15th Triennial World Congress, Barcelona, Spain

RISK-SENSITIVE STATE ESTIMATION FOR SINGULAR SYSTEMS [1]

Huanshui Zhang * **Lihua Xie** ** **Yeng Chai Soh** ** **Wei Wang** *

* *The Research Center of Information and Control, Dalian University of Technology, Dalian, P. R. China 116023*
e-mail: hszhang@hotmail.com
** *BLK S2, School of Electrical and Electronic Engineering Nanyang Technological University, Singapore 639798*

Abstract: This paper is concerned with finite horizon risk-sensitive filtering, prediction and smoothing problem for discrete-time singular systems. The problem is first converted to a minimax optimization of certain indefinite quadratic form. It is shown that a risk-sensitive estimator can be obtained by ensuring the minimum of the indefinite quadratic form to be maximum (minimum) when the risk-sensitivity parameter θ is negative (positive). An auxiliary state-space signal model and innovation sequences in Krein space are introduced to simplify our derivation. The finite horizon estimator is given based on a recursive Riccati equation by constructing a appropriate state space model.

Keywords: Risk-sensitive; estimation; Singular systems; Discrete-time ; Krein space.

1. INTRODUCTION

The analysis and design of linear singular systems have a rich and growing literature (Lewis and Mertzios, 1989, Cobb, 1989) in the last few years. Part of the motivations for this activity comes from applications arising from robotic, economic, electric and chemical systems (Lewis and Mertzios, 1989). More recently, recursive state estimation for singular systems has been the subject of several studies (Dai, 1989, Nikoukhah *et al.*, 1992, Zhang *et al.*, 1998). In (Dai, 1989), the singular system under consideration is first transformed into a normal form via state augmentation and then the filtering problem is solved using standard results for nonsingular systems. A general formulation of a discrete-time filtering problem for singular systems has been given by Nikoukhah *et al* (1992). By applying a "dual approach" to estimation, a so-called "3-block" form for the optimal filter is derived based on a Riccati difference equation of a 3-block form.

Using an innovation analysis method in time domain together with an output predictor, Zhang *et al* (1998) proposed a unified approach for filtering, smoothing and prediction problems for singular systems. The estimators are derived by using an ARMA innovation model in Hilbert space and calculated based on one spectral factorization.

It should be noted that the aforementioned work concentrated on the H_2 state estimation. A more general estimation problem is to minimize an ex-

[1] This work is supported by National Nature Science Foundation of P. R. China (60174017)

ponential function of the squared filtering error, or its expectation, thus penalizing all the higher order moments of the estimation error energy. This problem is termed a *risk-sensitive filtering problem.* The risk-sensitive performance was first used by Jacobson (1973) and pursued further by Whittle (1990). As opposed to the H_2 filtering (risk-neutral filtering) which minimizes a quadratic error criterion, risk-sensitive filtering robustifies the filter against plant and noise uncertainties. It is worth pointing out that all existing results on the risk sensitive estimation have been focused on nonsingular systems.

In this paper we shall consider the finite horizon risk-sensitive filtering, prediction and smoothing problems for discrete-time singular systems. The risk-sensitive estimation is then shown to be equivalent to a minimax optimization of an indefinite quadratic form. To derive the risk-sensitive estimator through the minimax optimization, an appropriate auxiliary stochastic state-space model and the innovation sequence in Krein space are introduced. The finite horizon risk sensitive estimator is given in terms of one recursive Riccati equation through state augmentation.

We end this section by recalling the definition of Krein space (Hassibi *et al.*, 1998).

Definition An abstract vector space $\{\mathcal{K}, \langle\cdot,\cdot\rangle\}$ that satisfies the following requirements is called a Krein Space:
1) $\mathcal{K}$ is a linear space over $\mathbf{C}$, the set of complex numbers.
2) There exists a bilinear form $\langle\cdot,\cdot\rangle \in \mathbf{C}$ on $\mathcal{K}$ such that
a) $\langle\mathbf{y},\mathbf{x}\rangle = \langle\mathbf{x},\mathbf{y}\rangle^*$
b) $\langle a\mathbf{x}+b\mathbf{y},\mathbf{u}\rangle = a\langle\mathbf{x},\mathbf{u}\rangle + b\langle\mathbf{y},\mathbf{u}\rangle$
3) The vector space $\mathcal{K}$ admits a direct orthogonal sum decomposition

$$\mathcal{K} = \mathcal{K}_+ \oplus \mathcal{K}_- \tag{1}$$

such that $\{\mathcal{K}_+, \langle\cdot,\cdot\rangle\}$ and $\{\mathcal{K}_-, \langle\cdot,\cdot\rangle\}$ are Hilbert spaces, and $\langle\mathbf{x},\mathbf{y}\rangle = 0$, for any $\mathbf{x} \in \mathcal{K}_+$ and $\mathbf{y} \in \mathcal{K}_-$.

It should be noted that Hilbert spaces satisfy not only 1) and 2), but also the requirement that $\langle\mathbf{x},\mathbf{x}\rangle > 0$, for $\mathbf{x} \neq 0$. In Krein Space we have $\mathbf{x} \neq 0$ such that $\langle\mathbf{x},\mathbf{x}\rangle = 0$ or $\langle\mathbf{x},\mathbf{x}\rangle < 0$.

Whenever a Krein Space element and a Euclidean space element satisfy the same set of constrains, we shall denote them by the same letter with the former being bold and the latter being normal.

2. PROBLEM STATEMENT

Consider a stochastic linear time-invariant system described by the following discrete-time model:

$$M\mathbf{x}(t+1) = \Phi\mathbf{x}(t) + \Gamma\mathbf{e}(t) \tag{2}$$
$$\mathbf{y}(t) = H\mathbf{x}(t) + \mathbf{v}(t) \tag{3}$$
$$\mathbf{z}(t) = L\mathbf{x}(t) \tag{4}$$

where $\mathbf{x}(t) \in R^n, \mathbf{e}(t) \in R^r, \mathbf{y}(t) \in R^m$, $\mathbf{v}(t) \in R^m$ and $\mathbf{z}(t) \in R^s$ represent the state, system stochastic noise, measurement output, measurement noise and the signal to be estimated, respectively. It is assumed that $\mathbf{e}(t)$ and $\mathbf{v}(t)$ are zero mean mutually uncorrelated Gaussian white noises with $\langle\mathbf{e}(k),\mathbf{e}(j)\rangle = E[\mathbf{e}(k)\mathbf{e}^T(j)] = Q_e\delta_{kj}$ and $\langle\mathbf{v}(k),\mathbf{v}(j)\rangle = E[\mathbf{v}(k)\mathbf{v}^T(j)] = Q_v\delta_{kj}$, where E is the mathematical expectation, δ_{kj} is the Kronecker delta and T denotes the transpose. We assume that $Q_v > 0$. The following assumptions will be made for (2)-(4).

Assumption 2.1 *M is singular but the system (2) is regular, i.e. $det(zM - \Phi) \not\equiv 0$.*

In this paper we consider a risk-sensitive estimation problem for the system (2)-(4). More precisely, the problem is stated as follows:

Finite Horizon Risk-Sensitive Estimation: Given a non-zero real scalar θ and the observation $\{\mathbf{y}(i)\}_{i=0}^t$, find an estimate of $\mathbf{z}(t+l) = L\mathbf{x}(t+l)$, denoted as $\check{\mathbf{z}}(t+l \mid t)$, $t = 0, 1, \cdots, N$, such that the following cost is minimized:

$$\min_{\{\check{\mathbf{z}}(t+l|t)\}} \left(-\frac{2}{\theta}log\left[Eexp\left(-\frac{\theta}{2}\mathbf{D}_N\right)\right]\right) \tag{5}$$

where $\mathbf{D}_N$ is given by

$$\mathbf{D}_N = \sum_{t=0}^{N} [\check{\mathbf{z}}(t+l \mid t) - L\mathbf{x}(t+l)]^T \times [\check{\mathbf{z}}(t+l \mid t) - L\mathbf{x}(t+l)] \tag{6}$$

Observe that the above estimation problems include three cases, i.e. $l = 0$, $l > 0$ and $l < 0$ which correspond to the cases of filtering, prediction and smoothing, respectively.

3. PRELIMINARIES

In this section, we shall first transform the singular signal model (2)-(3) into a nonsingular but non-causal signal model. We shall also establish an equivalence between the risk-sensitive estimation problem and a minimax optimization.

First, under Assumption 2.1, there exist nonsingular matrices Q_1 nd P_1 such that (Cobb, 1984)

$$Q_1MP_1 = \begin{bmatrix} I_{n_1} & 0 \\ 0 & M_1 \end{bmatrix}, \quad Q_1\Phi P_1 = \begin{bmatrix} \Phi_1 & 0 \\ 0 & I_{n_2} \end{bmatrix} \tag{7}$$

where $n_1 + n_2 = n$, M_1 is a nilpotent matrix with index λ_0, i.e. $M_1^{\lambda_0} = 0$, $M_1^{\lambda_0 - 1} \neq 0$.

The system (2)-(3) is restricted system equivalent (RSE) to (Zhang *et al.*, 1998, Cobb, 1984):

$$\bar{\mathbf{x}}_1(t+1) = \Phi_1 \bar{\mathbf{x}}_1(t) + \Gamma_1 \mathbf{e}(t) \tag{8}$$

$$M_1 \bar{\mathbf{x}}_2(t+1) = \bar{\mathbf{x}}_2(t) + \Gamma_2 \mathbf{e}(t) \tag{9}$$

$$\mathbf{y}(t) = H_1 \bar{\mathbf{x}}_1(t) + H_2 \bar{\mathbf{x}}_2(t) + \mathbf{v}(t) \tag{10}$$

where $\bar{\mathbf{x}}_1(t) \in R^{n_1}$ and

$$\bar{\mathbf{x}}(t) = \begin{bmatrix} \bar{\mathbf{x}}_1(t) \\ \bar{\mathbf{x}}_2(t) \end{bmatrix} = P_1^{-1} \mathbf{x}(t),$$

$$\begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} = Q_1 \Gamma, \quad [\, H_1 \quad H_2 \,] = H P_1$$

It follows from (9) that

$$\bar{\mathbf{x}}_2(t+1) = -\Gamma_2^{(1)} \mathbf{e}(t+1) - \Gamma_2^{(2)} \mathbf{e}(t+2) - \cdots - \Gamma_2^{(\lambda_0)} \mathbf{e}(t+\lambda_0) \tag{11}$$

where

$$\Gamma_2^{(i)} = M_1^{i-1} \Gamma_2, \quad i = 1, 2, \cdots, \lambda_0 \tag{12}$$

Hence, we have the following result.

Lemma 3.1 Under Assumption 2.1, the system (2) and (3) is RSE to:

$$\bar{\mathbf{x}}(t+1) = \bar{\Phi} \bar{\mathbf{x}}(t) + \bar{\Gamma}(q) \mathbf{e}(t) \tag{13}$$

$$\mathbf{y}(t) = \bar{H} \bar{\mathbf{x}}(t) + \mathbf{v}(t) \tag{14}$$

where

$$\bar{\Phi} = \begin{bmatrix} \Phi_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \bar{H} = [\, H_1 \quad H_2 \,],$$

$$\bar{\Gamma}(q) = \Gamma^{(0)} + \Gamma^{(1)} q + \cdots + \Gamma^{(\lambda_0)} q^{\lambda_0} \tag{15}$$

$$\Gamma^{(0)} = \begin{bmatrix} \Gamma_1 \\ 0 \end{bmatrix}, \quad \Gamma^{(i)} = - \begin{bmatrix} 0 \\ \Gamma_2^{(i)} \end{bmatrix} \quad (1 \leq i \leq \lambda_0) \tag{16}$$

and q is the forward shift operator, i.e. $q\mathbf{e}(t) = \mathbf{e}(t+1)$.

Proof: The result follows trivially from (8) and (11).

In the remainder of this section, we shall demonstrate that the risk-sensitive estimation problem can be converted into a minimax optimization problem.

For a given l, let $l_0 = max\{l + \lambda_0 - 1, \lambda_0 - 1\}$. Also, denote

$$e_t = col\{e(0), e(1), \cdots, e(t)\}$$

$$y_t = col\{y(0), y(1), \cdots, y(t)\} \tag{17}$$

for any $t \geq 0$.

We recall the following lemma (Hassibi *et al.*, 1998, Whittle, 1990).

Lemma 3.2 The risk-sensitive filtering problem is reduced to one of finding the estimator $\check{z}(t+l \mid t)$ such that

a) $\theta > 0 (risk - seeking)$:

$$\min_{\{\check{z}(t+l|t)\}} \left\{ \min_{e_{N+l_0}, \bar{x}_1(0)} J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N) \right\} \tag{18}$$

b) $\theta < 0 (risk - averse)$:

$$\max_{\{\check{z}(t+l|t)\}} \left\{ \min_{e_{N+l_0}, \bar{x}_1(0)} J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N) \right\} \tag{19}$$

where

$$\begin{aligned} & J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N) = \bar{x}_1(0) \Pi_1^{-1} \bar{x}_1^T(0) \\ & + \sum_{t=0}^{N+l_0} e^T(t) Q_e^{-1} e(t) + \sum_{t=0}^{N} [y(t) - \bar{H}\bar{x}(t)]^T Q_v^{-1} \\ & \times [y(t) - \bar{H}\bar{x}(t)] \\ & + \theta \sum_{t=0}^{N} [\check{z}(t+l \mid t) - \bar{L}\bar{x}(t+l)]^T [\check{z}(t+l \mid t) - \bar{L}\bar{x}(t+l)] \end{aligned} \tag{20}$$

with

$$\bar{L} = L P_1 \tag{21}$$

4. RISK-SENSITIVE FILTERING, PREDICTION AND SMOOTHING

In order to obtain an estimator which solves the optimization problem in (18) or (19), we first consider the problem of calculating $\min_{e_{N+l_0}, \bar{x}_1(0)} J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$. Observe from (20) that $J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$ can be rewritten as

$$\begin{aligned} & J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N) \\ & = \bar{x}_1^T(0) \Pi_1^{-1} \bar{x}_1(0) + \sum_{t=0}^{N+l_0} e^T(t) Q_e^{-1} e(t) + \\ & \sum_{t=0}^{N} v^T(t) Q_v^{-1} v(t) + \theta \sum_{t=0}^{N} v_l^T(t+l) v_l(t+l) \end{aligned} \tag{22}$$

where

$$\begin{aligned} v_l(t+l) &= \check{z}(t+l \mid t) - L x(t+l) \\ &= \check{z}(t+l \mid t) - \bar{L}\bar{x}(t+l) \end{aligned} \tag{23}$$

and $\bar{L}$ is as in (21).

Before proceeding further, we introduce the following short-hand notation

$$\bar{v}(t) = \begin{bmatrix} v(t) \\ v_l(t+l) \end{bmatrix}, \quad \bar{y}(t) = \begin{bmatrix} y(t) \\ \check{z}(t+l \mid t) \end{bmatrix} \quad (24)$$

$$\bar{v}_N = col\{\bar{v}(0), \cdots, \bar{v}(N)\} \quad (25)$$

$$\bar{y}_N = col\{\bar{y}(0), \quad \cdots, \quad \bar{y}(N)\} \quad (26)$$

$$R_{\bar{v}} = \overbrace{\begin{bmatrix} Q_v & 0 \\ 0 & \theta^{-1}I_s \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} Q_v & 0 \\ 0 & \theta^{-1}I_s \end{bmatrix}}^{(N+1) \text{ blocks}} \quad (27)$$

$$R_e = \overbrace{Q_e \oplus \cdots \oplus Q_e}^{(N+l_0) \text{ blocks}} \quad (28)$$

The use of (13)-(14), (17) and (23)- (26) gives

$$\begin{bmatrix} \bar{x}_1(0) \\ e_{N+l_0} \\ \bar{y}_N \end{bmatrix} = \Psi \begin{bmatrix} \bar{x}_1(0) \\ e_{N+l_0} \\ \bar{v}_N \end{bmatrix} \quad (29)$$

where Ψ is of the form:

$$\Psi = \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_{r(N+1)} & 0 \\ \Psi_2 & \Psi_1 & \Psi_0 \end{bmatrix} \quad (30)$$

with $\Psi_0 \in \mathbf{R}^{((m+s)(N+1))\times((m+s)(N+1))}$ invertable

From (22)-(29), it follows that

$$\begin{aligned} &J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N) \\ &= \begin{bmatrix} \zeta \\ \bar{y}_N \end{bmatrix}^T \begin{bmatrix} R_\zeta & R_{\zeta\bar{y}} \\ R_{\zeta\bar{y}}^T & R_{\bar{y}} \end{bmatrix}^{-1} \begin{bmatrix} \zeta \\ \bar{y}_N \end{bmatrix} \end{aligned} \quad (31)$$

where $\zeta = [\bar{x}_1^T(0) \;\; e_{N+l_0}^T]^T$ and

$$\begin{bmatrix} R_\zeta & R_{\zeta\bar{y}} \\ R_{\zeta\bar{y}}^T & R_{\bar{y}} \end{bmatrix} = \Psi \begin{bmatrix} \Pi_1 & 0 & 0 \\ 0 & R_e & 0 \\ 0 & 0 & R_{\bar{v}} \end{bmatrix} \Psi^T \quad (32)$$

Then we have the following results.

Lemma 4.1 Consider the system (2) and (3) and the associated cost (5). Then, $J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$ has a minimum with respect to $\{\bar{x}_1(0), e_{N+l_0}\}$ iff

$$R_\zeta - R_{\zeta\bar{y}} R_{\bar{y}}^{-1} R_{\zeta\bar{y}}^T > 0 \quad (33)$$

Furthermore, the minimum is calculated as

$$J_{l,N}^0(\bar{y}_N) = \bar{y}_N^T R_{\bar{y}}^{-1} \bar{y}_N \quad (34)$$

Proof: Omitted.

Lemma 3.1 indicates that the risk-sensitive estimation problem can be reduced to one of finding the estimator $\{\check{z}(t+l \mid t)\}$ such that $J_{l,N}^0(\bar{y}_N)$ is minimum for $\theta > 0$ (risk-seeking) or maximum for $\theta < 0$ (risk-averse). To compute the estimator, it is necessary for us to simplify (34). To this end, we shall construct a stochastic system with appropriate Gramians and introduce an associated innovation sequence.

4.1 Stochastic system and innovation sequence in Krein space

4.1.1. Stochastic system model in Krein space First, associated with the stochastic system (13) and (14) and in view of (23), we introduce

$$\check{\mathbf{z}}(t+l \mid t) = L\mathbf{x}(t+l) + \mathbf{v}_l(t+l) = \bar{L}\bar{\mathbf{x}}(t+l) + \mathbf{v}_l(t+l) \quad (35)$$

where $\mathbf{v}_l(t)$ is a white noise with zero mean and covariance $Q_{v_l} = \langle \mathbf{v}_l(t), \mathbf{v}_l(t)\rangle = \theta^{-1}I_s$, and is independent of $\mathbf{e}(t)$ and $\mathbf{v}(t)$. Putting together (3) with (35) and noting (24) yield

$$\begin{bmatrix} \mathbf{y}(t) \\ \check{\mathbf{z}}(t+l \mid t) \end{bmatrix} = \begin{bmatrix} q^{-l}\bar{H} \\ \bar{L} \end{bmatrix} \bar{\mathbf{x}}(t+l) + \bar{\mathbf{v}}(t) \quad (36)$$

Observe that

$$\langle \bar{\mathbf{v}}(i), \bar{\mathbf{v}}(j)\rangle = \begin{bmatrix} Q_v & 0 \\ 0 & \theta^{-1}I_s \end{bmatrix} \delta_{ij} = Q_{\bar{v}}\delta_{ij} \quad (37)$$

It is obvious that $Q_{\bar{v}}$ is indefinite when $\theta < 0$. In this case, (36) is no longer a stochastic system in Hilbert space but a stochastic system in Krein space. Using the Krein space model (2) and (36), noting (17) and (25), it is easy to verify that

$$\langle \mathbf{e}_N, \; \mathbf{e}_N\rangle = R_e, \quad \langle \bar{\mathbf{v}}_N, \; \bar{\mathbf{v}}_N\rangle = R_{\bar{v}} \quad (38)$$

Then the use of (29) gives

$$\left\langle \begin{bmatrix} \bar{\mathbf{x}}_1(0) \\ \mathbf{e}_N \\ \bar{\mathbf{y}}_N \end{bmatrix}, \begin{bmatrix} \bar{\mathbf{x}}_1(0) \\ \mathbf{e}_N \\ \bar{\mathbf{y}}_N \end{bmatrix} \right\rangle = \Psi \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & R_e & 0 \\ 0 & 0 & R_{\bar{v}} \end{bmatrix} \Psi^T \quad (39)$$

which together with (32) implies that

$$R_{\bar{y}} = \langle \bar{\mathbf{y}}_N, \; \bar{\mathbf{y}}_N\rangle \quad (40)$$

Then, by considering (34), the minimum of $J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$ is given by

$$J_{l,N}^0(\bar{y}_N) = \bar{y}_N^T \langle \bar{\mathbf{y}}_N, \bar{\mathbf{y}}_N\rangle^{-1} \bar{y}_N \quad (41)$$

4.1.2. Innovation sequence and an equivalent condition for minimality To simplify (41), we introduce the innovation sequence of $\bar{\mathbf{y}}_t(t) = \begin{bmatrix} \mathbf{y}(t) \\ \check{\mathbf{z}}(t+l \mid t) \end{bmatrix}$, i.e. the one-step ahead prediction error(Deng *et al.*, 1996)

$$\mathbf{w}(t) = \begin{bmatrix} \mathbf{w}_y(t) \\ \mathbf{w}_z(t) \end{bmatrix} = \bar{\mathbf{y}}(t) - \hat{\bar{\mathbf{y}}}_t(t \mid t-1)$$
$$= \begin{bmatrix} \mathbf{y}(t) \\ \check{\mathbf{z}}(t+l \mid t) \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{y}}(t \mid t-1) \\ \hat{\mathbf{z}}(t \mid t-1) \end{bmatrix} \tag{42}$$

where $\hat{\mathbf{y}}(t \mid t-1)$ and $\hat{\mathbf{z}}(t \mid t-1)$ are respectively obtained from the Krein space projections of $\mathbf{y}(t)$ and $\check{\mathbf{z}}(t+l \mid t)$ onto the linear space $L\{\{\bar{\mathbf{y}}(i)\}_{i=0}^{t-1}\}$. From (42), we obtain

$$\mathbf{w}(t) = L_{t0}\bar{\mathbf{y}}(0) + \cdots + L_{t,t-1}\bar{\mathbf{y}}(t-1) + \bar{\mathbf{y}}(t) \tag{43}$$

Then we have the following relation

$$\mathbf{w}_N = L_w \bar{\mathbf{y}}_N \tag{44}$$

where $\mathbf{w}_N = [\, \mathbf{w}^T(0) \quad \mathbf{w}^T(1) \quad \cdots \quad \mathbf{w}^T(N) \,]$ and

$$L_w = \begin{bmatrix} I_{m+s} & 0 & \cdots & 0 \\ L_{10} & I_{m+s} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ L_{N0} & L_{N1} & \cdots & I_{m+s} \end{bmatrix}$$

Note that L_w is square and nonsingular and that $R_{\bar{y}}$ can be calculated as

$$R_{\bar{y}} = \langle \bar{\mathbf{y}}_N, \bar{\mathbf{y}}_N \rangle = L_w^{-1} R_w L_w^{-T} \tag{45}$$

where

$$R_w = \overbrace{Q_w(0) \oplus \cdots \oplus Q_w(N)}^{(N+1) \text{ blocks}} \tag{46}$$

and $Q_w(t)$ is the covariance matrix of the innovation $\mathbf{w}(t)$.

We now present an equivalent result of Lemma 4.1 using the innovation sequence.

Lemma 4.2 Consider the system (2) and (3) and the associated cost (20). Then, $J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$ has a minimum with respect to $\{\bar{x}_1(0), e_{N+l_0}\}$ iff $Q_w(t)$ $(t = 0, 1, \cdots, N)$ and $Q_{\bar{v}}$ have the same inertia, where $Q_{\bar{v}}$ is as in (37).
Furthermore, the minimum of $J_{l,N}(\bar{x}_1(0), e_{N+l_0}; y_N)$ can be given in terms of the innovation $w(t)$ as

$$J_{l,N}^0(\bar{y}_N) = \sum_{t=0}^{N} [\, w_y^T(t) \quad w_z^T(t) \,] \, Q_w^{-1}(t) \begin{bmatrix} w_y(t) \\ w_z(t) \end{bmatrix} \tag{47}$$

In this case, the risk-sensitive estimation problem is to find the estimator $\check{z}(t+l \mid t)$ $(t = 0, 1, \cdots, N)$ such $J_{l,N}^0(\bar{y}_N)$ is minimum for $\theta > 0$ (risk-seeking) or maximum for $\theta < 0$ (risk-averse).

Proof: Omitted.

4.2 Finite horizon risk-sensitive estimation

In this section the finite horizon estimators are given based on the Lemma 4.2.
First, note that (13) can be rewritten as

$$\bar{\mathbf{x}}(t+l) = \bar{\Phi}\bar{\mathbf{x}}(t+l-1) + \bar{\Gamma}\bar{\mathbf{e}}_{t+l-1} \tag{48}$$

where $\bar{\Gamma} = [\, \Gamma^{(0)} \quad \cdots \quad \Gamma^{(\lambda_0)} \,]$ and

$$\bar{\mathbf{e}}_{t+l-1} = [\, \mathbf{e}^T(t+l-1) \quad \cdots \quad \mathbf{e}^T(t+l+\lambda_0-1) \,]^T$$

satisfy

$$\bar{\mathbf{e}}_{t+l} = \bar{E}_1 \bar{\mathbf{e}}_{t+l-1} + \bar{E}_2 \mathbf{e}(t+l+\lambda_0) \tag{49}$$

with $\bar{E}_2 = [\, 0 \quad 0 \quad \cdots \quad 0 \quad I_r \,]^T$

$$\bar{E}_1 = \begin{bmatrix} 0 & I_r & 0 & \cdots & 0 \\ 0 & 0 & I_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & I_r \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in R^{r(\lambda_0+1) \times r(\lambda_0+1)}$$

Then, the state space model of filtering $(l = 0)$ follows from (48)- (49) and (36)

$$\tilde{\mathbf{x}}(t+1) = \tilde{\Phi}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{e}}(t) \tag{50}$$
$$\begin{bmatrix} \mathbf{y}(t) \\ \check{\mathbf{z}}(t \mid t) \end{bmatrix} = \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} \tilde{\mathbf{x}}(t) + \tilde{\mathbf{v}}(t) \tag{51}$$

where $\tilde{\mathbf{v}}(t) = \bar{\mathbf{v}}(t)$,

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} \bar{\mathbf{x}}(t) \\ \bar{\mathbf{e}}_t \end{bmatrix}, \quad \tilde{\mathbf{e}}(t) = \begin{bmatrix} 0 \\ \bar{E}_2 \mathbf{e}(t+\lambda_0+1) \end{bmatrix},$$
$$\tilde{\Phi} = \begin{bmatrix} \bar{\Phi} & \bar{\Gamma} \\ 0 & \bar{E}_1 \end{bmatrix}, \quad \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} = \begin{bmatrix} \bar{H} & 0 \\ \bar{L} & 0 \end{bmatrix}$$

Similar, we obtain the following state space form for the case of $l > 0$

$$\tilde{\mathbf{x}}(t+1) = \tilde{\Phi}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{e}}(t) \tag{52}$$

$$\begin{bmatrix} \mathbf{y}(t) \\ \check{\mathbf{z}}(t+l \mid t) \end{bmatrix} = \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} \tilde{\mathbf{x}}(t) + \tilde{\mathbf{v}}(t) \tag{53}$$

where $\tilde{\mathbf{v}}(t) = \bar{\mathbf{v}}(t)$ and

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} \bar{\mathbf{x}}(t+l) \\ \bar{\mathbf{e}}_{t+l} \\ \mathbf{x}_m(t) \end{bmatrix}, \quad \tilde{\mathbf{e}}(t) = \begin{bmatrix} 0 \\ \bar{E}_2 \mathbf{e}(t+\lambda_0+l+1) \\ 0 \end{bmatrix}$$

$$\tilde{\Phi} = \begin{bmatrix} \bar{\Phi} & \bar{\Gamma} & 0 \\ 0 & \bar{E}_1 & 0 \\ H_m & 0 & M_m \end{bmatrix}, \quad \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} = \begin{bmatrix} 0 & 0 & T_m \\ \bar{L} & 0 & 0 \end{bmatrix}$$

with $T_m = [\, I_m \quad 0 \quad \cdots \quad 0 \,]$, $H_m = [\, 0 \quad \cdots \quad \bar{H} \,]^T$ and

$$M_m = \begin{bmatrix} 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in R^{ml \times ml}$$

For the case of smoothing $l < 0$, we obtain the state space model as

$$\tilde{\mathbf{x}}(t+1) = \tilde{\Phi}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{e}}(t) \qquad (54)$$

$$\begin{bmatrix} \mathbf{y}(t-l) \\ \check{\mathbf{z}}(t \mid t-l) \end{bmatrix} = \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} \tilde{\mathbf{x}}(t) + \tilde{\mathbf{v}}(t) \qquad (55)$$

where $\tilde{\mathbf{v}}(t) = \bar{\mathbf{v}}(t-l), \tilde{\mathbf{x}}(t) = [\, \bar{\mathbf{x}}^T(t-l) \quad \bar{\mathbf{e}}^T_{t-l} \quad \mathbf{x}^T_s(t) \,]^T$ and

$$\tilde{\mathbf{e}}(t) = \begin{bmatrix} 0 \\ \bar{E}_2\mathbf{e}(t+\lambda_0 - l + 1) \\ 0 \end{bmatrix}$$

$$\tilde{\Phi} = \begin{bmatrix} \bar{\Phi} & \bar{\Gamma} & 0 \\ 0 & \bar{E}_1 & 0 \\ L_s & 0 & M_s \end{bmatrix}, \quad \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} = \begin{bmatrix} \bar{H} & 0 & 0 \\ 0 & 0 & T_s \end{bmatrix}$$

with $T_s = [\, I_s \quad 0 \quad \cdots \quad 0 \,]$ and

$$M_s = \begin{bmatrix} 0 & I_s & 0 & \cdots & 0 \\ 0 & 0 & I_s & \cdots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \cdots & I_s \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad L_s = \begin{bmatrix} 0 \\ \vdots \\ \bar{L} \end{bmatrix}$$

Theorem 4.1 Consider the system (2)-(4) satisfying Assumptions 2.1 and 2.2. For any given $\theta > 0$, the risk-seeking estimation problem always has a solution. For any given $\theta < 0$, a solution for risk-averse estimation exists iff $Q_{\bar{v}} = \begin{bmatrix} Q_v & 0 \\ 0 & \theta^{-1}I \end{bmatrix}$ and

$$\begin{aligned} Q_w(t) &= \begin{bmatrix} Q_v & 0 \\ 0 & \theta^{-1}I \end{bmatrix} + \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} P_t \,[\, \tilde{H}^T \quad \tilde{L}^T \,] \\ &= \begin{bmatrix} Q_v + \tilde{H}P_t\tilde{H}^T & \tilde{H}P_t\tilde{L}^T \\ \tilde{L}P_t\tilde{H}^T & \theta^{-1}I + \tilde{L}P_t\tilde{L}^T \end{bmatrix} \end{aligned} \qquad (56)$$

have the same inertia for $t = 0, 1, \cdots N$, where

$$\begin{aligned} P_{t+1} = {} & \tilde{\Phi}P_t\tilde{\Phi}^T + Q_{\tilde{e}} - \\ & \tilde{\Phi}P_t \,[\, \tilde{H}^T \quad \tilde{L}^T \,]\, Q_w^{-1}(t) \begin{bmatrix} \tilde{H} \\ \tilde{L} \end{bmatrix} P_t\tilde{\Phi}^T \end{aligned} \qquad (57)$$

with $P_0 = \langle \tilde{\mathbf{x}}(0), \tilde{\mathbf{x}}(0) \rangle$ In both cases the optimal risk-sensitive filter, smoother or predictor is given by

$$\begin{aligned} \check{z}(t+l \mid t) = {} & \tilde{L}\hat{\tilde{x}}(t \mid t-1) + \tilde{L}P_t\tilde{H}^T(I + \tilde{H}P_t\tilde{H}^T)^{-1} \\ & \times[y(t) - \tilde{H}\hat{\tilde{x}}(t \mid t-1)] \end{aligned} \qquad (58)$$

$$\begin{aligned} \hat{\tilde{x}}(t+1 \mid t) = {} & \tilde{\Phi}\hat{\tilde{x}}(t \mid t-1) + \tilde{\Phi}P_t\tilde{H}^T](I + \tilde{H}P_t\tilde{H}^T)^{-1} \\ & \times[y(t) - \tilde{H}\hat{\tilde{x}}(t \mid t-1)] \end{aligned} \qquad (59)$$

5. CONCLUSION

In this paper, the finite horizon risk-sensitive estimation problem for linear discrete time-invariant singular systems has been solved for the first time. A simple derivation technique based on a Krein space signal model. The key behind the method is to convert the singular model into a nonsingular system through state augmentation. The result involves solving a Riccati difference equation. It should be noted that the calculation for smoothing and prediction estimation is complicated due to the state augmentation.

References

Lewis, F.L. and V.G. Mertzios, Eds.(1989). Recent advances in singular systems. *Special Issue, Circuits, Syst. and Signal Processing*, **8**.

Cobb (1984). Controllability, observability and duality in singular systems. *IEEE Trans. Automat. Control*, **29**, 1076-1082.

Dai, L. (1989) . Filtering and LQG problem for discrete-time stochastic singular systems. *IEEE.Trans Automat. Control*, **34**, 1105-1108.

Nikoukhah, R., S. Willsky and B. C. Levy (1992). Kalman filtering and Riccati equations for singular systems. *IEEE Trans. Automat. control*, **37**, 1325-1342.

Zhang, H., T. Y. Chai and X. J. Liu (1998). A unified approach to optimal estimation for discrete-time stochastic singular linear systems. *Automatica*, **34**, 777-781.

Deng, Z. L., H. S. Zhang, S. J. Liu and L. Zhou (1996). Optimal and self-tuning white noise estimators with applications to deconvolution and filtering problems. *Automatica, 32*, 199-216.

Hassibi,B., A.H. Sayed and T. Kailath (1998). *Indefinite Quadratic Estimation and Control: A Unified Approach to H_2 and H_∞ Theories,* SIAM Studies in Applied Mathematics series.

Jacobson, D. H..(1973) Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Trans. Automat. Contr.,* **18**, 24-131.

Whittle,P.(1990). *Risk-sensitive optial control,* John Wiley and Sons, Chichester and New York.

Hernandez-Hernandez, D., S. I. Marcus and P. J. Fard (1999). Analysis of a risk-sensitive control problem for hidden Markov chains. *IEE Transaction on Automatic Control*, **44**, 1093-1100.