2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Moving Horizon Estimation for Discrete-Time Linear Time-Invariant Systems Using Transfer Learning

Junyao Xie[a], Biao Huang[b,*]

Abstract— **In this article, we propose a novel moving horizon estimation method for discrete-time linear systems through transfer learning. Most moving horizon estimator designs require data from the considered systems of interest. However, practical processes might suffer from data availability issues, especially in a new or early operating environment. Motivated by the idea of transfer learning, this manuscript proposes a moving horizon estimator design using data from a similar but different system (i.e., source system) instead of the considered system (i.e., target system). Based on the data from the source system, we propose a novel moving horizon state estimation method for the target system and provide convergence and stability analyses. The state estimation error is upper bounded by a time-dependent sequence that is related to three types of similarities/differences between target and source systems, including initial conditions, disturbance levels, and model parameters. The effectiveness of the proposed approach is demonstrated through a numerical example.**

I. INTRODUCTION

Over the past decades, moving horizon estimation (MHE), or receding horizon estimation, has been an active research topic in the systems and control community. As an optimization-based technique, MHE has shown its capacity of addressing state and parameter estimation while handling constraints, and hence has been applied to various types of dynamical systems, including linear lumped parameter (finite-dimensional) systems [1]–[3], nonlinear lumped parameter systems [4]–[6], linear distributed parameter (infinite-dimensional) systems [7], hybrid systems [8]–[10], and so on. However, most of the MHE designs rely on data from the considered systems of interest, which might not be possible for some practical applications, e.g, those at early or new operating conditions. In this work, we propose a novel MHE method named transfer MHE by using concepts from transfer learning, and prove its convergence properties by using data borrowed from a similar system.

The fundamental motivation of transfer learning for machine learning (ML) was discussed in a NIPS-95 workshop on "Learning to Learn", where one theme was focused on developing methods using models previously learned for some problems as a basis when learning new, but related, problems [1]. The specific aspects and methods of transfer learning were comprehensively reviewed in [11], where four different classes of approaches were summarized, including instance-transfer (or output data transfer), feature-representation-transfer, parameter-transfer, and relational-knowledge-transfer, which provide important insights and concepts on the transfer learning. Transfer learning approaches have been widely used for fault diagnosis and soft sensing; see e.g., [12], [13]. Recently, the concepts of transfer learning were used for solving a linear system identification problem in [14], where state measurement data generated by a similar (but not identical) system as well as state measurement data from the true system were jointly utilized for identifying the true system in the presence of process noise. However, research work on transfer learning for state estimation remains under-explored in the literature.

In this work, we apply the concepts of transfer learning to the moving horizon estimator design of discrete-time linear time-invariant systems. More specifically, we consider the cases that there is no data available in the considered system (i.e., target system), and design a moving horizon estimator of the target system by leveraging data from a similar but different system (i.e., source system). To the best of authors' knowledge, it is the first contribution that combines moving horizon estimation and transfer learning and uses them for state estimation of a target system using data from a source system. The MHE framework that we use in this work is adopted from [3] considering that it is formulated as an entirely deterministic scheme using a fixed prior diagonal matrix (a user-designed parameter) as the arrival cost, which is different from the ones in [1], [2]. Moreover, we consider three types of similarities/differences between target and source systems, including initial conditions, disturbance levels, and model parameters. The contributions of this article are two-fold.

The first contribution is that we develop a novel moving horizon estimation approach based on the concepts of transfer learning. The typical MHE algorithms deploy a recursive least squares mechanism using equally weighted output samples. In this article, we propose a transfer MHE algorithm as a recursive weighted least squares algorithm that allows one to assign different weights to each output sample to express the transferability individually. Moreover, we propose a method to determine the weights via solving an optimization problem.

The second contribution is that we provide the convergence and stability analyses for the proposed transfer MHE approach. In particular, we have established the convergence and stability analyses in Theorem 2 in terms of the con-

[a]Junyao Xie is a postdoctoral fellow with Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Canada. junyao3@ualberta.ca

[b]Biao Huang is a professor with Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Canada. biao.huang@ualberta.ca

*Corresponding author.

[1]http://socrates.acadiau.ca/courses/comp/dsilver/NIPS95 LTL/transfer.workshop.1995.html

sidered three kinds of similarities/differences between target and source systems, including initial conditions, disturbance levels, and model parameters. We prove that the state estimation error is upper bounded by a time-dependent bounding sequence whose convergence depends on the considered three kinds of similarities/differences between target and source systems. In addition, we show that in Corollary 1 the transfer MHE using source-domain data would outperform the traditional MHE using target-domain data, provided that the difference between the model parameters of target and source systems are sufficiently small and the disturbance level of source-domain data is milder than that of the target-domain data.

The rest of this article is organized as follows. The problem statement of transfer moving horizon estimation is formulated in Section II. The moving horizon estimation algorithm design and its convergence and stability analyses are presented in Section III. Section IV validates the proposed theoretical results on a numerical example. Finally, concluding remarks are drawn in Section V.

Notations: We denote $v_n^m = \text{col}(v_m, v_{m+1}, \ldots, v_n)$ and $\{x_k\}_0^T = \{x_0, x_1, \ldots, x_T\}$. $\mathbb{N}_0$ denotes the set of non-negative integers and $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. For an element $u \in \mathbb{R}^n$, the symbol $\|u\|$ denotes the 2-norm of u in space $\mathbb{R}^n$, and $\|u\|_P$ represents the weighted norm of u, namely, $\|u\|_P = (u'Pu)^{1/2}$. Moreover, for a positive integer N, $\prod_{i=1}^N A_i = A_1 A_2 \ldots A_N$ denotes the ordered product of matrix A_i, $i = 1, 2, \ldots, N$.

II. Problem Statement

Consider a discrete-time linear system with plant and measurement disturbances as:

$$x_{k+1} = Ax_k + w_k \tag{1a}$$

$$y_k = Cx_k + v_k \tag{1b}$$

for $k \in \mathbb{N}_0$, where $x_k \in \mathbb{R}^n$ denotes the state (the initial state x_0 is unknown) and $y_k \in \mathbb{R}^p$ represents the output measurement at time k. Notations $w_k \in W \subset \mathbb{R}^n$ and $v_k \in V \subset \mathbb{R}^p$ represent plant and measurement disturbances, respectively. We assume that the statistics of initial state x_0, plant disturbance w_k, and output disturbance v_k are unknown, and consider them as deterministic variables of an unknown kind. If there are output data available from system (1), also termed as target system of interest, one can readily deploy the typical methods proposed in [1]–[3] for a moving horizon state estimator design. In this paper, we consider that there are no data available from target system (1), and we aim to address the moving-horizon state estimation of target system (1) using data from another system that is similar to but different from the target system.

Consider another discrete-time linear system with plant and measurement disturbances as:

$$x_{k+1}^s = (A + \delta A_k)x_k^s + w_k^s \tag{2a}$$

$$y_k^s = (C + \delta C_k)x_k^s + v_k^s \tag{2b}$$

for $k \in \mathbb{N}_0$, where $x_k^s \in \mathbb{R}^n$ denotes the state (the initial state x_0^s is unknown) and $y_k^s \in \mathbb{R}^p$ represents the output measurement at time k. Notations $w_k^s \in W^s \subset \mathbb{R}^n$ and $v_k^s \in V^s \subset \mathbb{R}^p$ represent the plant and measurement disturbances, respectively. To distinguish the notations from the target system (1), we use the superscript "s" to denote the system (2) (termed as the source system[2]). Similar to the target system, we assume that the statistics of initial state, plant disturbance, and output disturbance of source system (2) are unknown, and consider them as deterministic variables of an unknown kind. In addition, the matrices δA_k and δC_k are assumed to be known to the designer but time varying, indicating that the parameters of the target and source systems are different in a time-varying manner. Moreover, δA_k and δC_k are assumed to belong to the known compact sets $\mathcal{A}$ and $\mathcal{C}$, respectively.

Throughout this article, we consider the following estimation problem:

Problem 1: Estimate state $\{x_k\}_0^T$ of the target system (1) by using output measurements $\{y_k^s\}_0^T$ from source system (2).

Most existing references (e.g., [1]–[10]) address the state estimation problems using data coming from the considered system (i.e., target system). However, in practice, this might not be the case, where no (or very limited) data are available from the target system, while sufficient data available from similar systems (i.e., source systems). This motivates the present work. In particular, the formulated state estimation problem of target system (1) by using measurements from source system (2) enables one to account for the similarities and differences between source system (2) and target system (1) in terms of model parameters, initial conditions, as well as the disturbance levels. Intuitively, if the differences between source system (2) and target system (1) is sufficiently small, the state estimation of the target system using source-domain data can achieve relatively reasonable results.

For rigorous analysis, we make the following assumptions.

Assumption 1: W, V, W^s, and V^s are compact sets, and we denote that $r_w = \max_{w \in W} \|w\|$, $r_v = \max_{v \in V} \|v\|$, $r_w^s = \max_{w^s \in W^s} \|w^s\|$, and $r_v^s = \max_{v^s \in V^s} \|v^s\|$.

Assumption 2: System (1) is stable, and system (2) is quadratically stable, that is, there exists a positive definite matrix P such that is:

$$(A + \delta A)' P (A + \delta A) - P < 0, \ \forall \delta A \in \mathcal{A}$$

Assumption 3: The pair (A, C) is completely observable in N steps.

Assumption 1 ensures bounded plant and measurement disturbances in both target and source systems. Assumption 2 guarantees that the spectral radius of A is smaller than 1, and the spectral radius of $A + \delta A$ is smaller than 1 for all $\delta A \in \mathcal{A}$, namely,

$$a \triangleq \|A\| < 1 \text{ and } a_P \triangleq \max_{\delta A \in \mathcal{A}} \|A + \delta A\|_P < 1 \tag{3}$$

where we further denote $a_\delta \triangleq \max_{\delta A_k \in \mathcal{A}} \|\delta A_k\|_P$. For an operating procedure to verify Assumption 2, readers are referred to [15, Proposition 2]. The observability condition

[2]In transfer learning problems, target and source domains are commonly used terminologies. Here, we adopt the similar terminologies.

in Assumption 3 is commonly needed to show the stability analysis in observer/estimator design.

Remark 1: For simplicity, we consider target and source systems (1)-(2) by omitting control inputs. It is worth noting that in practice the target and source systems can possibly have totally different control inputs from close loops depending the specific regulation purposes, thus making it more challenging for a meaningful transfer.

III. Moving Horizon Estimation

In this section, we propose a novel moving horizon estimation framework based on the one proposed in [3]. The differences between our design and the one in [3] are twofold. First, our framework takes a weighted least squares framework instead of a least squares framework as in [3]. Second, we use the output measurement data from source system (2) (instead of the data from the considered system) for state estimation of target system (1). The main idea of the MHE design for system (1) is to minimize a quadratic cost J_t defined on the interval $[t-N,t]$ as:

$$J_t(\hat{x}_{t-N|t})=\mu\|\hat{x}_{t-N|t}-\bar{x}_{t-N}\|^2+\sum_{k=t-N}^{t}\|q_k y_k^s-C\hat{x}_{k|t}\|^2 \quad \text{(4a)}$$

$$\text{s.t.}: \ \hat{x}_{k+1|t}=A\hat{x}_{k|t}, \ k\in\{t-N,\ldots,t-1\} \quad \text{(4b)}$$

where $\|\cdot\|$ denotes the Euclidean norm, $\hat{x}_{k|t}$ represents the estimated state at time t, $\bar{x}_{t-N}$ denotes a priori estimate of x_{t-N} at time $t-1$, and μ is a non-negative scalar that is a design parameter expressing the belief in the state estimate $\bar{x}_{t-N}$ compared to the observation model. The second term in (4a) denotes the prediction error computed on the basis of the most recent measurements multiplied by weights q_k. The use of a weighted least squares is motivated by the fact that the target and source systems are similar but different, and we aim to compensate the difference via a weighted least squares framework with proper weights q_k. A similar weighted least squares has been reported in [14] for addressing a system identification problem.

Based on the proposed MHE framework (4), we formulate the following estimation problem:

Problem E_t**:** For a given priori state estimate $\bar{x}_{t-N}$ and output measurements $y_t^{s,t-N}$, find the optimal estimates $\hat{x}^o_{k|t}$ for $k\in\{t-N,\ldots,t\}$ that minimize the cost J_t in (4a) subject to constraint (4b).

When moving from the estimation problem E_t to the next estimation problem E_{t+1}, the optimal state estimate is propagated through:

$$\bar{x}_{t-N+1}=A\hat{x}^o_{t-N|t} \quad (5)$$

For brevity, we make the following notations. The output measurements of target system (1) in the interval $[t-N,t]$ are written as:

$$y_t^{t-N}=F_N x_{t-N}+H_N w_{t-1}^{t-N}+v_t^{t-N} \quad (6)$$

where F_N and H_N are defined as follows:

$$F_N=\begin{pmatrix} C\\ CA\\ \vdots\\ CA^N\end{pmatrix},\ H_N=\begin{pmatrix} 0 & 0 & \ldots & 0\\ C & 0 & \ldots & 0\\ CA & C & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ CA^{N-1} & CA^{N-2} & \ldots & C\end{pmatrix} \quad (7)$$

Considering that Assumption 3 holds, we denote

$$\delta_b \triangleq \lambda_{\min}(F_N'F_N)>0 \quad (8)$$

Similarly, one can define the output measurements sequence of source system (2) in the interval $[t-N,t]$ as:

$$\begin{aligned} y_t^{s,t-N}&=\mathcal{F}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)x^s_{t-N}\\ &\quad+\mathcal{H}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)w_{t-1}^{s,t-N}+v_t^{s,t-N} \end{aligned} \quad (9)$$

where $\mathcal{F}_N$ and $\mathcal{H}_N$ are denoted as follows:

$$\mathcal{F}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)=\begin{pmatrix} (C+\delta C_{t-N})\\ (C+\delta C_{t-N+1})(A+\delta A_{t-N})\\ \vdots\\ (C+\delta C_t)\prod_{i=1}^N(A+\delta A_{t-i})\end{pmatrix} \quad (10)$$

$$\mathcal{H}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)=\begin{pmatrix} 0 & 0 & \ldots & 0\\ (C+\delta C_{t-N+1}) & 0 & \ldots & 0\\ (C+\delta C_{t-N+2})(A+\delta A_{t-N+1}) & (C+\delta C_{t-N+1}) & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ (C+\delta C_t)\prod_{i=1}^{N-1}(A+\delta A_{t-i}) & (C+\delta C_t)\prod_{i=1}^{N-2}(A+\delta A_{t-i}) & \ldots & (C+\delta C_{t-N+1})\end{pmatrix} \quad (11)$$

From (6) and (9) along with notation $Q=\text{diag}\{q_{t-N},q_{t-N+1},\ldots,q_t\}$, we define "transfer innovation" as the distance between the disturbance-free output measurements of target system (1) and the disturbance-free weighted output measurements of source system (2) as:

$$\begin{aligned} &\|Q\mathcal{F}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)x^s_{t-N}-F_N x_{t-N}\|\\ \le&\|(Q\mathcal{F}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)-F_N)x^s_{t-N}\|\\ &+\|F_N(x^s_{t-N}-x_{t-N})\|\\ \le&\gamma\|x^s_{t-N}\|+f_{\max}\|x^s_{t-N}-x_{t-N}\|\\ \le&c_0\|x_0^s\|_P+c_1\|x_0^s-x_0\|_P+c_2 r_w^p+c_3 r_w^{s,p} \end{aligned} \quad (12)$$

where the last equality holds due to $\|x^s_{t-N}\|\le 1/\sqrt{\lambda_{\min}(P)}(a_p^{t-N}\|x_0^s\|+\frac{1-a_p^{t-N}}{1-a_p}r_w^{s,p})$, and we denote that $f_{\max}=\|F_N\|$, $\gamma=\max_{\delta A_{t-N}^{t-1}\in\mathcal{A}^{N+1},\delta C_{t-N}^t\in\mathcal{C}^{N+1}}\|Q\mathcal{F}_N(\delta A_{t-N}^{t-1},\delta C_{t-N}^t)-F_N\|$, and

$$c_0=\frac{1}{\sqrt{\lambda_{\min}(P)}}\left(\gamma a_p^{t-N}+f_{\max}a_\delta\sum_{i=0}^{t-N-1}a_p^i a^{t-N-1-i}\right)$$

$$c_1=\frac{f_{\max}}{\sqrt{\lambda_{\min}(P)}}\times a^{t-N},\ c_2=\frac{f_{\max}}{\sqrt{\lambda_{\min}(P)}}\times\frac{1-a^{t-N}}{1-a}$$

$$c_3 = \frac{1}{\sqrt{\lambda_{\min}(P)}} \times \Bigg(\gamma \frac{1-a_p^{t-N}}{1-a_p} + f_{\max} \times \Big(\frac{1-a^{t-N}}{1-a} + a_\delta \sum_{i=1}^{t-N-1} \frac{1-a_p^i}{1-a_p} a^{t-N-1-i}\Big)\Bigg)$$

Remark 1: The defined "transfer innovation" is dependent on both system parametric difference (reflected by δA_k and δC_k) and the distance between actual initial conditions of target and source systems. Motivated by that the terminology "innovation" is often used in Kalman filter, the definition "transfer innovation" enables us to quantify the transferrable output from source system (2) to facilitate state estimation of target system (1). As expected, it would approach to zero, as the similarity between the target and source domains increases. In the extreme case where target and source domains are identical, such term would be equal to zero, implying we handle two identical systems, and the MHE design reduces to the typical MHE framework with $Q = I$.

For any given initial conditions, we can minimize the transfer innovation with respect to Q as follows:

$$Q = \arg \min_{\delta A_{t-N}^{t-1}, \delta C_{t-N}^{t}, Q} \|Q\mathcal{F}_N(\delta A_{t-N}^{t-1}, \delta C_{t-N}^{t}) - F_N\| \quad (13)$$

where $\delta A_i \in \mathcal{A}$ for $i = t-N, \ldots, t-1$ and $\delta C_i \in \mathcal{C}$ for $i = t-N, \ldots, t$.

Proposition 1: Suppose that $\mu > 0$ or that Assumption 3 is verified. Then, problem E_t has a unique solution given by:

$$\hat{x}^o_{t-N|t} = (\mu I + F_N' F_N)^{-1} \times \left[\mu \bar{x}_{t-N} + F_N' Q y_t^{s,t-N}\right] \quad (14)$$

for $t = N, N+1, \ldots$.

Proof of Proposition 1: The proof can be similarly done by following that of [3, Proposition 1], and hence it is omitted.

In the following, we first revisit an existing result (i.e., [3, Theorem 1]) on stability analysis of MHE in a slightly different form, by using data coming from the considered system, namely, target system (1). The norm of the state estimation error is upper bounded by a sequence that decreases if suitable conditions on the scalar parameter of the cost function are satisfied. This result will be used for performance comparison with the proposed MHE later on.

Theorem 1: Under Assumptions 1 and 3, the norm of the state estimation error of target system (1) using data $\{y_k\}$ from target system (2) is bounded as $\|e_{t-N}\|^2 = \|x_{t-N} - \hat{x}_{t-N}\|^2 \le \eta_{t-N}$, for $t = N, N+1, \ldots$, and the sequence $\{\eta_t\}$ is given by

$$\eta_t = \bar{a}\eta_{t-1} + \bar{b}, \; \eta_0 = \bar{\kappa} \quad (15)$$

for $t = 1, 2, \ldots$, where

$$\bar{\kappa} = \frac{4}{\mu + \delta_b}(\mu\|x_0 - \bar{x}_0\|^2 + \bar{c}), \; \bar{a} = \frac{8\mu a^2}{\mu + \delta_b}$$
$$\bar{b} = \frac{8(\mu r_w^2 + \bar{c})}{\mu + \delta_b}, \; \bar{c} = N^2 r_w^2 \|H_N\|^2 + (N+1)^2 r_v^2$$

If μ is selected such that $\bar{a} < 1$, then

1) the sequence $\{\eta_t\}$ converges to $\eta_{ss} = \lim_{t\to+\infty} \eta_t = \frac{\bar{b}}{1-\bar{a}}$
2) if $\eta_t > \eta_{ss}$, then $\eta_{t+1} < \eta_t$; the sequence $\{\eta_t\}$ is strictly decreasing if $\eta_0 > \frac{\bar{b}}{1-\bar{a}}$. □

Proof of Theorem 1: The proof can be similarly done by following that of [16, Theorem 1]. □

By the following theorem, we show that the stability analysis of the proposed MHE using data coming from source system (2). In a similar manner to Theorem 1, we show that norm of the state estimation error is upper bounded by a sequence that decreases if suitable conditions on the scalar parameter of the cost function are satisfied.

Theorem 2: Under Assumptions 1-3, the norm of the state estimation error of target system (1) using data $\{y_k^s\}$ from source system (2) is bounded as $\|e_{t-N}\|^2 = \|x_{t-N} - \hat{x}_{t-N}\|^2 \le \zeta_{t-N}$, for $t = N, N+1, \ldots$, and the bounding sequence $\{\zeta_t\}$ is given by

$$\zeta_t = \bar{a}\zeta_{t-1} + \hat{b}, \; \zeta_0 = \kappa \quad (16)$$

for $t = 1, 2, \ldots$, where

$$\kappa = \frac{4}{\mu + \delta}(\mu\|x_0 - \bar{x}_0\|^2 - 2\mu r_w^2) + \hat{b}$$
$$\hat{b} = d_1 r_w^2 + d_2 (r_w^s)^2 + d_3 (r_v^s)^2 + d_4 \|x_0^s\|_P^2 + d_5\|x_0^s - x_0\|_P^2 + d_6 (r_w^p)^2 + d_7 (r_w^{s,p})^2$$

and $\bar{a}$ is given in Theorem 1, where $d_1 = \frac{8\mu}{\mu+\delta_b}$, $d_2 = \frac{12 q_{\max}^2 \|\mathcal{H}_N\|^2 N^2}{\mu+\delta_b}$, $d_3 = \frac{12 q_{\max}^2 (N+1)^2}{\mu+\delta_b}$, $d_4 = \frac{48 c_0^2}{\mu+\delta_b}$, $d_5 = \frac{48 c_1^2}{\mu+\delta_b}$, $d_6 = \frac{48 c_2^2}{\mu+\delta_b}$, and $d_7 = \frac{48 c_3^2}{\mu+\delta_b}$, with definition $q_{\max} = \|Q\|$ and notations $\bar{c}_i$, for $i = 0, \ldots, 3$ given in (12).

If μ is selected such that $\bar{a} < 1$, then

1) the sequence $\{\zeta_t\}$ converges to $\zeta_{ss} = \lim_{t\to+\infty} \zeta_t = \frac{1}{\mu+\delta_b - 8\mu a^2} \times \Big[8\mu r_w^2 + 12(r_w^s)^2 q_{\max}^2 \|H_N\|^2 N^2 + 12 (r_v^s)^2 q_{\max}^2 (N+1)^2 + \frac{48 (r_w^{s,p})^2}{\lambda_{\min}(P)} \Big(\frac{\gamma}{1-a_p} + \frac{f_{\max}}{1-a} \times (1 + \frac{a_\delta}{1-a_p})\Big)^2 + \frac{48 f_{\max}^2 (r_w^p)^2}{\lambda_{\min}(P)(1-a)^2}\Big]$, where we denote $r_w^p = \max_{w\in W} \|w\|_P$ and $r_w^{s,p} = \max_{w^s \in W^s} \|w^s\|_P$.
2) if $\zeta_t > \zeta_{ss}$, then $\zeta_{t+1} < \zeta_t$; the sequence $\{\zeta_t\}$ is strictly decreasing if $\zeta_0 > \frac{\hat{b}}{1-\bar{a}}$. □

Proof of Theorem 2: The rationale of the proof is motivated by that of [3, Theorem 1]. More specifically, we seek to find the upper and lower bounds for the cost function J_t.

From the optimality principle, we note that $J_t(\hat{x}_{t-N|t})$ is upper bounded by $J_t(x_{t-N|t})$ as follows:

$$\begin{aligned}
&J_t(\hat{x}_{t-N|t}) \le J_t(x_{t-N|t}) \\
&= \mu\|x_{t-N|t} - \bar{x}_{t-N}\|^2 + \|Q y_t^{s,t-N} - F_N x_{t-N}\|^2 \\
&= \mu\|A(x_{t-N-1} - \bar{x}_{t-N-1}) - w_{t-N-1}\|^2 \\
&\quad + \|Q\mathcal{F}_N x_{t-N}^s - F_N x_{t-N} + Q\mathcal{H}_N w_{t-1}^{s,t-N} + Q v_t^{s,t-N}\|^2 \\
&\le 2\mu\|A(x_{t-N-1} - \hat{x}_{t-N-1})\|^2 + 2\mu\|w_{t-N-1}\|^2 \\
&\quad + 3\|Q\mathcal{F}_N x_{t-N}^s - F_N x_{t-N}\|^2 + 3\|Q\mathcal{H}_N w_{t-1}^{s,t-N}\|^2 \\
&\quad + 3\|Q v_t^{s,t-N}\|^2 \\
&\le 2\mu a^2 \|e_{t-N-1}\|^2 + 2\mu r_w^2 + 3 q_{\max}^2 \|\mathcal{H}_N\|^2 N^2 (r_w^s)^2 \\
&\quad + 3 q_{\max}^2 (N+1)^2 (r_v^s)^2 + 12 c_0^2 \|x_0^s\|_P^2 + 12 c_1^2 \|x_0^s - x_0\|_P^2 \\
&\quad + 12 c_2^2 (r_w^p)^2 + 12 c_3^2 (r_w^{s,p})^2
\end{aligned} \quad (17)$$

where (9) is used in the above derivation, and for brevity, we drop the dependence of $\mathcal{F}_N$ and $\mathcal{H}_N$ on δA_{t-N}^{t-1} and δC_{t-N}^{t}.

By applying the triangular inequality, we derive the following lower bound for $J_t(\hat{x}_{t-N|t})$ as:

$$\begin{aligned}
&J_t(\hat{x}_{t-N|t})\\
&= \mu\|\hat{x}_{t-N|t} - \bar{x}_{t-N}\|^2 + \|Qy_t^{s,t-N} - F_N\hat{x}_{t-N}\|^2\\
&\geq \mu\Big[\frac{1}{2}\|x_{t-N} - \hat{x}_{t-N|t}\|^2 - \|x_{t-N} - \bar{x}_{t-N}\|^2\Big]\\
&\quad + \frac{1}{2}\|F_N x_{t-N} - F_N\hat{x}_{t-N|t}\|^2 - \|Qy_t^{s,t-N} - F_N x_{t-N}\|^2\\
&\geq \frac{\mu+\delta_b}{2}\|x_{t-N} - \hat{x}_{t-N|t}\|^2\\
&\quad - \mu\|A(x_{t-N-1} - \hat{x}_{t-N-1}) - w_{t-N-1}\|^2\\
&\quad - 3\|Q\mathcal{F}_N x_{t-N}^s - F_N x_{t-N}\|^2 - 3\|Q\mathcal{H}_N w_{t-1}^{s,t-N}\|^2\\
&\quad - 3\|Qv_t^{s,t-N}\|^2\\
&\geq \frac{\mu+\delta_b}{2}\|e_{t-N}\|^2 - 2\mu a^2\|e_{t-N-1}\|^2 - 2\mu r_w^2\\
&\quad -3q_{\max}^2\|\mathcal{H}_N\|^2 N^2 (r_w^s)^2 - 3q_{\max}^2(N+1)^2(r_v^s)^2\\
&\quad -12c_0^2\|x_0^s\|_P^2 - 12c_1^2\|x_0^s - x_0\|_P^2 - 12c_2^2(r_w^p)^2\\
&\quad -12c_3^2(r_w^{s,p})^2
\end{aligned} \tag{18}$$

Then by combining the upper and lower bounds in (17) and (18), we can derive the following inequality:

$$\begin{aligned}
\|e_{t-N}\|^2 \leq{}& \frac{8\mu a^2}{\mu+\delta_b}\|e_{t-N-1}\|^2 + d_1 r_w^2 + d_2(r_w^s)^2 + d_3(r_v^s)^2\\
&+ d_4\|x_0^s\|^2 + d_5\|x_0^s - x_0\|^2 + d_6(r_w^p)^2 + d_7(r_w^{s,p})^2
\end{aligned}$$

where the notations of d_i, $i = 1, \ldots, 7$, are given in Theorem 2. Through further algebraic manipulation, one can derive the bounding sequence $\{\zeta_t\}$ given by (16). The proof is completed. □

Theorem 2 provides a converging bounding sequence for the norm of the estimation error under the condition $\bar{a} < 1$. If the prior estimate of the initial condition is chosen properly, a fast converging estimation result can be expected.

Remark 2: From Theorem 2, we note that the bounding sequence $\{\zeta_t\}$ of the state estimation error depends on the plant disturbance levels of target and source systems (i.e., r_w^s and r_w), measurement disturbance level of source system (i.e., r_v^s) (as source data is used), and the transfer innovation. Compared with a typical MHE design using data from the target system, the bounding sequence of state estimation error of the the proposed transfer MHE design is additionally related to the system parametric difference (reflected by δA_k and δC_k) and the distance between actual initial conditions of target and source systems. It is worth noting that Theorems 1 and 2 share the same $\bar{a}$ that determines the stability of the upper bounding sequences $\{\eta_t\}$ and $\{\zeta_t\}$.

Although transfer MHE introduce additional error component related to the system parametric difference and the distance between actual initial conditions of target and source systems, the transfer MHE design allows one to achieve better estimation results provided that the disturbance levels of the adopted source-system data is milder compared to those in target systems in a typical MHE design. More specifically, we will discuss such case in the following corollary.

Corollary 1: Suppose that Assumptions 1-3 are verified. The steady-state upper bound of the state estimation by using the proposed transfer MHE design in Theorem 2 is smaller than that of using the typical MHE design in Theorem 1 if there exist r_w, r_v, r_w^s, $r_v^s \in \mathbb{R}$ such that the following holds: $3q_{\max}^2\big((r_w^s)^2\|H_N\|^2 N^2 + (r_v^s)^2(N+1)^2\big) + \frac{12(r_w^{s,p})^2}{\lambda_{\min}(P)}\left(\frac{\gamma}{1-a_p} + \frac{f_{\max}}{1-a}\times\left(1+\frac{a_\delta}{1-a_p}\right)\right)^2 + \frac{12f_{\max}^2(r_w^p)^2}{\lambda_{\min}(P)(1-a)^2} < 2(N^2 r_w^2\|H_N\|^2 + (N+1)^2 r_v^2)$, where the notations are given in Theorem 2. □

Proof of Corollary 1: Corollary 1 descends directly from Theorems 1-2 along with $\frac{\hat{b}}{1-\bar{a}} < \frac{\bar{b}}{1-\bar{a}}$. □

Under the condition considered in Corollary 1, the transfer MHE in Theorem 2 would achieve a smaller upper bound for state estimation error compared to the typical MHE design in Theorem 1.

Remark 3: It is worth noting that the proposed transfer MHE in the disturbance-free cases (i.e., $r_w = 0$, $r_w^s = 0$, and $r_v^s = 0$) is an asymptotically stable observer (as its estimation error converges asymptotically to zero). Such results are consistent with the typical MHE design, e.g., [3, Corollary 1]. Due to the data availability issues, the formulated transfer state estimation problem is more challenging compared to the typical MHE design in [3]. The proposed transfer MHE enables to leverage the data from a different system for the state estimation of the target system with convergence guarantee.

IV. Numerical Example

In this section, a numerical simulation is provided to demonstrate the effectiveness of the proposed transfer moving horizon state estimation method. More specifically, we consider the target system (1) with the following parameters:

$$A = \begin{pmatrix} 1 & 0.5 \\ -0.125 & 0.9 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

For the source system (2), we first consider two cases having the following parameters: Case 1: $A + \delta A_k = A - 0.001I$, $C + \delta C_k = C - 0.001\begin{pmatrix} 1 & 1 \end{pmatrix}$, and Case 2: $A + \delta A_k = A - 0.01I$, $C + \delta C_k = C - 0.01\begin{pmatrix} 1 & 1 \end{pmatrix}$. In both cases, we consider the initial condition for the target system as $x_0 = [1; 1]$. The parameters of the proposed MHE method are taken as: $N = 10$ and $\mu = 0.001$. In each simulation run of MHE, the priori guess of the initial condition is taken as zero. In Case 1, we consider the initial condition of the source system is $0.98x_0$, along with disturbance norm bounds: $r_w^s = 0.001$ and $r_v^s = 0.001$. In Case 2, we consider the initial condition of the source system is $0.95x_0$, along with disturbance bounds: $r_w^s = 0.01$, and $r_v^s = 0.01$. The estimation results in both cases by using the proposed MHE method are illustrated in Fig. 1. It can be observed that in both cases, the estimated states are similar to the actual ones. Compared to Case 2, the estimation results in Case 1 are better, because of the milder disturbance level, similar initial condition and model parameters.

In Tables I-III, we carry out a sensitivity analysis of the proposed MHE algorithm, subject to different parameter values. The estimation performance is assessed in terms of the root mean square error (i.e., RMSE), and the results are obtained as averaged RMSE values of 100 runs. As shown in Table I, the estimation performance error increases with the increase of the model parameter difference. It can be observed in Table II that as the plant and measurement disturbance levels increase, the estimation performance of the proposed MHE degrades. The influence of the distance between the initial conditions of source and target systems on the estimation performance is shown in Table III, where the estimation error increases as the distance between the initial conditions of source and target systems enlarges. The results verify the proposed theoretical results (e.g., Theorem 2 and Remark 2).

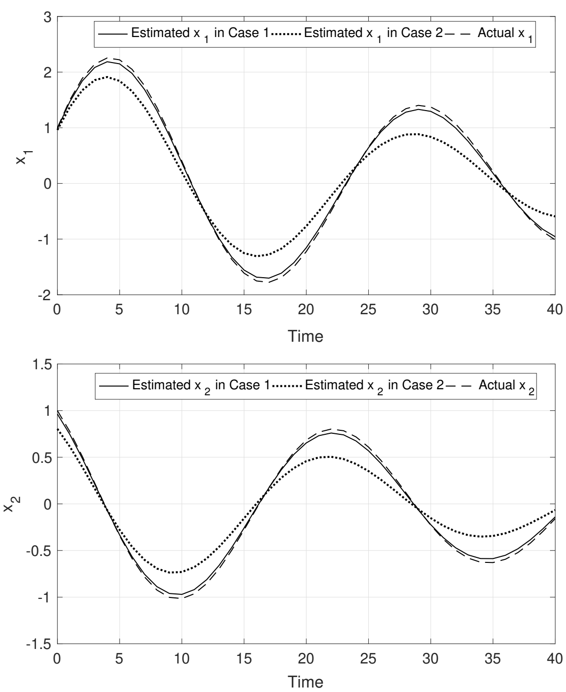

Fig. 1: State estimation results of x_1 (up) and x_2 (bottom)

TABLE I: Estimation performance with different δ in $A + \delta A_k = A - \delta I$, $C + \delta C_k = C - \delta[1, 1]$ along with $x_0^s = 1.05x_0$, $r_w^s = 0.002$, and $r_v^s = 0.002$

δ	0.001	0.01	0.1
RMSE of estimation of x_1	0.0404	0.2320	1.0551
RMSE of estimation of x_2	0.0134	0.1504	0.5658

TABLE II: Estimation performance with different r_w^s and r_v^s along with $\delta = 0.002$ and $x_0^s = 1.05x_0$

r_w^s, r_v^s	0.001	0.01	0.1
RMSE of estimation of x_1	0.0361	0.0463	0.3172
RMSE of estimation of x_2	0.0172	0.0239	0.1853

TABLE III: Estimation performance with different β in $x_0^s = \beta x_0$ along with $r_w^s = 0.002$, $r_v^s = 0.002$, and $\delta = 0.002$

β	1.05	1.25	1.50
RMSE of estimation of x_1	0.0365	0.2563	0.5595
RMSE of estimation of x_2	0.0175	0.1061	0.2470

V. Conclusion

In this work, a transfer moving horizon estimator was designed for linear discrete-time finite-dimensional systems by using data from a similar but different system. The idea of transfer learning was used for addressing the formulated problem. A novel MHE algorithm was proposed in a weighted least squares form. The convergence/stability analyses of the proposed MHE design were conducted in terms of similarities and differences between the source and target systems/data, including initial conditions, disturbance levels, and model parameters. A numerical simulation example was provided to verify the proposed theoretical results. An extension to time-varying systems constitutes future work.

References

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

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

[3] A. Alessandri, M. Baglietto, and G. Battistelli, "Receding-horizon estimation for discrete-time linear systems," *IEEE Trans. Autom. Control*, vol. 48, no. 3, pp. 473–478, 2003.

[4] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations," *IEEE Trans. Autom. Control*, vol. 48, no. 2, pp. 246–258, 2003.

[5] A. Alessandri, M. Baglietto, and G. Battistelli, "Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes," *Automatica*, vol. 44, no. 7, pp. 1753–1765, 2008.

[6] V. M. Zavala, "Stability analysis of an approximate scheme for moving horizon estimation," *Comput. Chem. Eng.*, vol. 34, no. 10, pp. 1662–1670, 2010.

[7] J. Xie, J.-P. Humaloja, C. R. Koch, and S. Dubljevic, "Constrained receding horizon output estimation of linear distributed parameter systems," *IEEE Trans. Autom. Control*, 2022.

[8] G. Ferrari-Trecate, D. Mignone, and M. Morari, "Moving horizon estimation for hybrid systems," *IEEE Trans. Autom. Control*, vol. 47, no. 10, pp. 1663–1676, 2002.

[9] A. Alessandri, M. Baglietto, and G. Battistelli, "Receding-horizon estimation for switching discrete-time linear systems," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1736–1748, 2005.

[10] Y. Guo and B. Huang, "Moving horizon estimation for switching nonlinear systems," *Automatica*, vol. 49, no. 11, pp. 3270–3281, 2013.

[11] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010.

[12] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans Ind. Informat.*, vol. 15, no. 4, pp. 2446–2455, 2018.

[13] J. Xie, B. Huang, and S. Dubljevic, "Transfer learning for dynamic feature extraction using variational Bayesian inference," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 11, pp. 5524–5535, 2021.

[14] L. Xin, L. Ye, G. Chiu, and S. Sundaram, "Identifying the dynamics of a system by leveraging data from similar systems," in *2022 American Control Conference (ACC)*. IEEE, 2022, pp. 818–824.

[15] A. Alessandri, M. Baglietto, and G. Battistelli, "Robust receding-horizon state estimation for uncertain discrete-time linear systems," *Syst. Control. Lett.*, vol. 54, no. 7, pp. 627–643, 2005.

[16] A. Alessandri and M. Awawdeh, "Moving-horizon estimation with guaranteed robustness for discrete-time linear systems and measurements subject to outliers," *Automatica*, vol. 67, pp. 85–93, 2016.