The Noise Covariances of Linear Gaussian Systems with Unknown Inputs Are Not Uniquely Identifiable Using Autocovariance Least-squares

He Kong, Salah Sukkarieh, Travis J. Arnold, Tianshi Chen, and Wei Xing Zheng

Abstract—Existing works in optimal filtering for linear Gaussian systems with arbitrary unknown inputs assume perfect knowledge of the noise covariances in the filter design. This is impractical and raises the question of whether and under what conditions one can identify the noise covariances of linear Gaussian systems with arbitrary unknown inputs. This paper considers the above identifiability question using the correlation-based autocovariance least-squares (ALS) approach. In particular, for the ALS framework, we prove that (i) the process noise covariance $Q$ and the measurement noise covariance $R$ cannot be uniquely jointly identified; (ii) neither $Q$ nor $R$ is uniquely identifiable, when the other is known. This not only helps us to have a better understanding of the applicability of existing filtering frameworks under unknown inputs (since almost all of them require perfect knowledge of the noise covariances) but also calls for further investigation of alternative and more viable noise covariance methods under unknown inputs. Especially, it remains to be explored whether the noise covariances are uniquely identifiable using other correlation-based methods. We are also interested to use regularization for noise covariance estimation under unknown inputs, and investigate the relevant property guarantees for the covariance estimates. The above topics are the main subject of our current and future work.

Index Terms—Estimation; Arbitrary unknown input; Kalman filter; Noise covariance estimation.

I. INTRODUCTION

The last few decades have witnessed much progress in the development of parameter identification techniques and their applications in practice (see, e.g., [1]-[4]). Popular state estimation methods include the well-known Kalman filter (KF) and adaptive KF [5], moving horizon estimation [6]-[9], etc. Despite existing methods’ versatility, their performance might be compromised under unmodeled dynamics whose models or statistical properties are hard to obtain. Typical scenarios include system faults, abrupt/jumping noises, arbitrary vehicle tires/ground interactions, and systems with network-induced effects or attacks (see the in-depth discussions in [10]-[19] and the references therein). Hence, estimation under arbitrary unknown inputs (whose models or statistical properties are not assumed to be available), also called unknown input decoupled estimation, has received much attention in the past. A seminal work on unknown input decoupled estimation is due to Hautus [20] where it has been shown that the strong detectability criterion is necessary and sufficient for the existence of a stable observer for estimating the state/unknown input.

Works on the filtering case, e.g., [21]-[25], have similar rank matching and system being minimum phase requirements as in [20]. Extensions to cases with rank-deficient shaping matrices have been discussed in [26]-[28]. It has also been shown in the above works that for unbiased and minimum variance estimation of the state/unknown input, the initial guess of the state must be unbiased. Recently, connections between the above-mentioned results and KF of systems within which the unknown input is taken to be a white noise of unbounded variance, have been established in [29]. There are also works dedicated to alleviating the strong detectability conditions and the unbiased initialization requirement, and the incorporation of norm constraints (see [30]-[32] and the references therein).

However, most existing filtering works mentioned above assume that the process and measurement noise covariances (denoted as $Q$ and $R$, respectively) are perfectly known for the optimal filter design. This raises the question of whether and under what conditions one can identify $Q/R$ from real data. We believe that addressing the identifiability issue of noise covariances under arbitrary unknown inputs is important because in practice the noise covariances are not known a priori and have to be identified from real closed-loop data where there might be unknown system uncertainties such as faults, etc. Another relevant application is tracking of targets whose motions might be subject to abrupt disturbances (in the form of unknown inputs), as considered in our recent work [33].

To our best knowledge, [34]-[35] are the only existing works on identification of stochastic systems under unknown inputs. However, in [34]-[35], the unknown inputs are assumed to be a wide-sense stationary process with rational power spectral density or deterministic but unknown signals. We do not make such assumptions here. Also, we aim to investigate

1The strong detectability concept was also introduced in [20]. The two criteria, as discussed in [20], are equivalent for discrete-time systems, but differ for continuous systems. Here we focus on the filtering case of discrete-time systems.
the identifiability of the original noise covariances for linear Gaussian systems with unknown inputs. This is in contrast to the work in [34] where the measurement noise covariance of the considered system is assumed to be known, and the input autocorrelations are identified from the output data and then used for input realization and filter design. Our work is also different from subspace identification where the stochastic parameters of the system are estimated and used to calculate the optimal estimator gain [36].

Note that noise covariance estimation is a special parameter identification question that is of lasting interest for the control community, and the literature is fairly mature. Existing noise covariance estimation methods can be classified as Bayesian, maximum likelihood, covariance matching, correlation-based techniques, etc., (see [37]-[40] and the references therein). Especially, the autocovariance least-squares (ALS) framework in [41]-[44] is a popular correlation-based method that has gained much attention in the recent literature [45]-[47]. The main concept of ALS is to design a stable filter without having to know the true noise covariances. Since the filter is suboptimal, the innovations will be correlated, based on which the noise covariance estimation question can be transformed to a standard least-squares optimization problem.

Still, most noise covariance estimation methods mentioned above have not considered the case with unknown inputs. This observation motivates us to study the identifiability of $Q/R$ for systems under unknown inputs. Especially, we discuss the correlation-based ALS framework and show that (i) to apply the ALS framework for the problem at hand, one has to apply a linear transformation to the innovation so that it is decoupled from the unknown inputs (see discussions in Section III. A); (ii) the ALS problem for jointly estimating $Q$ and $R$ does not have a unique solution (see in Theorem 2); (iii) the ALS problem for estimating $Q$ or $R$, when the other is known, does not have a unique solution (see Corollary 1).

The above findings reveal that the noise covariances are in general not uniquely identifiable using the ALS approach. This not only helps us to better understand the applicability of existing filtering frameworks under unknown inputs (since almost all of them require perfect knowledge of the noise covariances) but also calls for further investigation of alternative and more viable noise covariance methods under unknown inputs. Especially, it remains to be explored whether the noise covariances are uniquely identifiable using other correlation-based methods (see [39] and the references therein). We are also interested to use regularization for noise covariance estimation under unknown inputs, and investigate the relevant property guarantees for the covariance estimates (see [43] and the references therein). The above topics are the main subject of our current and future work.

The remainder of the paper is organized as follows. In Section II, we recall preliminaries on estimation of systems with unknown inputs. Section III contains our major results. Section IV verifies the theoretical results with numerical examples. Section V concludes the paper. Notation: $A^T$ denotes the transpose of matrix $A$. $\mathbb{R}^n$ stands for the $n$-dimensional Euclidean space. $I_n$ stands for identity matrices of $n$ dimension. $\mathbb{C}$ and $|z|$ denote the field of complex numbers and the absolute value of a given complex number $z$, respectively. $[a_1, \ldots, a_n]$ denotes $[a_1^T \cdots a_n^T]^T$, where $a_1, \ldots, a_n$ are scalars/vectors/matrices of appropriate dimensions.

II. PRELIMINARIES AND PROBLEM STATEMENT

We consider the discrete-time linear time-invariant (LTI) model of the plant:

$$
\begin{align*}
\begin{cases}
x_{k+1} &= Ax_k + Bd_k + Gw_k \\
y_k &= Cx_k + Dd_k + v_k
\end{cases}
\end{align*}
$$

where $x_k \in \mathbb{R}^n$, $d_k \in \mathbb{R}^q$, and $y_k \in \mathbb{R}^p$ are the state, the unknown input, and the output, respectively; $w_k \in \mathbb{R}^g$ and $v_k \in \mathbb{R}^p$ represent zero-mean uncorrelated Gaussian process and measurement noises with covariances $Q$ and $R$, respectively; $A, B, G, C,$ and $D$ are real and known matrices with appropriate dimensions. (When the noise shaping matrix $G$ is unknown, one needs to identify $\hat{G} = GQG^T$ and $R$.) The analysis of this paper can be directly extended to this case; the pair $(A, C)$ is assumed to be detectable; we also assume that the initial state $x_0$ is independent of the noises. Without loss of generality, we assume $n \geq g$ and $G \in \mathbb{R}^{n \times g}$ to be of full column rank (when this is not the case, one can remodel the system to obtain a full rank shaping matrix $\hat{G}$).

For system (1), a major question of interest is the existence condition of an observer/filter that can estimate the state/unknown input with asymptotically stable error, using only the output. To address this question, concepts such as strong detectability and strong estimator have been rigorously discussed in [20]. As remarked in [20], the term “strong” is to emphasize that state estimate has to be obtained without knowing $d_k$.

Theorem 1: ([20]) The following statements hold true: (i) the system (1) has a strong estimator if and only if it is strongly detectable; (ii) the system (1) is strongly detectable if and only if

$$
\text{rank}
\begin{pmatrix}
CB \\
D
\end{pmatrix}
= \text{rank}(D) + \text{rank}
\begin{pmatrix}
B \\
D
\end{pmatrix},
$$

and all its invariant zeros are stable, i.e.,

$$
\text{rank}
\begin{pmatrix}
zm - A & -B \\
C & D
\end{pmatrix}_{M(z)}
= n + \text{rank}
\begin{pmatrix}
B \\
D
\end{pmatrix},
$$

for all $z \in \mathbb{C}$ and $|z| \geq 1$.

Conditions (2) and (3) are the so-called rank matching and minimum phase requirements, respectively. Note that Theorem 1 holds for both the deterministic and stochastic cases (hence we use “estimator” instead of KF/full state Luenberger observer). For system (1), the noise covariances $Q$ and $R$ are usually not available, and have to be identified from data. However, all existing filtering methods in the literature adopt the assumption of knowing $Q$ and $R$ exactly, which is not practical. This raises the question of whether and under what conditions one can identify $Q$ and/or $R$. Of particular interest in this paper is the correlation-based ALS method. Especially, the questions of interests are formally stated as follows.
Problem 1: Given system (1) with $A, B, G, C,$ and $D$ known, we aim to investigate the following questions: under the assumption that system (1) satisfies the strong detectability condition in Theorem 1 using the ALS approach, whether and under what conditions one can (a) uniquely jointly identify $Q$ and $R$; (b) uniquely identify $Q$ or $R$, assuming the other covariance to be known.

III. IDENTIFIABILITY ANALYSIS OF Q/R USING THE ALS FRAMEWORK

This section contains our major results. In particular, we will prove that (i) $Q$ and $R$ cannot be uniquely jointly identified; (ii) neither $Q$ nor $R$ is not uniquely identifiable, when the other is known.

A. The filter and the choice of innovation model

When the system (1) is strongly detectable, and $Q$ and $R$ are known, one can design an unbiased and optimal filter for estimating the state/unknown input. When $Q/R$ are not known, one can still design an unbiased and stable (but not optimal) filter. To do so for the system (1), we assume $\text{rank}(D) = q$, and adopt the framework of (22). Other methods in (21), (25)-(28), etc., can be considered similarly and we will not elaborate these extensions further. To be specific, the proposed filter implements the following steps recursively after initialization:

1. Unknown input estimation:
   \[ \hat{d}_k = F(y_k - C\hat{x}_{k|k-1}); \] \hspace{1cm} (4)

2. Measurement update:
   \[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + L(y_k - C\hat{x}_{k|k-1}); \] \hspace{1cm} (5)

3. Time update:
   \[ \hat{x}_{k+1|k} = A\hat{x}_{k|k} + B\hat{d}_k. \] \hspace{1cm} (6)

It was shown in (22) that steps in (4) and (6) give unbiased estimates of the unknown input and state, respectively, if and only if the initial guess of the state is unbiased, and $F \in \mathbb{R}^{q \times p}$ and $L \in \mathbb{R}^{n \times p}$ satisfy
\[ [F, L] D = [I_q, 0]. \] \hspace{1cm} (7)

The stability of the filter (4)-(6) can also be guaranteed under conditions (2) (see, e.g., (23), (31)). Define
\[
\begin{align*}
\tilde{y}_k &= y_k - C\hat{x}_{k|k-1}; \quad \hat{d}_k = d_k - \tilde{d}_k, \\
\tilde{x}_{k|k} &= x_k - \hat{x}_{k|k-1}; \quad \tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k},
\end{align*}
\] \hspace{1cm} (8)

as the innovation, the unknown input estimation error, the filtered state error, and the state prediction error, respectively. Based on (1) and (4)-(6), we have
\[
\begin{align*}
\tilde{y}_k &= Dd_k + C\tilde{x}_{k|k-1} + v_k, \\
\tilde{d}_k &= -FC\tilde{x}_{k|k-1} - Fv_k.
\end{align*}
\] \hspace{1cm} (9)

The state-space model to be used for computing the autocovariance in the next subsection is given as follows:
\[
\begin{align*}
\tilde{x}_{k+1|k} &= A\tilde{x}_{k|k-1} + \tilde{G}[w_k, v_k] \\
\tilde{y}_k &= L\tilde{y}_k + \tilde{L}\tilde{x}_{k|k-1} + Lv_k
\end{align*}
\] \hspace{1cm} (10)

where $\tilde{L} = LC, \; A_c = A - KC, \; \tilde{G} = \begin{bmatrix} G & -K \end{bmatrix}, \; K = AL + BF$. It is worthwhile remarking that we have chosen $\tilde{y}_k$ as the output of the error dynamics model in (10), rather than $y_k$ and $d_k$ in (9). This is because $\tilde{y}_k$ is affected by the unknown inputs, for which we do not have to assume any statistical properties; $\tilde{d}_k$ is the unknown input estimation error, which can be used for analysis, but cannot be obtained from data processing (since we do not know the true $d_k$). On the contrary, $\hat{y}_k$ is a linear transformation of the standard innovation $\tilde{y}_k$, and decoupled from $d_k$, thus can be obtained from data processing.

B. $Q$ and $R$ are not uniquely jointly identifiable

When the filter in (4)-(6) is stable, the steady-state estimation error covariance $P$ satisfies
\[ P = A_cPA_c^T + GQG^T + KLRK^T. \] \hspace{1cm} (11)

The autocovariance is defined as the expectation of the data with its lagged version of itself, i.e., $\mathcal{E}(Y_k Y_{k+j}^T)$. Suppose that $\{Y_1, Y_2, \ldots, Y_{N_d}\}$ is the innovations calculated from (10), where $N_d$ is the number of data points. Denote
\[
\mathcal{Y} = 
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_{N_d}
\end{bmatrix}
\in \mathbb{R}^{N_d \times \tilde{n}},
\] where $\tilde{n} = N_d - N + 1$, $N$ represents the window size used in computing the autocovariance. If the steady-state distribution of $\tilde{x}_{k|k-1}$ is used as the initial condition, with (10), we can obtain
\[
\mathcal{E}(Y_k Y_k^T) = \tilde{L}P\tilde{L}^T + LRL^T \in \mathbb{R}^{n \times n},
\] \hspace{1cm} (12)

\[
\mathcal{E}(Y_{k+j} Y_k^T) = \tilde{L}A_c^j P\tilde{L}^T - L\Theta_1^{-1}KRL^T,
\] \hspace{1cm} with $j \geq 1$, and $\tilde{L}$ being defined in (10). Note that the above autocovariance expressions are independent of $k$.

Consider the first column block of the full autocovariance matrix of the innovation process over a window of length $N$:
\[
\mathcal{R}(N) = \mathcal{E} \begin{bmatrix}
Y_k Y_k^T \\
\vdots \\
Y_{k+N-1} Y_k^T
\end{bmatrix} R_{11}^T,
\] \hspace{1cm} (13)

where $\Theta_1 = \begin{bmatrix} \tilde{L}, \tilde{L}A_c, \ldots, \tilde{L}A_c^{N-2} \end{bmatrix}$. The above formulas is the ideal way to compute $\mathcal{R}(N)$. Since the process in (10) is driven by Gaussian noises, it is ergodic, and a practical way of approximating the expectation is to use the time average
\[
\mathcal{R}^*(N) = \frac{1}{N_d - N + 1} \mathcal{Y}_{\text{row}}^T \mathcal{Y}_{\text{row}},
\] \hspace{1cm} (14)

where $\mathcal{Y}_{\text{row}} = \begin{bmatrix} I_n & 0 & \cdots & 0 \end{bmatrix} \mathcal{Y}$. Define $X_\mathcal{R}$ as the outcome of applying the vectorization operator to matrix $X$. Denote
\[
b = (\mathcal{R}^*(N)).s,
\] \hspace{1cm} (15)
In the following, we employ standard definition and properties of the Kronecker product. We apply the vec operator on both sides of (13) and use the fact \((AXB)_s = (B^T \otimes A)X_s\) to obtain
\[
(R(N))_s = (\tilde{L} \otimes \Theta)P_s + (L \otimes \Upsilon)R_s.
\] (16)

From (11), one has
\[
P_s = (A_c \otimes A_c)P_s + (G \otimes G)Q_s + (K \otimes K)R_s.
\] (17)

By following similar steps in (41)-(42), we then have the ALS problem formulation for identifying \(Q\) and \(R\):
\[
\Xi^* = \arg \min_{\Xi} \|H\Xi - b\|_W^2,
\] (18)

where \(b\) is defined in (15), \(W > 0\) is a weighting matrix, and
\[
H = [H_1 \ H_2], \quad \Xi = [Q_s^T \ R_s^T]^T.
\] (19)

If \(H\) in (19) is of full column rank, then the solution of (18) exists and is unique; moreover, desirable properties such as unbiasedness, asymptotic convergence of the covariance estimates, as established in (41)-(43), can be obtained. However, as it will be shown next, it is impossible for \(H\) in (19) to be of full column rank.

**Theorem 2:** Assume that \(N \geq n + 1\); the system (11) satisfies the strong detectability condition in Theorem 1 and a filter of the form (4)-(6) has been designed for it. Then the following statements hold true:

(i) \(H\) is of full column rank if and only if
\[
\text{rank}(A) = \text{rank}(C) = n, \text{rank}(L) = p; \tag{20}
\]

(ii) it is impossible for condition (20) to hold, i.e., the ALS problem (18) does not have a unique solution.

**Proof.** (i) This part follows from similar arguments with those of (44) Chap. 3., and is included here for completeness. Suppose that \(H\Xi = 0\), with \(\Xi\) being defined in (19). An equivalent representation of \(H\Xi = 0\) is \(\Theta P\tilde{L}^T + \Upsilon R\tilde{L}^T = 0\), which can be further expressed as
\[
\tilde{L} P\tilde{L}^T + LRL^T = 0, \quad \Theta_1 A_c P \Sigma C^T L^T = \Theta_1 K R L^T, \tag{21}
\]

where \(\Theta_1\) is defined in (14). Given that the system (11) satisfies the strong detectability condition in Theorem 1 if one designs a filter of the form (4)-(6) for it, one has that \(A_c = A - K C\) is Schur stable (see, e.g., (41)). Hence, \((A_c, \tilde{L})\) is detectable, because \(A_c = A_c - 0 \cdot L\) is stable. It follows that \(\Theta_1\) is of full column rank, when \(N \geq n + 1\). We first prove sufficiency. Based on (21), we have
\[
R = -C P C^T, \quad A_c P C^T - K R = 0,
\]
when \(L\) is of full rank. The above two equalities lead to
\[
A P C^T = 0. \tag{22}
\]

If \(A\) is nonsingular and \(C\) is of full column rank, then we have \(P = 0, R = 0\), which further implies that \(Q = 0\). As a result, the null space of \(H\) only has one element, i.e., the zero vector. In other words, \(H\) is of full column rank. We next prove necessity. Firstly, assume that \(A\) is singular, i.e., there exists a nonzero vector \(z\) such that \(Az = 0\). Set \(P = zz^T\) so that (22) holds. If we select \(R = -C P C^T\), then one has that the two equalities in (21) hold. Moreover, from (17), one has that
\[
\begin{align*}
(G \otimes G)Q_s = (I_n - A_c \otimes A_c)P_s - (K \otimes K)R_s, \\
\Rightarrow Q_s = M^T (M M^T)^{-1} \mu,
\end{align*}
\]

where we have used the assumption that \(G\) (and hence \(M\)) is of full column rank. Note that, \(\mu\) might or might not be zero. However, for either case, we have \(R_s \neq 0\). Hence, there exists a nonzero vector \(z_1\) such that \(L z_1 = 0\). If we set \(R = z_1 z_1^T, P = 0\), then the two equalities in (21) hold. Then by following similar arguments with the above, there exists a nonzero element \(\Xi = [Q_s^T \ R_s^T]^T\) in the null space of \(H\). The necessity of full column rankness of \(C\) can also be proved similarly.

(ii) Assume that condition (20) holds. For the unbiased filter design, one must have \(LD = 0 \Rightarrow D = 0\), which contradicts the assumption that \(D\) is of full column rank. As a result, it is impossible for condition (4) to hold, and the ALS problem (18) does not have a unique solution. This completes the proof.

**Remark 1:** The results in Theorem 2 can be considered as generalizations of those in (41)-(43) to the case with unknown inputs. Although Theorem 2 is established for the case with direct feedthrough, for the case without feedthrough, i.e., \(D = 0\) in (1), the same statements regarding the non-identifiability of \(Q\) and \(R\) can also be obtained (this can be shown by taking the filtering framework of (24) and similar steps above). For the situation without feedthrough, the innovation model to be used for calculating the autocovariance is different from (10), since in this case only the filtered state error is unbiased (i.e., the predicted state error is biased). Detailed proofs are omitted here due to limited space.

**C. Neither \(Q\) nor \(R\) is uniquely identifiable when the other is known**

We next consider part (b) of Problem 1, i.e., identifiability of \(Q\) or \(R\) when the other is known. We firstly consider the case of estimating \(Q\) when \(R\) is known. Denote
\[
b_Q = b - H_2 R_s, \tag{23}
\]
where \(b\) and \(H_2\) are defined in (15) and (19), respectively. By following similar steps to those of the previous subsection, we have the following ALS problem formulation for identifying \(Q\):
\[
\Xi_Q = \arg \min_{\Xi} \|H_1 \Xi_Q - b_Q\|^2_{W_Q} \tag{24}
\]

where \(\Xi_Q = Q_s, H_1\) is defined in (18), \(b_Q\) is defined in (23), and \(W_Q > 0\). For the case of estimating \(R\) when \(Q\) is known, we denote
\[
b_R = b - H_1 Q_s, \tag{25}
\]
where \(b\) and \(\mathcal{H}_1\) are defined in (15) and (19), respectively. By following similar steps to those of the previous subsection, we then have the following ALS problem formulation for identifying \(R\):

\[
\Xi_{R}^{*} = \arg \min_{\Xi_{R}} \| \| H_2 \Xi_R - b_R \|_F^2 \]

where \(\Xi_R = R_s\), \(H_2\) is defined in (18), \(b_R\) is defined in (23), and \(\Xi_R > 0\). We then have the following result.

**Corollary 1:** Assume that \(N > n + 1\); the system \(\mathbf{1}\) satisfies the strong detectability condition in Theorem 1 and a filter of the form (4)-(6) has been designed for it. Then neither of the ALS problems (24) and (26) has a unique solution; in other words, neither \(Q\) nor \(R\) is uniquely identifiable when the other is known.

**Proof.** For the ALS problems (24) and (26) to have a unique solution, it is necessary that \(\mathcal{H}_1\) and \(\mathcal{H}_2\) are of full column rank, respectively. From Theorem 2, we know that \(\text{rank}(L) = p\) is a necessary condition for the full column rankness of both \(\mathcal{H}_1\) and \(\mathcal{H}_2\). However, as discussed in the proof of Theorem 2, \(L\) cannot have full column rank. In other words, \(\mathcal{H}_1\) and \(\mathcal{H}_2\) cannot be of full column rank. This completes the proof.

Corollary 1 together with Theorem 2 provide a complete and negative answer to Problem 1. For the above-mentioned scenarios where the ALS problems are ill-posed and the non-identifiability of noise covariances is obtained, a natural idea is to use regularization to introduce further constraints to uniquely determine the solution [43]. However, a key question to be answered is whether some desirable properties can be guaranteed for the covariance estimates. A full investigation of the above questions is a subject of our current and future work. The results in this paper can be readily extended to the case with correlated noises. They can also be generalized to other more complex scenarios, e.g., linear time varying systems as in [45], although the solution uniqueness conditions of the corresponding least-squares will be harder to analyze.

**IV. EXAMPLES**

In this section, we use some numerical examples to verify the theoretical results. For simplicity, we take \(G = I_n\) in the plant model (1). Assume that the plant model (1) is has the following system matrices

\[
A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.
\]

It can be easily verified that conditions (2)–(5) are satisfied. To design the filter in (4)-(6), we select

\[
F = \begin{bmatrix} 1 & 0.5 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix},
\]

such that condition (7) is met and \(\mathcal{H}_c\) in (10) is stable. Note that the requirements in (20) are all satisfied except that \(L\) is not of full column rank (because it cannot be, as proved in Theorem 2). Hence, \(\mathcal{H}\) in (15) cannot be of full column rank. To double confirm, select \(N = 10\), it can be checked that \(\text{rank}(\mathcal{H}) = 2, \text{rank}(\mathcal{H}_1) = 1, \text{rank}(\mathcal{H}_2) = 2\), where \(\mathcal{H} \in \mathbb{R}^{40 \times 8}, \mathcal{H}_1 \in \mathbb{R}^{40 \times 4}\) and \(\mathcal{H}_2 \in \mathbb{R}^{40 \times 4}\). This validates Theorem 2 and Corollary 1.

**V. CONCLUSIONS**

The past few decades have witnessed much progress in optimal filtering for stochastic systems with arbitrary unknown inputs and Gaussian noises. However, the existing works assume perfect knowledge of the noise covariances in the filter design, which is impractical. In this paper, for linear Gaussian systems under unknown inputs, we have investigated the identifiability question of the process and measurement noises covariances (i.e., \(Q\) and \(R\)) using the correlation-based ALS method. In particular, we have shown that the ALS problem for estimating \(Q\) and/or \(R\) does not have a unique solution. The above findings reveal that the noise covariances are in general not uniquely identifiable using the ALS approach. This not only helps us to have a better understanding of the applicability of existing filtering frameworks under unknown inputs (since almost all of them require perfect knowledge of the noise covariances) but also calls for further investigation of alternative and more viable noise covariance methods under unknown inputs. Especially, it remains to be explored whether the noise covariances are uniquely identifiable using other correlation-based methods. We are also interested to use regularization for noise covariance estimation under unknown inputs, and investigate the relevant property guarantees for the covariance estimates. The above topics are the main subject of our current and future work.

**VI. ACKNOWLEDGMENT**

The authors would like to thank the reviewers and Editors for their constructive suggestions which have helped to improve the quality and presentation of this paper significantly.

**REFERENCES**

[1] K. Mahata and T. Söderström, Improved estimation performance using known linear constraints, *Automatica*, Vol. 40, No. 8, pp. 1307–1318, 2004.

[2] G. C. Goodwin, J. C. Agüero, J. S. Welsh, J. I. Yuz, G. J. Adams, and C. R. Rojas, Robust identification of process models from plant data, *Journal of Process Control*, Vol. 18, No. 9, pp. 810–820, 2008.

[3] T. Söderström, L. Wang, R. Pintelon, and J. Schoukens, Can errors-in-variables systems be identified from closed-loop experiments, *Automatica*, Vol. 49, No. 2, pp. 681–684, 2013.

[4] B. Q. Mu, H. F. Chen, L. Y. Wang, and G. Yin, Characterization and identification of matrix fraction descriptions for LTI systems, *SIAM Journal on Control and Optimization*, Vol. 52, No. 6, pp. 3694–3721, 2014.

[5] F. Gustafsson, Adaptive filtering and change detection, John Wiley & Sons, 2000.

[6] 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.

[7] H. Kong and S. Sukkarieh, Metamorphic moving horizon estimation, *Automatica*, Vol. 97, pp. 167–171, 2018.

[8] H. Kong and S. Sukkarieh, Suboptimal receding horizon estimation via noise blocking, *Automatica*, Vol. 98, pp. 66–75, 2018.

[9] A. Alessandri and G. Battistelli, Moving horizon estimation: Open problems, theoretical progress, and new application perspectives, *Int. Journal of Adaptive Control and Signal Processing*, Vol. 34, No. 6, pp. 703–705, 2020.

[10] A. Cristofaro and T. A. Johansen, Fault tolerant control allocation using unknown input observers, *Automatica*, Vol. 50, No. 7, pp. 1891–1897, 2014.

[11] Z. Gao, X. Liu, and M. Chen, Unknown input observer based robust fault estimation for systems corrupted by partially-decoupled disturbances, *IEEE Trans. on Industrial Electronics*, Vol. 63, No. 4, pp. 2537–2547, 2016.
