Simultaneous input & state estimation, singular filtering and stability ⋆

Mohammad Ali Abooshahab [a] Mohammed M.J. Alyaseen [b] Robert R. Bitmead [b] Morten Hovd [a]

[a] *Department of Engineering Cybernetics, Norwegian University of Science & Technology, 7491 Trondheim, Norway.*

[b] *Department of Mechanical & Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, MS-0411, La Jolla CA 92093-0411, USA.*

Abstract

Input estimation is a signal processing technique associated with deconvolution of measured signals after filtering through a known dynamic system. Kitanidis and others extended this to the simultaneous estimation of the input signal and the state of the intervening system. This is normally posed as a special least-squares estimation problem with unbiasedness. The approach has application in signal analysis and in control. Despite the connection to optimal estimation, the standard algorithms are not necessarily stable, leading to a number of recent papers which present sufficient conditions for stability. In this paper we complete these stability results in two ways in the time-invariant case: for the square case, where the number of measurements equals the number of unknown inputs, we establish exactly the location of the algorithm poles; for the non-square case, we show that the best sufficient conditions are also necessary. We then draw on our previous results interpreting these algorithms, when stable, as singular Kalman filters to advocate a direct, guaranteed stable implementation via Kalman filtering. This has the advantage of clarity and flexibility in addition to stability. En route, we decipher the existing algorithms in terms of system inversion and successive singular filtering. The stability results are extended to the time-varying case directly to recover the earlier sufficient conditions for stability via the Riccati difference equation.

Key words: stability, state estimation, input estimation, singular filtering

1 Introduction

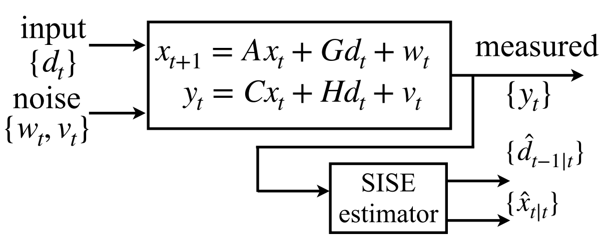

Fig. 1. SISE estimation as posed in this paper.

Figure 1 depicts the Simultaneous Input and State Estimation (SISE) problem formulation: using the measured output signal $y_t \in \mathbb{R}^p$, construct estimates of both plant input signal $d_t \in \mathbb{R}^m$ and plant state $x_t \in \mathbb{R}^n$. Plant matrices $[A, G, C, H]$ are assumed known and w_t and v_t are independent zero-mean white noises of known covariances, $Q \geq 0$ and $R > 0$. The idiosyncrasy of this specification lies in the *feature* that no specific statistical model of the signal d_t is provided. — it is effectively treated as white noise. Denote the d_t-to-y_t transfer function

$$Z_d(z) = H + C(zI - A)^{-1}G. \tag{1}$$

Input reconstruction or estimation from output measurements is a standard signal processing problem of deconvolution [1–3] and, when possible, may be performed by $Z_d(z)$ inversion. SISE augments this by including state estimation.

A number of related SISE algorithms has been developed, each reflecting the delay structure of Z_d in the latency, ℓ, of the input estimate, $\hat{d}_{t|t+\ell}$. Thus: Kitanides [4] and Gillijns & de Moor [5] present an algorithm for delay one, i.e. $H = 0$ and rank $CG = m$; Gillijns & de Moor [6] also provide an algorithm for delay zero, rank $H = m$; and Yong,

⋆ The authors would like to thank for financial support from CINELDI – the Centre for Intelligent Electricity Distribution, Research Council of Norway project no. 257626/E20. Author 2 thanks Kuwait University for funding. This paper was not presented at any IFAC meeting. Corresponding author R.R. Bitmead. Tel. +1 858 822 3477.

Email addresses: mohammad.ali.abooshahab@ntnu.no (Mohammad Ali Abooshahab), malyasee@eng.ucsd.edu (Mohammed M.J. Alyaseen), rbitmead@ucsd.edu (Robert R. Bitmead), morten.hovd@ntnu.no (Morten Hovd).

Zhu & Frazzoli [7] an algorithm (ULISE) for mixed zero-one delays, where H might be less than full rank rank but $\begin{bmatrix} H & CG \end{bmatrix} = m$. Other variations are possible [8,9] and time-variation and more complicated delay structure are readily incorporated because of the connections to Kalman filtering. We do not rederive these algorithms here but, rather, study their stability properties. In Section 5, we present a novel construction for guaranteed stable SISE based on inner-outer factorization and the recovery of an estimate of an all-pass-filtered version of d_t from which d_t can be recovered by different non-real-time means.

In each algorithm, the absence of a statistical model for d_t is accommodated by $\hat{d}_{t|t+\ell}$ being a finite impulse response of the filtered or predicted output error, as appropriate for ℓ. The d_t estimate is then used in the innovations of the state estimator. Such a strategy also appears in the *white noise estimators* which seek to reconstruct the process noise, w_t, from smoothed state estimates [10]. These rapid estimators are related to singular filtering; Kalman filtering with fixed process noise covariance $Q \geq 0$ and measurement noise covariance $R = 0$ exactly [11]. This corresponds to noise-free measurements and, in this case, a number of elements is estimated using geometric construction. The limiting filter with $Q \geq 0$ and $R \stackrel{+}{\to} 0$ is also the singular filter when the plant is minimum phase. This is explored in the dual minimum variance control context using the Return Difference Equality in [12]. In [13], SISE, when stable, is shown to conform to the Kalman filter with fixed $R > 0$ and $Q^{-1} \to 0$, which is an equivalent description of this limiting filter. The connection to SISE with be drawn again here, since singular filtering also is connected with system inversion and its allied stability concerns. Indeed, one of our proofs draws directly from Loop Transfer Recovery/cheap/singular control [14].

There have been a great number of recent papers on these algorithms, with a recent subset [15–17,7] focusing on conditions for stability. Since SISE presumes that no model is available for the disturbance signal, the algorithm proceeds without an explicit description of the disturbance signals' statistical properties. Thus techniques such as extended state observers [18] and augmented Kalman filters [19,20] are inapplicable, as both rely on disturbance models. A recent paper by the authors [13] establishes that, when stable, the linear system SISE algorithm of [5] coincides with the Kalman filter with $\{d_t\}$ modeled by white Gaussian noise of unbounded variance. Various approaches consider first estimating the state [15] and then using the state recursion to reconstruct d_t, or estimating the disturbance first and then reconstructing the state [5–7]. These methods rely on geometric approaches and system inversion, although there is a strong overlap with least-squares state estimation concepts of unbiasedness and optimality.

SISE algorithms go back to least to Kitanidis [4] with antecedents [21,10,22] concentrating on input signal reconstruction. Here, we follow the formulation from Yong, Zhu and Frazzoli [7], which in turn builds on [5,6]. We consider linear time-invariant systems to add clarity and to explore the connection to optimal estimation before extending to uniformly time-varying systems.

Input estimation is a signal processing technique for the recovery of signals after filtering through a known dynamic system. Examples include the estimation of rainfall given river flow and the calculation of salinity in the ocean accommodating for sensor dynamics [23]. Here, the central objective is to estimate the driving disturbance signal d_t and there is little interest in the sensor state. The algorithm should be stable, however. Our particular driving problem, on the other hand, is the estimation of generator states in part of a power grid when the interconnection signals are unknown [24]. Here the priority is to estimate network generator states in the face of unmodeled and unmeasured consumption, which is treated as the disturbance signal. In spite of these distinct objectives, the same algorithms are used.

Contributions & organization

Our objective in this paper is to attempt to bring some clarity and unity to this picture by establishing precisely the connection to system inversion and optimal estimation by deriving necessary and sufficient conditions for stability using explicit system inverse formulæ and algebraic Riccati equations, starting with the time-invariant case. Earlier stability conditions were sufficient only but derived in the time-varying situation. We recover these. Further, when stability is not achieved by these SISE algorithms, we propose a modification based on inner-outer factorization, which maintains state estimation performance at the expense of simple disturbance recovery. This can be compared with the techniques advanced in [15] for approximate system inversion with delay. Beyond this work, we know of no other which addresses estimation when the stability conditions fail.

Section 2 presents the SISE problem for a linear time-invariant system. Section 3 studies the zero direct feedthrough case and the corresponding SISE of [5] and shows that, in the square case where the number of measurements equals the number of disturbance channels, the input estimator is the inverse of the d_t-to-y_t system and the state estimator is a plant simulation. Stability depends on the transmission zeros of the former system. These necessary and sufficient stability conditions then are extended to the non-square case with more measurements. This involves the Riccati difference equation and a detectability condition. Section 4 expands this analysis to the full-rank direct feedthrough case and comments on the non-full-rank case of [7]. Section 5 draws connections to earlier works of singular filtering and introduces an accommodation to circumvent stability issues using the inner-outer factorization. It also contains the extension to time-varying systems via the Riccati equation. Section 6 reinforces the connections to system inversion and concludes. The Appendix contains the proofs.

2 Problem statement

SISE algorithms have been formulated for linear time-varying systems [5–7] and for nonlinear time-varying systems [23,25]. However for clarity of development, we consider the linear, time-invariant system with zero known control input,

$$x_{t+1} = Ax_t + Gd_t + w_t, \tag{2}$$
$$y_t = Cx_t + Hd_t + v_t, \tag{3}$$

with $x_t \in \mathbb{R}^n$, $d_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$. Zero-mean white noises $\{w_t\}$ and $\{v_t\}$ are independent and independent from $\{d_t\}$ and x_0. The covariance of w_t is $Q \geq 0$ and the covariance of v_t is $R > 0$. Denote the signal measurements $\mathbf{Y}^t \triangleq \{y_t, y_{t-1}, \ldots, y_0\}$. The aim is to produce from $\mathbf{Y}^t$, a recursive filtered state estimate, $\hat{x}_{t|t}$, and filtered and/or smoothed estimates, $\hat{d}_{t|t}$ or $\hat{d}_{t-1|t}$, depending on the properties of G and $\begin{bmatrix} C & H \end{bmatrix}$. We make the following assumption.

Assumption 1 *System (2-3) has $[A, C]$ observable, $[A, Q]$ reachable, and $R > 0$.*

Full-rank direct feedthrough, i.e. $\text{rank}(H) = m$, is treated in [6]; zero direct feedthrough, $H = 0$, in [5] with $\text{rank}(CG) = m$; and, [7] provides a generalization, ULISE, with mixed rank properties between H and CG. A noise-free variant is treated in [15].

3 Zero direct feedthrough

For $H = 0$ in (3), SISE from [5] is the recursion.

$$X_t = AP_{t-1}A^T + Q, \tag{4}$$
$$K_t = X_tC^T(CX_tC^T + R)^{-1}, \tag{5}$$
$$M_t = [G^TC^T(CX_tC^T + R)^{-1}CG]^{-1} \times G^TC^T(CX_tC^T + R)^{-1}, \tag{6}$$
$$P_t = (I - K_tC)\left[(I - GM_tC)X_t \times (I - GM_tC)^T + GM_tRM_t^TG^T\right] + K_tRM_t^TG^T, \tag{7}$$
$$\hat{d}_{t-1|t} = M_t(y_t - CA\hat{x}_{t-1|t-1}), \tag{8}$$
$$\hat{x}_{t|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t} + K_t \times (y_t - CA\hat{x}_{t-1|t-1} - CG\hat{d}_{t-1|t}), \tag{9}$$
$$\text{cov}(x_t|\mathbf{Y}^t) = P_t, \tag{10}$$

under the following structural condition.

Assumption 2

$$\textit{rank}\, CG = m. \tag{11}$$

An immediate observation is that SISE contains no specific information related to a model for the unmeasured disturbance d_t. Indeed, it is frequently claimed that signal $\{d_t : t = 0, 1, \ldots\}$ possesses no model whatsoever. Although, for bounded covariance X_t, i.e. when the algorithm is stable, the authors derived this version of SISE in [13] as a Kalman filter with $\{d_t\}$ modeled as a white noise process of unbounded variance. We shall return to this point later. Evidently, Assumption 2 requires $p \geq m$ and $\text{rank}\, C \geq \text{rank}\, G = m$. Firstly, we treat the square case, $p = m$, where the number of measurements equals the dimension of the disturbance input. Then we shall derive more general results.

3.1 Square zero-feedthrough case

From Assumption 2 when $p = m$, CG is invertible. Since, from (6), $M_tCG = I$ or $M_t = (CG)^{-1}$, we have

$$\hat{d}_{t-1|t} = (CG)^{-1}(y_t - CA\hat{x}_{t-1|t-1}), \tag{12}$$
$$0 = y_t - CA\hat{x}_{t-1|t-1} - CG\hat{d}_{t-1|t}, \tag{13}$$
$$\hat{x}_{t|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t}, \tag{14}$$
$$= [I - G(CG)^{-1}C]A\hat{x}_{t-1|t-1} + G(CG)^{-1}y_t. \tag{15}$$

This estimation algorithm:

- is time-invariant;
- does not depend on Q or R, the noise variances;
- is independent from the covariance calculations;
- has zero $\hat{x}_{t|t}$ innovations (13), (9).

SISE reduces to (12-15). Note that, using the matrix inversion lemma, we may rewrite the SISE y_t-to-$\hat{d}_{t-1|t}$ transfer function as

$$(CG)^{-1} - (CG)^{-1}CA(zI - A + G(CG)^{-1}CA)^{-1}G(CG)^{-1} = \left[CG + CA(zI - A)^{-1}G\right]^{-1}, = \left[zC(zI - A)^{-1}G\right]^{-1}, \tag{16}$$

which is the inverse of the one-step-advanced d_t-to-y_t transfer function $Z_d(z)$ in (1). The filtered state estimate error satisfies

$$\tilde{x}_{t|t} \triangleq x_t - \hat{x}_{t|t}, = [I - G(CG)^{-1}C]A\tilde{x}_{t-1|t-1} + [I - G(CG)^{-1}C]w_{t-1} - G(CG)^{-1}v_t.$$

The stability of SISE, i.e. the boundedness of the covariance of $\tilde{x}_{t|t}$, depends on the eigenvalues of $[I - G(CG)^{-1}C]A$.

Theorem 1 *For system (2-3) with $p = m$ and subject to Assumption 2, the eigenvalues of the SISE estimator system matrix, $\left[I - G(CG)^{-1}C\right] A$, lie at the transmission zeros of the square transfer function $zC(zI - A)^{-1}G$. Accordingly, the SISE estimator is asymptotically stable if and only if the transmission zeros of $T(z)$ all lie inside the unit circle.*

The proof of this theorem follows immediately from (16). An alternate is given in the Appendix for completeness and to establish connections to singular filtering. We note that condition (11) in Assumption 2 implies that $zC(zI-A)^{-1}G$ possesses exactly n finite transmission zeros with exactly m at zero.

We see that, in the square case, the poles of SISE can be located precisely at the transmission zeros of the d_t-to-y_t transfer function. SISE therefore is performing system inversion to recover $\hat{d}_{t-1|t}$ from $\mathbf{Y}^t$. The dependent recursion (14) for $\hat{x}_{t|t}$ is a simulation of the state equation (2) driven by $\hat{d}_{t-1|t}$. Effectively all the information in $\mathbf{Y}^t$ is used in generating the disturbance estimate, leaving simulation (14) to generate the state estimate.

When SISE is stable, it was shown in [13] that the state estimation algorithm implements a Kalman filter with a model for $\{d_t\}$ as a white noise of unbounded variance, D. In this case, the state estimation problem has driving noise variance $Q+GDG^T$ and measurement noise variance R. The identical filter, but not the covariances, will be achieved by taking driving noise GDG^T for finite D and $R \to 0$. That is, SISE is a singular filter. The connection to [14] in the proof is to the equivalent result in Loop Transfer Recovery for LQG control. When one selects $R = 0$, as opposed to $R \to 0$ from above, then the poles are placed at the transmission zeros. The limiting operation, on the other hand places the poles at the stable transmission zeros and the inverses of the unstable transmission zeros [11].

3.2 Non-square zero-feedthrough case

From Assumption 2, we take $p \geq m$ and make a transformation of the output signal as follows. This is a variation on the technique of [7]. Take the singular value decomposition of $p \times m$ CG.

$$\begin{aligned}\texttt{svd}(CG) &= U\Sigma V^T,\\ &= \begin{bmatrix} U_m \; U_{p-m}\end{bmatrix}\begin{bmatrix}\Sigma\\ 0\end{bmatrix} V^T.\end{aligned}$$

Define the $p \times p$ transformation

$$\mathcal{T} = \begin{bmatrix} U_m^T - U_m^T R U_{p-m}(U_{p-m}^T R U_{p-m})^{-1} U_{p-m}^T \\ U_{p-m}^T \end{bmatrix}, \quad (17)$$

and transform the original output signal, call it $\bar{y}_t$,

$$y_t = \mathcal{T}\bar{y}_t = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} x_t + \begin{bmatrix} v_{1,t} \\ v_{2,t}\end{bmatrix}, \quad (18)$$

yielding

$$\det C_1 G \neq 0, \quad C_2 G = 0, \quad \operatorname{cov}\begin{bmatrix} v_{1,t} \\ v_{2,t}\end{bmatrix} = \begin{bmatrix} R_1 & 0 \\ 0 & R_2\end{bmatrix}.$$

Theorem 2 *For system (2-3) with $p \geq m$ and subject to Assumptions 1 and 2, if and only if the pair $[A(I - G(C_1G)^{-1}C_1), C_2]$ is detectable then the filtered state covariance, P_t, is bounded and converges to a limit P_∞ as $t \to \infty$.*

The corresponding gain matrices, K_∞ and M_∞, yield the limiting SISE system matrix, $(I - K_\infty C)(I - GM_\infty C)A$, with all its eigenvalues strictly inside the unit circle.

The proof of this result appears in the Appendix and is based on proving that the state covariance satisfies a Riccati Difference Equation. Although this condition is not strictly the same as the condition in [7], the theorem condition implies theirs. Hence, their condition is also necessary. Theorem 2 similarly extends the condition in [16]. The sufficient stability result in [25] is predicated on P_t being bounded a priori. We already know from Theorem 1 the eigenvalues of $A(I - G(C_1G)^{-1}C_1)$ are stable if and only if $zC(zI - A)^{-1}G$ is minimum-phase.

We see that, when $p > m$, the surfeit of measurements beyond those strictly needed to produce $\hat{d}_{t-1|t}$ are brought to bear on estimating x_t. The stability of SISE depends on either the square case yielding stability via Theorem 1, i.e. via stable transmission zeros, or there being sufficient information in the additional measurements to stabilize the estimator.

When SISE is stable, then the algorithm implements a singular filter, as explained above. However, now the corresponding singular filter is *partially singular*, a term introduced in [26]. That is, the process noise variance is finite but the measurement noise variance, R, is less than full rank rather than zero. The approach of [26], under the banner of stable optimal filtering, in this case involves precisely a succession of a singular estimator and followed by a regular estimator, as in SISE. The result in [13] derives this stable (partially) singular filter when the plant satisfies the conditions of Theorem 2.

4 Nonzero direct feedthrough

When $H \neq 0$ in (3), SISE alters. Gillijns and De Moor [6] provide a SISE algorithm, subject to the following, for the time-invariant case.

Assumption 3 *Rank* $H = m$.

Subject to this assumption, the SISE formulation for time-invariant system (2-3) is

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t-1}, \tag{19}$$

$$P^x_{t|t-1} = \begin{bmatrix} A & G \end{bmatrix} \begin{bmatrix} P^x_{t-1|t-1} & P^{xd}_{t-1|t-1} \\ P^{dx}_{t-1|t-1} & P^d_{t-1|t-1} \end{bmatrix} \begin{bmatrix} A^T \\ G^T \end{bmatrix} + Q,$$

$$\tilde{R}_t = CP^x_{t|t-1}C^T + R,$$

$$M_t = (H^T\tilde{R}_t^{-1}H)^{-1}H^T\tilde{R}_t^{-1},$$

$$\hat{d}_{t|t} = M_t(y_t - C\hat{x}_{t|t-1}), \tag{20}$$

$$P^d_{t|t} = (H^T\tilde{R}_t^{-1}H)^{-1},$$

$$K_t = P^x_{k|k-1}C^T\tilde{R}_t^{-1},$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(y_t - C\hat{x}_{t|t-1} - H\hat{d}_{t|t}), \tag{21}$$

$$P^x_{t|t} = P^x_{t|t-1} - K_t(\tilde{R}_t - HP^dH^T)K_t^T,$$

$$P^{xd}_{t|t} = \left(P^{dx}_{t|t}\right)^T = -K_tHP^d_{t|t}.$$

When rank $H < m$, [7] provide ULISE, a carefully developed SISE algorithm which uses the singular value decomposition as in Subsection 3.2 but more widely to handle the more complicated interaction between filtered and smoothed estimates for d_t.

4.1 Square full-rank case

As with the $H = 0$ case, we consider first rank $H = m$ and $m = p$. That is H is invertible and, since $M_tH = I$, $M_t = H^{-1}$. Then SISE reduces to the recursion

$$\hat{d}_{t|t} = H^{-1}(y_t - C\hat{x}_{t|t-1})$$

$$0 = y_t - C\hat{x}_{t|t-1} - H\hat{d}_{t|t}, \tag{22}$$

$$\hat{x}_{t+1|t} = A\hat{x}_{t-1|t-1} + G\hat{d}_{t-1|t-1}, \tag{23}$$

$$= (A - GH^{-1}C)\hat{x}_{t|t-1} + GH^{-1}y_t,$$

$$\tilde{x}_{t+1|t} = (A - GH^{-1}C)\tilde{x}_{t|t-1} + w_t - GH^{-1}v_t.$$

Theorem 3 *For system (2-3) subject to Assumption 3, the eigenvalues of the SISE estimator system matrix, $A - GH^{-1}C$, lie at the transmission zeros of the square transfer function $H + C(zI - A)^{-1}G$. Accordingly, the SISE estimator is asymptotically stable if and only if these transmission zeros all lie inside the unit circle.*

Proof: Applying the matrix inversion lemma to the square transfer function between d_t and y_t,

$$\left[H + C(zI - A)^{-1}G\right]^{-1} = H^{-1} - H^{-1}C(zI - A + GH^{-1}C)^{-1}GH^{-1}.$$

The poles of the square direct feedthrough SISE lie at the transmission zeros of the d_t to y_t transfer function. □

Again, this result adds necessity to that of [7] in this case. Further, the result does not rely on optimality arguments. As in the square zero feedthrough case, the SISE estimator is time-invariant and independent from Q and R, and the state estimate filter innovations is zero. The condition rank $H = m$ ensures that all n transmission zeros are finite. We note again that the state innovations sequence (22) is zero and the filter (23) simulates $\hat{x}_{t+1|t}$ from $\hat{d}_{t|t}$.

4.2 Non-square full-rank case

The careful derivation of ULISE to accommodate rank $H \leq m$ is a central contribution of [7] and involves separation into subspaces. Take the singular value decomposition of $p \times m$ H possessing rank r.

$$\texttt{svd}(H) = U\Sigma V^T,$$

$$= \begin{bmatrix} U_r & U_{p-r} \end{bmatrix} \begin{bmatrix} \bar{H} & 0 \\ 0 & 0 \end{bmatrix} V^T.$$

Matrices take on the $(r, p - r)$ structure.

$$H = \begin{bmatrix} \bar{H} & 0 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \quad G = \begin{bmatrix} G_1 & G_2 \end{bmatrix},$$

$$K_t = \begin{bmatrix} K_{1,t} & K_{2,t} \end{bmatrix}, \quad M_t = \begin{bmatrix} M_{1,t} & M_{2,t} \end{bmatrix}.$$

As earlier in (17) and (18), define the $p \times p$ transformation

$$\mathcal{T} = \begin{bmatrix} U_r^T - U_r^TRU_{p-r}(U_{p-r}^TRU_{p-r})^{-1}U_{p-r}^T \\ U_{p-r}^T \end{bmatrix},$$

and transform the original output signal, call it $\bar{y}_t$,

$$y_t = \mathcal{T}\bar{y}_t = \begin{bmatrix} \bar{C}_1 \\ \bar{C}_2 \end{bmatrix} x_t + \begin{bmatrix} \bar{H} \\ 0 \end{bmatrix} d_t + \begin{bmatrix} \bar{v}_{1,t} \\ \bar{v}_{2,t} \end{bmatrix}.$$

yielding $\det \bar{H} \neq 0$ and $\operatorname{cov} \begin{bmatrix} \bar{v}_{1,t} \\ \bar{v}_{2,t} \end{bmatrix} = \begin{bmatrix} \bar{R}_1 & 0 \\ 0 & \bar{R}_2 \end{bmatrix}$. When rank $H = m$, $\bar{H}$ is $m \times m$ and we have the following result stemming from $M_{1,t}\bar{H} = I_m$.

Theorem 4 *Subject to Assumptions 1 and 3, $p \geq m$, SISE with feedthrough is stable if only if $[A - G\bar{H}^{-1}\bar{C}_1, \bar{C}_2]$ is detectable.*

The proof of this result is in the Appendix. This extends the detectability condition [1] of Theorem 5 of [7] to a necessary and sufficient condition for stability of SISE in this case. It

[1] Note that [7] uses p to denote our m.

also is the analog of Theorem 2 for the full-rank feedthrough case.

An alternative way to view necessity is to write the system matrix of SISE as

$$\begin{aligned} A - [AK_t + (G - AK_tH)M_t]C \\ = A - G\bar{H}^{-1}\bar{C}_1 \\ - (AK_{2,t} + GM_{2,,t} - AK_{1,t}\bar{H}M_{2,t})\bar{C}_2. \end{aligned}$$

For this matrix to be stable, a multiple of $\bar{C}_2$ must stabilize $A - G\bar{H}^{-1}\bar{C}_1$.

4.3 Less than full rank feedthrough

We build again on the decomposition above of [7] and make the following assumption.

Assumption 4 *rank* $\bar{C}_2 G_2 = m -$ *rank* $\bar{H}$.

This assumption guarantees that the transfer function $Z_d(z)$ in (1) has delay no greater than one and so the input might be reconstructed from filtered and one-step-smoothed data. In [7], the authors derive a sufficient condition for stability which we now extend to necessity.

Theorem 5 *Subject to Assumptions 1 and 4, general feedthrough SISE is stable if only if* $[A - G_1\bar{H}^{-1}\bar{C}_1, \bar{C}_2]$ *is detectable.*

This detectability condition is shown in [7] to be sufficient for stability by using the filter recursion for $\hat{x}_{t|t}$. If one calculates the alternative recursive prediction, $\hat{x}_{t|t-1}$, then it is evident that the ULISE system matrix is again of the form

$$\begin{aligned} (A - G_1\bar{H}^{-1}\bar{C}_1)(I - \tilde{L}_t\bar{C}_2)(I - G_2M_{2,t}\bar{C}_2) \\ = A - G_1\bar{H}^{-1}\bar{C}_1 + W_t\bar{C}_2, \end{aligned}$$

for appropriate W_t. Evidently, this can be stable only if the detectability condition holds.

The development in [7] is conducted in the time-varying case and uses corresponding time-varying decompositions and uniform detectability conditions [33]. To be fair, in this more challenging time-varying case, sufficiency of the stability condition is all that is conceivable. The restriction to time-invariant systems admits both the necessity and the possibility of using transfer functions to define transmission zeros, a connection made in [9].

5 Upshots

5.1 Stability and singular filtering

The preceding analysis provides necessary and sufficient conditions for the stability of linear SISE algorithms. Further, for the square cases, it yields the precise locations on the algorithm poles and demonstrates that the emphasis is on d_t-to-y_t system inversion to recover d_t followed by best efforts to estimate the state. We have pointed out the successive estimation nature of SISE, as have others. The question remains as to actions to be taken when SISE proves to be unstable, noting these central properties:

(i) SISE is stable when the d_t-to-y_t system is stably invertible.
(ii) When SISE is stable, it corresponds (at least in the zero feedthrough case) to a singular Kalman filter.
(iii) Subject to: detectability of $[A, C]$, stabilizability of $[A, Q^{\frac{1}{2}}]$, and $R \xrightarrow{+} 0$; this limiting Kalman filter is a stable estimator by construction.
(iv) The stability conditions for the limiting Kalman filter are more relaxed than Assumptions 2, 3 or 4, which pertain for the exact $R = 0$ singular filter.
(v) When SISE proves to be unstable, it (of course) differs from the Kalman filter.

5.2 Kalman filtering for input and state estimation

Denote the transfer function from d_t to y_t by $T(z)$. As derived in earlier sections, when $T(z)$ has all its transmission zeros inside the unit circle, then SISE is guaranteed stable and is equivalent to a specific stable singular Kalman filter. Compute a discrete-time inner-outer factorization [2] [28,29].

$$T(z) = T_o(z)T_i(z),$$

with $m \times m$ $T_i(z)$ an inner function, i.e. stable and all-pass, and $p \times m$ $T_o(z)$ an outer function, i.e. all transmission zeros inside the unit circle.

We note the following from the construction of the inner-outer factors.

Lemma 1 *If* $T(z)$ *has realization*

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + Gd_t + w_t, \\ y_t &= Cx_t + Du_t + Hd_t + v_t, \end{aligned}$$

then $T_o(z)$ *has realization*

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + \check{G}\check{d}_t + w_t, \\ y_t &= Cx_t + Du_t + \check{H}\check{d}_t + v_t, \end{aligned}$$

where $\check{d}_t$ *is the output of* $T_i(z)$.

That is, for the same initial conditions, $T_o(z)$ and $T(z)$ have the same states and have realizations which differ only in the G and H matrices. This factorization is depicted in Figure 2.

[2] Strictly speaking, this is a co-inner-co-outer factorization because of the ordering of T_i and T_o [27]. It can be obtained from the inner-outer factorization of the transpose of $T(z)$.

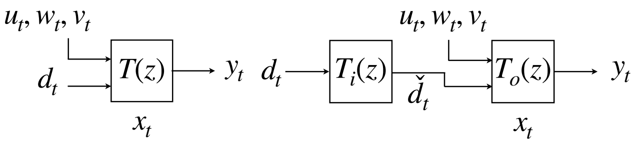

Fig. 2. System inner-outer factorization

Applying SISE or the singular/limiting Kalman filter for the outer function $T_o(z)$ to the signal $\{y_t\}$ yields estimates of $\{\check{d}_t\}$ and $\{x_t\}$ via a stable algorithm with guaranteed statistical and optimality properties. This follows since T_o is stably invertible by construction. Further, this stability depends on the standard assumptions above for Kalman filter stability. If one uses SISE, then depending on the delay properties of $T_o(z)$, i.e. its behavior as $z \to \infty$, a variation of the algorithm and Assumption 2, 3, or 4 might be needed to accommodate d_t-to-y_t invertibility. This is discussed further in Section 6.

To recover estimates for original system inputs $\{d_t\}$ from those for $\{\check{d}_t\}$ requires deconvolution (input estimation) without state estimation for the maximum-phase but stable system T_i. If delay is not an issue, then this can proceed stably via a fixed-interval smoother or reverse-time input estimation.

If the state estimates of x_t themselves are the objective, then the reconstruction of $\check{d}_t$ versus d_t is immaterial. This is the nature of the problem addressed in partially-known power system state estimation [24].

The singular filters derived by Shaked and co-authors [11,26,30] rely on the Return Difference Equality and spectral factorization for their calculation. In the case where the transmission zeros are unstable, the filter solution replaces them by their inverses akin to the inner-outer factorization.

The Kalman filter of [13] for $T_o(z)$ may be derived from the state-space model below with appropriate covariances,

$$\begin{bmatrix} x_{t+1} \\ \check{d}_{t+1} \end{bmatrix} = \begin{bmatrix} A & \check{G} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \check{d}_t \end{bmatrix} + \begin{bmatrix} I_n & 0 \\ 0 & I_m \end{bmatrix} \begin{bmatrix} w_t \\ \delta_t \end{bmatrix},$$
$$y_t = \begin{bmatrix} C & \check{H} \end{bmatrix} \begin{bmatrix} x_t \\ \check{d}_t \end{bmatrix} + v_t,$$

or using the direct construction as in [13], which avoids an explicit model for d_t but yields the same filter.

Marro and Zattoni [15] provide guidance on the recovery of the disturbance input signal when the d_t-to-y_t system is non-minimum-phase. Their approach involves the approximate inversion of this system using a long delay to accommodate the nominal instability of this inverse. Such techniques are reminiscent of those advanced in [31]. While the approach in [15] centers on state-estimation first, their development is geometric and noise free and so, it is unclear how this affects performance. Of course, the geometric analysis throws up the same initial reliance on minimum-phase zeros for stability and exact inversion.

5.3 *Extension to time-varying systems*

Developments so far have been limited to the time-invariant case and have availed themselves of concepts of transmission zeros, stable invertibility and inner-outer factorization, each of which is problematic to extend to time-varying systems. However, since alternative results have been phrased for the time-varying case, we consider this extension now, relying on examination of SISE recursions via Riccati difference equations in the proofs of Theorems 2 and 4.

Appealing to [5,6] for the time-varying SISE algorithms in the case of Theorem 2 and zero direct feedthrough, Riccati equation (28) becomes

$$X_{t+1} = \bar{A}_t X_t \bar{A}_t^T - \bar{A}_t X_t C_{2,t}^T (C_{2,t} X_t C_{2,t}^T + R_{2,t})^{-1} \times (\bar{A}_t X_t C_{2,t}^T)^T + \bar{Q}_t,$$

where,

$$\bar{A}_t = A_t (I - G_{t-1} (C_{1,t} G_{t-1})^{-1} C_{1,t}),$$
$$\bar{Q}_t = A_t G_{t-1} (C_{1,t} G_{t-1})^{-1} R_{1,t} (G_{t-1} (C_{1,t} G_{t-1})^{-1})^T A_t^T + Q_t,$$

and, in the case of full-rank feedthrough, (30) becomes

$$X_{t+1} = \hat{A}_t X_t \hat{A}_t^T - (\hat{A}_t X_t \bar{C}_{2,t}^T)(\bar{C}_{2,t} X_t \bar{C}_{2,t} + \bar{R}_{2,t})^{-1} \times (\hat{A}_t X_t \bar{C}_{2,t}^T)^T + \hat{Q}_t,$$

where,

$$\hat{A}_t = A_t - G_t \bar{H}_t^{-1} \bar{C}_{1,t}, \quad \hat{Q}_t = Q_t + G_t \bar{H}_t^{-1} \bar{R}_{1,t} \bar{H}_t^{-T} G_t^T.$$

with now time-varying quantities $\{A_t, G_t, \dots, \}$. We may appeal to standard sufficient results, e.g. [32,19] and Theorem 5.3 in [33], on the exponential stability of the Kalman filter subject to uniform reachability and detectability. Subject to the uniform satisfaction of time-varying equivalents of Assumptions 1, 2 and/or 3 as appropriate, this extends these stability conditions to the uniformly time-varying case. We note that the development of ULISE in [7] also is time-varying and draws on precisely these techniques in [33] to establish sufficient conditions for stability.

5.4 *Extension to more complicated delay structure*

Both [8] and [9] study delay structures for the d_t-to-y_t system which are more detailed than Assumption 2-plus-$\{H = 0\}$ or Assumption 3. They consider reconstruction of smoothed disturbance estimates $\hat{d}^{\ell}_{t-\ell|t}$ for the successive components of d_t accessible with increasing smoothing

$\ell = 0, 1, \ldots, L$. The algorithms are developed in the time-invariant case, iterate over the delay using singular value decompositions and have stability properties tied to the transmission zeros. The inner-outer decomposition above could also be applied to these algorithms to ensure estimator stability.

6 System Inversion and SISE

For the square cases of SISE satisfying Assumptions 1 or 2, we were able to demonstrate that the SISE d_t-estimator implements exactly the left inverse of the y_t-to-d_t system. The simultaneous x_t-estimate is the state of the inverse system, the stability of which depends on the transmission zeros of the original system.

Conditions for left invertibility of a linear time-invariant system are provided by Sain and Massey [34] and for stable invertibility by Moylan [35] via the Rosenbrock system matrix. Both papers construct the inverse system. Moreover, in [34,36], left invertibility with delay, L, is studied, where stacked measurements $\begin{bmatrix} y_t^T & y_{t+1}^T & \ldots & y_{t+L}^T \end{bmatrix}^T$ are used to estimate d_t. Marro and Zattoni blend into this picture stable approximate inversion with delay.

From [34], we see that, for any $p \geq m$:

- Assumption 2 is the left invertibility condition for $C(zI - A)^{-1}G$ with delay one.
- Assumption 3 is the left invertibility condition for $H + C(zI - A)^{-1}G$ with delay zero.
- Assumption 4 is the left invertibility condition for $H + C(zI - A)^{-1}G$ with delay one.

Given the input recovery objective of SISE, this is not surprising. But it is interesting to tie these ideas more closely. It is reasonable to expect that related connections to invertibility and the delay structure of the d_t-to-y_t system in the time-invariant case should hold for the more complicated circumstances considered in [8,9], where estimator stability properties is also proven tied to the transmission zeros.

It is worth remarking that many presentations of SISE algorithms make connections to 'unbiasedness' and 'optimality' of the state estimate. As [15,13] demonstrate, the probabilistic concept of unbiasedness is really tied to a geometric property of the algorithms and the nature of certain subspaces. The optimality of the state estimates is within the class of estimators already satisfying the geometric constraints. As is evident from, say, Theorems 1 and 2 and the Riccati equation proof, there is no degree of freedom left for the state estimator in the square case and limited degrees of freedom in the non-square case. Indeed, in the square cases, (14-23) show that $\hat{x}_{t|t}$ is computed by system simulation using the estimated input; the measurements play no further part. The detectability conditions on Theorems 2 and 4 show how the remaining degrees of freedom are used in the Riccati difference equations (28) and (30).

In conclusion, the paper attempts to unify the collection of SISE algorithms by revealing their explicit connections to system inversion to recover the otherwise unmodeled disturbance input d_t followed by their 'best efforts' subsequent estimation of the state x_t. The result has been to develop necessary and sufficient conditions for stability, at least in the linear time-invariant case, in terms of the transmission zeros of the d_t-to-y_t plant and then the detectability of the subsequent state estimator. As already stated, the prospect for extension of necessity conditions to the time-varying case is dim and the interpretation via zeros problematic.

Acknowledgement

The authors are grateful for the care in reading and careful insights provided by the reviewers.

Appendix

Proof of Theorem 1

We loosely follow a calculation from Maciejowski [14]. From (15) and the system equations the (time-invariant) transfer function from d_t to $\hat{x}_{t|t}$ via y_t is given by

$$\begin{aligned}\Psi(z) &= \left\{zI - \left[I - G(CG)^{-1}C\right]A\right\}^{-1} zG(CG)^{-1} \\ &\qquad \times C(zI - A)^{-1}G, \\ &= \left\{zI - \left[I - \Pi\right]A\right\}^{-1} z\Pi(zI - A)^{-1}G,\end{aligned} \tag{24}$$

where we have used $\Pi \triangleq G(CG)^{-1}C$. Write

$$\begin{aligned}&[zI - (I - \Pi)A]^{-1} z\Pi \\ &\quad = [zI - (I - \Pi)A]^{-1} [z\Pi - zI + (I - \Pi)A] + I, \\ &\quad = -[zI - (I - \Pi)A]^{-1} (I - \Pi)(zI - A) + I.\end{aligned}$$

Then, since $(I - \Pi)G = 0$,

$$\begin{aligned}\Psi(z) = \\ &- [zI - (I - \Pi)A]^{-1} (I - \Pi)(zI - A)(zI - A)^{-1}G \\ &\qquad + (zI - A)^{-1}G, \\ &= -[zI - (I - \Pi)A]^{-1} (I - \Pi)G + (zI - A)^{-1}G, \\ &= (zI - A)^{-1}G.\end{aligned} \tag{25}$$

From (24), $\Psi(z)$ is the product of two transfer functions and nominally should have $2n$ poles; those at the eigenvalues of A and those at the eigenvalues of $(I - \Pi)A$. The transfer function $zC(zI - A)^{-1}G$ has McMillan degree n with n finite transmission zeros. We see from (25) that only poles at the eigenvalues of A are present in Ψ. This implies that the poles due to the eigenvalues of $(I - \Pi)A$ cancel the transmission zeros of $zC(zI - A)^{-1}G$.

Proof of Theorem 2

Define the following quantities.

$$\mathcal{H}_t = GM_t - K_tCGM_t + K_t, \quad Z = C_1G,$$
$$Y = (CX_tC^T + R)^{-1} = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix},$$

where Y is divided conformably with C and v_t in (18).

From (8) and (9) the filtered prediction error satisfies

$$\begin{aligned}\tilde{x}_t &\triangleq x_t - x_{t|t} \\ &= (I - \mathcal{H}_tC)A\tilde{x}_{t-1} + (I - \mathcal{H}_tC)w_{t-1} - \mathcal{H}_tv_t.\end{aligned}$$

Whence,

$$\begin{aligned}P_{t|t} &= \text{cov}(x_t|\mathbf{Y_t}) \\ &= (I - \mathcal{H}_tC)(AP_{t-1|t-1}A^T + Q)(I - \mathcal{H}_tC)^T \\ &\quad + \mathcal{H}_tR\mathcal{H}_t^T.\end{aligned}$$

Using (4) yields

$$X_{t+1} = A\left((I - \mathcal{H}_tC)X_t(I - \mathcal{H}_tC)^T + \mathcal{H}_tR\mathcal{H}_t^T\right)A^T + Q. \qquad (26)$$

We show that this discrete Lyapunov equation is also a Riccati difference equation by substituting for $\mathcal{H}_t$ using $CG = \begin{bmatrix} Z^T & 0 \end{bmatrix}^T$.

$$\begin{aligned}\mathcal{H}_t &= G\left(\begin{bmatrix} Z^T & 0 \end{bmatrix} Y \begin{bmatrix} Z \\ 0 \end{bmatrix}\right)^{-1} \begin{bmatrix} Z^T & 0 \end{bmatrix} Y - X_t \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} \\ &\quad Y \begin{bmatrix} Z \\ 0 \end{bmatrix} \left(\begin{bmatrix} Z^T & 0 \end{bmatrix} Y \begin{bmatrix} Z \\ 0 \end{bmatrix}\right)^{-1} \begin{bmatrix} Z^T & 0 \end{bmatrix} Y \\ &\quad + X_t \begin{bmatrix} C_1^T & C_2^T \end{bmatrix} Y \\ &= \begin{bmatrix} GZ^{-1} & GZ^{-1}Y_1^{-1}Y_2 \end{bmatrix} + X_t \begin{bmatrix} 0 & C_2(Y_3 - Y_2^TY_1^{-1}Y_2) \end{bmatrix}. \qquad (27)\end{aligned}$$

Using partitioned matrix inversion with Y gives

$$\begin{aligned}(Y_3 - Y_2^TY_1^{-1}Y_2) &= (C_2X_tC_2^T + R_2)^{-1} \\ Y_1^{-1}Y_2 &= -(C_1X_tC_2^T)(C_2X_tC_2^T + R_2)^{-1}.\end{aligned}$$

Substituting this into (27) and (26) gives the following Riccati difference equation.

$$X_{t+1} = \bar{A}X_t\bar{A}^T - (\bar{A}X_tC_2^T)(C_2X_tC_2^T + R_2)^{-1} \times (\bar{A}X_tC_2^T)^T + \bar{Q}, \quad (28)$$

where,

$$\begin{aligned}\bar{A} &= A(I - G(C_1G)^{-1}C_1), \\ \bar{Q} &= AG(C_1G)^{-1}R_1(G(C_1G)^{-1})^TA^T + Q. \qquad (29)\end{aligned}$$

Appealing to Theorem 14.3.1 [37] (p. 510), provided $[\bar{A}, \bar{Q}^{\frac{1}{2}}]$ is stabilizable and $[\bar{A}, C_2]$ is detectable, then X_t converges to the maximal solution of the algebraic Riccati equation, which is stabilizing.

Now, since by assumption $[A, Q^{\frac{1}{2}}]$ is stabilizable, there exists a $\mathcal{K}$ such that $A - Q^{\frac{1}{2}}\mathcal{K}$ is stable. Taking,

$$\bar{Q}^{\frac{1}{2}} = \begin{bmatrix} Q^{\frac{1}{2}} & AG(C_1G)^{-1}R_1^{\frac{1}{2}} \end{bmatrix},$$

and $\bar{\mathcal{K}} = \begin{bmatrix} \mathcal{K}^T & C_1^TR_1^{-\frac{T}{2}} \end{bmatrix}^T$,

$$\begin{aligned}\bar{A} - \bar{Q}^{\frac{1}{2}}\bar{\mathcal{K}} &= A(I - G(C_1G)^{-1}C_1) \\ &\quad - \begin{bmatrix} Q^{\frac{1}{2}} & AG(C_1G)^{-1}R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathcal{K} \\ R^{-\frac{1}{2}}C_1 \end{bmatrix}, \\ &= A - Q^{\frac{1}{2}}\mathcal{K}.\end{aligned}$$

Thus, $\bar{A} - \bar{Q}^{\frac{1}{2}}\bar{\mathcal{K}}$ also is stable by construction. So stabilizability of $[A, Q^{\frac{1}{2}}]$ implies stabilizability of $[\bar{A}, \bar{Q}^{\frac{1}{2}}]$.

Proof of Theorem 4

The proof parallels that of Theorem 2. Substitute (20) and (21) into (19) to yield

$$\hat{x}_{t+1|t} = (A - \mathcal{L}_tC)\hat{x}_{t|t-1} + \mathcal{L}_ty_t,$$

where $\mathcal{L}_t = AK_t - AK_tHM_t + GM_t$. Then

$$\begin{aligned}\tilde{x}_{t+1|t} &\triangleq x_t - \hat{x}_{t+1|t}, \\ &= (A - \mathcal{L}_tC)\tilde{x}_{t|t-1} + w_t - \mathcal{L}_tv_t, \\ X_{t+1} &= (A - \mathcal{L}_tC)X_t(A - \mathcal{L}_tC)^T + \mathcal{L}_tR\mathcal{L}_t^T + Q,\end{aligned}$$

with $X_{t+1} \triangleq \text{cov}(x_{t+1}|\mathbf{Y}^t)$. Dividing K_t and M_t conformably with C^T: $K_t = \begin{bmatrix} K_{1,t} & K_{2,t} \end{bmatrix}$, $M_t = \begin{bmatrix} M_{1,t} & M_{2,t} \end{bmatrix}$, one arrives directly at the following Riccati difference equation.

$$X_{t+1} = \hat{A}X_t\hat{A}^T - \hat{A}X_t\bar{C}_2^T(\bar{C}_2X_t\bar{C}_2 + \bar{R}_2)^{-1}\bar{C}_2X_t\hat{A}^T + \hat{Q}, \qquad (30)$$

where,

$$\hat{A} = A - G\bar{H}^{-1}\bar{C}_1, \quad \hat{Q} = Q + G\bar{H}^{-1}\bar{R}_1\bar{H}^{-T}G^T.$$

The proof follows as that for Theorem 2 using [37].

References

[1] N. Wiener. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. Wiley, New York, 1949.

[2] Jerry M. Mendel and C.S. Burrus (eds). *Maximum-Likelihood Deconvolution*. Signal Processing & Digital Filtering. Springer Verlag, 1990.

[3] Wail A. Mousa. *Advanced Digital Signal Processing og Seismic Data*. Cambridge University Press, 2020.

[4] P.K. Kitanidis. Unbiased minimum-variance linear state estimation. *Automatica*, 23(6):775–778, 1987.

[5] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time sysems. *Automatica*, 43:111–116, 2007.

[6] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time sysems with direct feedthrough. *Automatica*, 43:934–937, 2007.

[7] Sze Zhong Yong, Minghui Zhu, and Emilio Frazzoli. A unified filter for simultaneous input and state estiamtion of linear discrete-time stochastic systems. *Automatica*, 63:321–329, 2016.

[8] S. Sundaram and C. N. Hadjicostis. Partial state observers for linear systems with unknown inputs. *Automatica*, 44(12):3126–3132, 2008.

[9] S.Z. Yong, M. Zhu, and E. Frazzoli. Simultaneous input and state estimation with a delay. In *Proceedings of the 54th IEEE Conference on Decision and Control*, pages 468–475, Osaka Japan, 2015.

[10] J.M. Mendel. White-noise estimators for seismic data processing in oil exploration. *IEEE Transactions on Automatic Control*, AC-22(5):694–706, October 1977.

[11] Uri Shaked. Explicit solution to the singular discrete-time stationary linear filtering problem. *IEEE Transactions on Automatic Control*, AC-30(1):34–47, 1985.

[12] K.J. Åström and B. Wittenmark. *Computer Controlled Systems: Theory and Design - Third Edition*. Prentice-Hall, Upper Saddle River, New Jersey, 1997.

[13] R.R. Bitmead, M. Hovd, and M.A. Abooshahab. A Kalman-filtering derivation of simultaneous input and state estimation. *Automatica*, 108(108478), 2019.

[14] J. M. Maciejowski. Asymptotic recovery for discrete-time systems. *IEEE Transactions on Automatic Control*, AC-30:602–605, 1985.

[15] Giovanni Marro and Elena Zattoni. Unknown-state, unknown-input reconstruction in discrete-time nonminimum-phase systems: geometric methods. *Automatica*, 46:815–822, 2010.

[16] Huazhen Fang, Yang Shi, and Jingang Yi. On stable simultaneous input and state estimation for discrete-time linear systems. *International Journal of Adaptive Control & Signal Processing*, 25:671–686, 2011.

[17] Huazhen Fang and Raymond A. de Callafon. On the asymptotic stability of minimum-variance unbiased input and state estimation. *Automatica*, 48:3183–3186, 2012.

[18] Bao-Zhu Guo and Zhi-Liang Zhao. *Active disturbance rejection control for nonlinear systems: an introduction*. John Wiley & Sons (Asia), Singapore, 2016.

[19] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[20] D. Simon. *Optimal State Estimation: Kalman, H_∞, and nonlinear approaches*. John Wiley & Sons, Hoboken NJ, 2006.

[21] J.D. Glover. The linear estimation of completely unknown signals. *IEEE Transactions on Automatic Control*, AC-14(6):766–769, 1969.

[22] P. Sanyal and C.N. Shen. Bayes' decision rule for rapid detection and adaptive estimation scheme with space applications. *IEEE Transactions on Automatic Control*, AC-19(3):228–231, 1974.

[23] Huazhen Fang, R.A. de Callafon, and J. Cortés. Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation. *Automatica*, 49:2805–2812, 2013.

[24] Mohammad Ali Abooshahab, Morten Hovd, and Robert R. Bitmead. Disturbance and state estimation in partially known power networks. In *IEEE Conference on Control Technology and its Application*, Hong Kong, August 2019.

[25] Hunmin Kim, Pinyao Guo, Minghui Zhu, and Peng Liu. Simultaneous input and state estimation for stochastic nonlinear systems with additive unknown inputs. *Automatica*, 111:108588, 2020.

[26] B. Priel and U. Shaked. Sequential decomposition of the partially singular discrete time filtering problem. In *Proceedings of the 25th IEEE Conference on Decision and Control*, pages 1048–1053, Athens, Greece, December 1986.

[27] Bruce A. Francis. *A course in H_∞ theory*. Number 88 in LNCIS. Springer-Verlag, Berlin, 1987.

[28] Michael Green. On inner-outer factorization. *Systems & Control Letters*, 11:93–97, 1988.

[29] Vlad Ionecsu and Cristian Oară. Spectral and inner-outer factorizations for discrete-time systems. *IEEE Transactions on Automatic Control*, 41(12):1840–1845, 1996.

[30] U. Shaked and E. Soroka. A simple solution to the singular linear minimum-variance estimation problem. *IEEE Transactions on Automatic Control*, AC-32(1):81–84, 1987.

[31] B. Widrow and E. Wallach. *Adaptive Inverse Control*. Prentice-Hall Inc, Englewood Cliffs, NJ, 1994.

[32] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, Inc., New York, 1970.

[33] B.D.O. Anderson and J.B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM J. Control Optimization*, 19(1):20–32, 1981.

[34] Michael K. Sain and James L. Massey. Invertibility of linear time-invariant dynamical systems. *IEEE Transactions on Automatic Control*, AC-14(2):141–149, 1969.

[35] P.J. Moylan. Stable inversion of linear systems. *IEEE Transactions on Automatic Control*, 22(1):74–78, February 1977.

[36] Shreyas Sundaram. Inversion of linear systems. *https://engineering.purdue.edu/∼sundara2/talks/invert.pdf*, accessed June 2020.

[37] T. Kailtah, A. Sayed, and B. Hassibi. *Linear Estimation*. Prentice-Hall, Upper Saddle River, NJ, USA, 2000.