\max}(\gamma_{t,k}^\top z_{r|r}^\top) |\hat{x}_0^- - x_0|^2 \\ &\leq \frac{s}{\lambda_{\min}(\Sigma_0^-)} |\hat{x}_0^- - x_0|^2 =: \varepsilon^2, \end{aligned} \quad (\text{B.10})$$

where the last inequality is due to (B.9). Therefore, for a given $\varepsilon > 0$, we can choose $\delta = \sqrt{\frac{\lambda_{\min}(\Sigma_0^-)}{s}} \varepsilon$ which results in $|\hat{x}_t^{\text{cm}} - x_t| \leq \varepsilon$ when $|x_0 - \hat{x}_0^-| \leq \delta$ for all $N \geq n$ and $t \geq N+1$. This completes the first part of the proof. For the second part, we consider (B.8) and take limit $t \rightarrow \infty$, we get

$$\lim_{t \rightarrow \infty} \text{tr}(\Sigma_t^{\text{cm}}) = \lim_{k \rightarrow \infty} \text{tr} \left(\sum_{j=0}^{k-1} \gamma_{t,j}^\top V_j \gamma_{t,j} + \gamma_{t,k}^\top \Sigma_r^{\text{cm}} \gamma_{t,k} \right) \leq s,$$

which results in $\text{tr}(\gamma_{t,k-1}^\top V_{k-1} \gamma_{t,k-1}) \rightarrow 0$ as $k \rightarrow \infty$. By substituting $V_{k-1} = \sum_{i=0}^N \ell_{i|t-(k-1)(N+1)}$, we conclude that $\text{tr}(\gamma_{t,k-1}^\top z_{N|t-(k-1)(N+1)}^\top Q z_{N|t-(k-1)(N+1)} \gamma_{t,k-1}) \rightarrow 0$. Since $Q > 0$, we get $\text{tr}(z_{N|t-(k-1)(N+1)}^\top \gamma_{t,k-1} \gamma_{t,k-1}^\top z_{N|t-(k-1)(N+1)}) \rightarrow 0$ as $k \rightarrow \infty$. Now, we consider the expression (B.10) and substitute $\gamma_{t,k} = A^\top z_{N|t-(k-1)(N+1)} \gamma_{t,k-1}$ to get $\hat{x}_t^{\text{cm}} - x_t = (A^\top z_{N|t-(k-1)(N+1)}^\top z_{r|r}^\top (\hat{x}_0^- - x_0))$. Since $s \geq \text{tr}(\Sigma_t^{\text{cm}}) \geq \text{tr}(z_{r|r}^\top \Sigma_0^- z_{r|r})$, we get $|z_{r|r}|_F \leq \sqrt{\frac{s}{\lambda_{\min}(\Sigma_0^-)}}$. We have

$$\begin{aligned} |\hat{x}_t^{\text{cm}} - x_t| &= |(z_{r|r} A^\top z_{N|t-(k-1)(N+1)}^\top \gamma_{t,k-1})^\top (\hat{x}_0^- - x_0)| \\ &\leq |z_{N|t-(k-1)(N+1)}^\top \gamma_{t,k-1}|_F |z_{r|r}|_F |A|_F |\hat{x}_0^- - x_0| \\ &\leq |z_{N|t-(k-1)(N+1)}^\top \gamma_{t,k-1}|_F \sqrt{\frac{s}{\lambda_{\min}(\Sigma_0^-)}} |A|_F |\hat{x}_0^- - x_0|, \end{aligned}$$

which implies $|\hat{x}_t^{\text{cm}} - x_t| \rightarrow 0$ because $|z_{N|t-(k-1)(N+1)}^\top \gamma_{t,k-1}|_F \rightarrow 0$ as $t \rightarrow \infty$. This completes the second part of the proof. $\square$

Proof of Theorem 3. If $Q - \Sigma_0^- \geq 0$, $\text{tr}(\Sigma_t^{\text{cf}})$ is a monotonically increasing sequence due to Lemma 3. We get a feasible control sequence due to Assumption 2. Therefore, due to optimality

$\text{tr}(\Sigma_t^{\text{cf}}) \leq \text{tr} \left(\Gamma_0(z_t) + \alpha_t^\top R \alpha_t + \sum_{i=0}^{t-1} \ell_i \right)$, where $\ell_i = z_i^\top Q z_i + \alpha_i^\top R \alpha_i$, $\alpha_{i+1} = K_i z_i$ and $z_{i+1} = (A^\top + C^\top K_i) z_i$. Due to Assumption 2, we have $\ell_0 \leq \lambda_0 Q$, and for $t \geq 1$, $\ell_t \leq \lambda_t z_{t-1}^\top Q z_{t-1} \leq \lambda_t \ell_{t-1} \leq \lambda_t \lambda_{t-1} \dots \lambda_0 Q$. Let us define $\rho_t := \lambda_t \lambda_{t-1} \dots \lambda_0$, then

$$\begin{aligned} \text{tr}(\Sigma_t^{\text{cf}}) &\leq \text{tr} \left(\Gamma_0(z_t) + \ell_0 + \alpha_t^\top R \alpha_t + \sum_{i=1}^{t-1} \ell_i \right) \\ &\leq \text{tr} \left(z_t^\top (\Sigma_0^- - Q) z_t + \ell_0 + \sum_{i=1}^t \ell_i \right) \\ &\leq \text{tr} \left(\lambda_0 Q + \sum_{i=1}^t \ell_i \right) \leq \text{tr} \left(\sum_{i=0}^t \rho_i Q \right) = \text{tr}(Q) \sum_{i=0}^t \rho_i. \end{aligned}$$

Since $\frac{\rho_{i+1}}{\rho_i} = \lambda_{i+1} < 1$ for each i , there exists $\bar{\rho} > 0$ such that $\sum_{i=0}^t \rho_i < \bar{\rho}$ for each t . Therefore, $\text{tr}(\Sigma_t^{\text{cf}}) \leq \bar{\rho} \text{tr}(Q)$ for each t . Since $\text{tr}(\Sigma_t^{\text{cf}})$ is a monotonically increasing sequence and is bounded above, there exists some $s' > 0$ such that (24) holds. This completes the first part of the proof.

For the second case, the stabilizing condition of Lemma 4 is satisfied, and $\text{tr}(\Sigma_t^{\text{cf}}) \geq 0$ is monotonically decreasing for all $t \geq n+1$. Therefore, there exists some $s' \geq 0$ such that (24) holds. This completes the second part of the proof. $\square$

References

- Alessandri, A., Baglietto, M., & Battistelli, G. (2003). Receding-horizon estimation for discrete-time linear systems. *IEEE Transactions on Automatic Control*, 48(3), 473–478.
- Alessandri, A., Baglietto, M., & Battistelli, G. (2010). A maximum-likelihood Kalman filter for switching discrete-time linear systems. *Automatica*, 46(11), 1870–1876.
- Alessandri, A., & Gaggero, M. (2020). Fast moving horizon state estimation for discrete-time systems with linear constraints. *International Journal of Adaptive Control and Signal Processing*, 34(6), 706–720.
- Andersson, J. A. E., Gillis, J. B., Rawlings, J. B., & Diehl, M. (2019). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1), 1–36.
- Åström, K. J. (1970). *Introduction to stochastic control theory*. New York and London: Academic Press.
- Bakolas, E. (2018a). Constrained minimum variance control for discrete-time stochastic linear systems. *Systems & Control Letters*, 113, 109–116.
- Bakolas, E. (2018b). Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints. *Automatica*, 91, 61–68.
- Bertsekas, D. P. (2005). *Rollout algorithms for constrained dynamic programming: Lab. for information and decision systems report 2646*.
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.
- Brembeck, J. (2019). Nonlinear constrained moving horizon estimation applied to vehicle position estimation. *Sensors*, 19(10), 2276.
- Chatterjee, D., Ramponi, F., Hokayeb, P., & Lygeros, J. (2012). On mean square boundedness of stochastic linear systems with bounded controls. *Systems & Control Letters*, 61(2), 375–380.
- Copp, D. A., & Hespanha, J. P. (2017). Simultaneous nonlinear model predictive control and state estimation. *Automatica*, 77, 143–154.
- Darouach, M., & Zasadzinski, M. (1997). Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, 33(4), 717–719.
- Darouach, M., Zasadzinski, M., & Boutayeb, M. (2003). Extension of minimum variance estimation for systems with unknown inputs. *Automatica*, 39(5), 867–876.
- Farina, M., Ferrari-Trecate, G., & Scattolini, R. (2010). Distributed moving horizon estimation for linear constrained systems. *IEEE Transactions on Automatic Control*, 55(11), 2462–2475.
- Flayac, E. (2019). *Coupled methods of nonlinear estimation and control applicable to terrain-aided navigation* (Ph.D. thesis), Paris Saclay.
- Garcia T., J. F., Marquez-Ruiz, A., Botero C., H., & Angulo, F. (2018). A new approach to constrained state estimation for discrete-time linear systems with unknown inputs. *International Journal of Robust and Nonlinear Control*, 28(1), 326–341.
- Gharbi, M., & Ebenbauer, C. (2018). A proximity approach to linear moving horizon estimation. *IFAC-PapersOnLine*, 51(20), 549–555.
- Goodwin, G., Seron, M. M., & De, D. (2006). *Constrained control and estimation: an optimisation approach*. Springer Science & Business Media.

Haseltine, E. L., & Rawlings, J. B. (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial and Engineering Chemistry Research*, 44(8), 2451–2460.

Haverbeke, N. (2011). *Efficient numerical methods for moving horizon estimation* (Ph.D. thesis), Katholieke Universiteit Leuven, Heverlee, Belgium.

Jazwinski, A. (1968). Limited memory optimal filtering. *IEEE Transactions on Automatic Control*, 13(5), 558–563.

Keerthi, S. S., & Gilbert, E. G. (1988). Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations. *Journal of Optimization Theory and Applications*, 57(2), 265–293.

Kim, J. W., Mehta, P. G., & Meyn, S. P. (2019). What is the Lagrangian for nonlinear filtering? In *58th conf. on decision and control* (pp. 1607–1614). IEEE.

Ko, S., & Bitmead, R. R. (2007). State estimation for linear systems with state equality constraints. *Automatica*, 43(8), 1363–1368.

Kong, H., Shan, M., Su, D., Qiao, Y., Al-Azzawi, A., & Sukkarieh, S. (2020). Filtering for systems subject to unknown inputs without a priori initial information. *Automatica*, 120, Article 109122.

Kong, H., & Sukkarieh, S. (2018). Metamorphic moving horizon estimation. *Automatica*, 97, 167–171.

Kwon, B. K., Han, S., Kwon, O. K., & Kwon, W. H. (2007). Minimum variance FIR smoothers for discrete-time state space models. *IEEE Signal Processing Letters*, 14(8), 557–560.

Kwon, W. H., Kim, P. S., & Han, S. H. (2002). A receding horizon unbiased FIR filter for discrete-time state space models. *Automatica*, 38(3), 545–551.

Kwon, W. H., Kim, P. S., & Park, P. (1999). A receding horizon Kalman FIR filter for discrete time-invariant systems. *IEEE Transactions on Automatic Control*, 44(9), 1787–1791.

Liew, C. K. (1976). Inequality constrained least-squares estimation. *Journal of the American Statistical Association*, 71(355), 746–751.

Löfberg, J. (2004). YALMIP: A toolbox for modeling and optimization in MATLAB. In *Int. symposium on computer aided control systems design* (pp. 284–289). IEEE.

Makkapati, V. R., Rajpurohit, T., Okamoto, K., & Tsiotras, P. (2020). Covariance steering for discrete-time linear-quadratic stochastic dynamic games. In *59th conf. on decision and control* (pp. 1771–1776). IEEE.

Mishra, P. K., Chatterjee, D., & Quevedo, D. E. (2017). Output feedback stable stochastic predictive control with hard control constraints. *IEEE Control Systems Letters*, 1, 382–387.

Mishra, P. K., Chatterjee, D., & Quevedo, D. E. (2020). Stochastic predictive control under intermittent observations and unreliable actions. *Automatica*, 118, Article 109012.

Morabito, B., Kögel, M., Bullinger, E., Pannocchia, G., & Findeisen, R. (2015). Simple and efficient moving horizon estimation based on the fast gradient method. *IFAC-PapersOnLine*, 48(23), 428–433.

Mortensen, R. E. (1968). Maximum-likelihood recursive nonlinear filtering. *Journal of Optimization Theory and Applications*, 2(6), 386–394.

Muske, K. R., Rawlings, J. B., & Lee, J. H. (1993). Receding horizon recursive state estimation. In *American control conf.* (pp. 900–904). IEEE.

Pavon, M., & Wets, R. J. B. (1982). The duality between estimation and control from a variational viewpoint: The discrete time case. In *Algorithms and theory in filtering and control* (pp. 1–11). Springer.

Pearson, J. D. (1966). On the duality between estimation and control. *SIAM Journal on Control*, 4(4), 594–600.

Pluymsers, B., Roobrouck, L., Buijs, J., Suykens, J. A. K., & De Moor, B. (2005). Constrained linear MPC with time-varying terminal cost using convex combinations. *Automatica*, 41(5), 831–837.

Rao, C. V. (2000). *Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems* (Ph.D. thesis), University of Wisconsin-Madison.

Rao, C. V., Rawlings, J. B., & Lee, J. H. (2001). Constrained linear state estimation—a moving horizon approach. *Automatica*, 37(10), 1619–1628.

Risbeck, M. J., & Rawlings, J. B. (2016). MPCTools: Nonlinear model predictive control tools for CasADi (Octave interface). URL <https://bitbucket.org/rawlings-group/octave-mpctools>.

Schneider, R., Hannemann-Tamás, R., & Marquardt, W. (2015). An iterative partition-based moving horizon estimator with coupled inequality constraints. *Automatica*, 61, 302–307.

Simon, D. (2010). Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8), 1303–1318.

Snyders, J. (1973). On the error matrix in optimal linear filtering of stationary processes. *IEEE Transaction on Information Theory*, 19(5), 593–599.

Sui, D., & Johansen, T. A. (2014). Linear constrained moving horizon estimator with pre-estimating observer. *Systems & Control Letters*, 67, 40–45.

Sui, D., Johansen, T. A., & Feng, L. (2010). Linear moving horizon estimation with pre-estimating observer. *IEEE Transactions on Automatic Control*, 55(10), 2363–2368.

Teixeira, B. O. S., Chandrasekar, J., Torres, L. A. B., Aguirre, L. A., & Bernstein, D. S. (2007). State estimation for equality-constrained linear systems. In *46th conf. on decision and control* (pp. 6220–6225). IEEE.

Todorov, E. (2008). General duality between optimal control and estimation. In *47th conf. on decision and control* (pp. 4286–4292). IEEE.

Toh, K., Todd, M. J., & Tütüncü, R. H. (2012). On the implementation and usage of SDPT3—a Matlab software package for semidefinite-quadratic-linear programming, version 4.0. In *Handbook on semidefinite, conic and polynomial optimization* (pp. 715–754). New York, NY, USA: Springer.

Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1), 25–57.

Yang, C., & Blasch, E. (2009). Kalman filtering with nonlinear state constraints. *IEEE Transactions on Aerospace and Electronic Systems*, 45(1), 70–84.

Zhao, S., Shmaliy, Y. S., Huang, B., & Liu, F. (2015). Minimum variance unbiased FIR filter for discrete time-variant systems. *Automatica*, 53, 355–361.



Prabhat K. Mishra is a post-doctoral research associate at Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA. He was awarded the Ph.D. degree in 2019 from Indian Institute of Technology Bombay, India. His doctoral study was supported by Ministry of Human Resource and Development, India, and he received a Swiss Government Excellence Scholarship in 2017 to study at EPFL, Switzerland, for a year. His research interests include predictive, adaptive and learning based control.



Girish Chowdhary is an associate professor and Donald Biggar Willet faculty fellow, and ACES Office of Research Faculty fellow at UIUC. He holds a joint appointment with Agricultural and Biological Engineering and Computer Science, and has affiliate appointments in Aerospace Engineering, Electrical and Computer Engineering. He is a full member of the Coordinated Science Laboratory. He is the director of the Field Robotics and Engineering Research Hub (FRESH), the associate director for the National AI institute (AIFARMS) and the Chief Scientist of the Illinois Autonomous Farm. He received his Ph.D. in Aerospace Engineering with a focus on robotics and autonomy from Georgia Institute of Technology (2010). He was a postdoc at the Laboratory for Information and Decision Systems (LIDS) of the Massachusetts Institute of Technology (2011–2013), and an assistant professor at Oklahoma State University's Mechanical and Aerospace Engineering department (2013–2016). He has a bachelor's degree in Aerospace Engineering with honors from the Royal Melbourne Institute of Technology in Melbourne Australia (2003). After his bachelors, he also worked with the German Aerospace Center's (DLR's) Institute of Flight Systems for around three years (2003–2006). His work on neural network based adaptive flight control has led to several key advances to flight control and his Dave Ward memorial award by the Aerospace Guidance and Controls Committee. He won the AFOSR Young Investigator Award in 2015, and several best paper awards. He is author of over 90 publications in autonomy and robotics and a PI on NSF, Air Force Office of Scientific Research (AFOSR), NASA, Advanced Research Projects Agency-Energy (ARPA-E), and Department of Energy (DOE) grants and Office of Naval Research (ONR) Multidisciplinary University Research Initiatives (MURI). Dr. Chowdhary is an Associate Fellow of the American Institute of Aeronautics and Astronautics (AIAA), and a senior member of IEEE.



Prashant G. Mehta received the Ph.D. degree in applied mathematics from Cornell University, Ithaca, NY, USA, in 2004. He is a Professor of mechanical science and engineering with the University of Illinois at Urbana-Champaign, IL, USA. Prior to joining Illinois, he was a Research Engineer with the United Technologies Research Center (UTRC), East Hartford, CT, USA. His current research interests are in nonlinear filtering. Dr. Mehta was the recipient of the Outstanding Achievement Award at UTRC for his contributions to the modeling and control of combustion instabilities in jet-engines. His students were the recipients of the best student paper awards at the IEEE Conference on Decision and Control for 2007, 2009, and 2019, respectively and were finalists for these awards in 2010 and 2012. He has served on the editorial boards for the ASME Journal of Dynamic Systems, Measurement, and Control and the Systems and Control Letters. He currently serves on the editorial board for the IEEE Transactions on Automatic Control.