Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

The behavioral approach to discrete-time deterministic Kalman filtering on the di-polynomial ring

Osamu Kaneko and Takao Fujii

Abstract— **In this paper, we address the discrete time deterministic Kalman filtering within a behavioral setting. In the continuous time case, Fagnani and Willems developed this issue by using the novel idea based on two-variable polynomial matrices and quadratic differential forms in [1] and [12]. We expand the Kalman filtering problem into the discrete time case by using two-variable di-polynomial matrices and quadratic difference forms studied in [3], [4] and [5]. Here we derive a sufficient condition for the latent variables of the filter to estimate the observed data and the state variable of the system in the sense that the sum of squared error signals between observed variables and estimated ones is minimized deterministically. By using this condition, we then clarify the structure of the optimal filter with respect to the notion of the state. And then we also provide the procedure for implementing the optimal filter as a real-time algorithm. Finally, we give an illustrative example in order to show the validity of our results.**

I. INTRODUCTION

In this paper, we address the discrete time deterministic Kalman filtering within a behavioral setting. In [1] and [12], Fagnani and Willems developed this issue by using the novel idea based on two-variable polynomial matrices and quadratic differential forms (cf. [13]) in continuous time. From the practical points of view, the filtering problem is deeply related to how to treat the sampled data, which implies that it is meaningful to consider this issue in the discrete time case. From the theoretical points of view, there are some critical differences between the continuous time and the discrete time case. One of such issues is that polynomial matrices used in discrete time case consist of terms of not only non-negative but also positive powers with respect to the indeterminate (we refer to such polynomials as "di-polynomials" or "Laurent polynomials" in this paper). Since the dipolynomial matrix ring is peculiar to the discrete time case, it is to be expected that the di-polynomials may be useful tools to study essential features of discrete time dynamical systems and to derive important theoretical results. Thus, it is significant to study discrete time systems on the dipolynomial ring. Another different point is on the notion of the relationship between storage functions and state variables. In [1] and [12], the fact that every storage function is a quadratic function of the state ([10]) is used for the derivation of the main theorem and the implementation issue. In discrete time, as shown in [5], not every storage function is a quadratic function of the state. This fact means that another strategies should be used to derive the important theoretical results in discrete time Kalman filtering.

From these reasons, this paper addresses the discrete time deterministic Kalman filtering within a behavioral setting. Here, we approach to this problem on two-variable dipolynomial matrices and quadratic difference forms. We also derive a sufficient condition for the filter to estimate the observed data and the state variable of the system in the sense that the sum of squared error signals between the observed data and estimated ones is minimized. By using this condition, we then clarify the structure of the optimal filter. We also provide the implementation procedure for deterministic Kalman filtering.

This work was not supported by any organization

Graduate School of Engineering Science, Osaka University , Machikaneyama 1-3, Toyonaka, Osaka, Japan kaneko@ft-lab.sys.es.osaka-u.ac.jp

II. PRELIMINARIES

A. Notation

In this paper we denote the set of real numbers with $\mathbb{R}$ and the set of integers with $\mathbb{Z}$. The space of $\mathtt{n}$ dimensional real vectors is denoted by $\mathbb{R}^{\mathtt{n}}$, and the space of $\mathtt{w} \times \mathtt{m}$ real matrices, by $\mathbb{R}^{\mathtt{w}\times\mathtt{m}}$. If $A \in \mathbb{R}^{\mathtt{w}\times\mathtt{m}}$, then $A^{\mathrm{T}} \in \mathbb{R}^{\mathtt{m}\times\mathtt{w}}$ denotes its transpose. Whenever the size of a matrix or a vector is not specified, a bullet $\bullet$ is used. In order to enhance readability, when dealing with a vector space $\mathbb{R}^{\bullet}$ whose elements are commonly denoted with w, we use the notation $\mathbb{R}^{\mathtt{w}}$; similar considerations hold for matrices representing linear operators on such spaces.

The set consisting of all sequences from $\mathbb{Z}$ to $\mathbb{R}^{\mathtt{w}}$ is denoted with $(\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$. For $w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ and $T \in \mathbb{Z}$, $w_{(T)}$ denotes the value of w at time T. On $(\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$, we define the backwards shift operator $\sigma : (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}} \to (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ as $(\sigma w)_{(t)} := w_{(t+1)}$ for all $t \in \mathbb{Z}$. The set of square-summable $\mathtt{w}$-dimensional vector time series is denoted with $l_2(\mathbb{Z}, \mathbb{R}^{\mathtt{w}})$, i.e. $w \in l_2(\mathbb{Z}, \mathbb{R}^{\mathtt{w}})$ if $\sum_{t=-\infty}^{\infty} w_{(t)}^{\mathrm{T}} w_{(t)} < \infty$. For $T \in \mathbb{Z}$, let $l_2^T(\mathbb{Z}, \mathbb{R}^{\mathtt{w}})$ denote the set of $\mathtt{w}$-dimensional vector time series which is square summable from $-\infty$ to T, i.e., $w \in l_2^T(\mathbb{Z}, \mathbb{R}^{\mathtt{w}})$ if $\sum_{t=-\infty}^{T} w_{(t)}^{\mathrm{T}} w_{(t)} < \infty$.

The ring of polynomials with real coefficients in the indeterminate ξ is denoted by $\mathbb{R}[\xi]$; the ring of two-variable polynomials with real coefficients in the indeterminates ζ and η is denoted by $\mathbb{R}[\zeta, \eta]$. Similarly, the ring of "di-polynomials" with both positive and negative powers of ξ is denoted with $\mathbb{R}[\xi^{-1}, \xi]$. In addition, the ring of dipolynomials with only nonpositive powers of ξ is denoted by $\mathbb{R}[\xi^{-1}]$. Let $\mathbb{R}[\zeta, \zeta^{-1}, \eta, \eta^{-1}]$ denote the set of two-variable di-polynomials with both positive and negative powers of the indeterminates ζ and η. Similarly, we use the notations, $\mathbb{R}[\zeta^{-1}, \eta]$, $\mathbb{R}[\zeta, \eta^{-1}]$, and so on. The matrix versions of these rings of size $\mathtt{w} \times \mathtt{m}$ are denoted by $\mathbb{R}^{\mathtt{w}\times\mathtt{m}}[\xi]$, $\mathbb{R}^{\mathtt{w}\times\mathtt{m}}[\zeta, \eta]$,

$\mathbb{R}^{\mathtt{w}\times\mathtt{m}}[\xi^{-1},\xi]$, $\mathbb{R}^{\mathtt{w}\times\mathtt{m}}[\xi^{-1}]$, and $\mathbb{R}^{\mathtt{w}\times\mathtt{m}}[\zeta,\zeta^{-1},\eta,\eta^{-1}]$ respectively. Let $\mathbb{R}_s^{\mathtt{w}\times\mathtt{w}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$ denote the set of two-variable di-polynomial matrices satisfying $\Phi(\zeta^{-1},\zeta,\eta^{-1},\eta) = \Phi(\eta^{-1},\eta,\zeta^{-1},\zeta)^{\mathrm{T}}$.

B. Behavioral system theory ([2], [9], [11])

A discrete time dynamical system is defined as a triple $\Sigma = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathfrak{B})$, where $\mathbb{Z}$ is the discrete time axis, $\mathbb{R}^{\mathtt{q}}$ is the signal space, and $\mathfrak{B}$ is the (manifest) behavior. A dynamical system $\Sigma = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathfrak{B})$ is linear, time-invariant and complete if and only if Σ is representable by a kernel representation $R(\sigma)w = 0$ with $R(\xi) \in \mathbb{R}^{\bullet\times\mathtt{q}}[\xi]$ for all $w \in \mathfrak{B}$. In addition to manifest variables w, there are many cases in which some auxiliary variables, say ℓ, are required to describe a dynamics. It is called a latent variable and a dynamical system with latent variables is defined as a quadruple $\Sigma_a = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathbb{R}^{\bullet},\mathfrak{B}_a)$, where $\mathbb{R}^{\bullet}$ is the signal space of ℓ and $\mathfrak{B}_a \subseteq (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}} \times (\mathbb{R}^{\bullet})^{\mathbb{Z}}$ is the full behavior.

A linear time-invariant complete system Σ is controllable (in a behavioral sense) if and only if it is representable by an image representation $w = M(\sigma)\ell$ with $M(\xi) \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}[\xi]$ for all $(w,\ell) \in \mathfrak{B}_a$, where ℓ is the latent variable. In $w = M(\sigma)\ell$, if $w = 0$ implies $\ell = 0$, then ℓ is said to be observable from w, which is equivalent to that $M(\lambda)$ is full column rank for all nonzero $\lambda \in \mathbb{C}$. In such a case, $w = M(\sigma)\ell$ is called an observable image representation. A controllable system has many observable image representations. If $M(\xi)$ is full column rank, then there exists an nonsingular matrix $P \in \mathbb{R}^{\mathtt{q}\times\mathtt{q}}$ such that $PM(\xi) = \mathrm{col}\left[\begin{array}{c} U(\xi)^{\mathrm{T}}, Y(\xi)^{\mathrm{T}} \end{array}\right]^{\mathrm{T}}$, where $\det(U(\xi)) \neq 0$ and $Y(\xi)U(\xi)^{-1}$ is proper. We can regard $u := U(\sigma)\ell$ ($y := Y(\sigma)\ell$) as inputs (outputs, respectively). In this paper, we assume that a system is discrete-time, linear, time-invariant, complete and controllable.

Although it is enough to use the polynomial matrix ring for describing difference equations in the discrete time systems, there are some cases in which using the di-polynomial ring enables us to take a comprehensive and panoramic view of theoretical feature which is peculiar to discrete time systems as stated in the introduction. In such cases, we often use mathematical representations induced by di-polynomial matrices, e.g., $w = M(\sigma^{-1})\ell$ and so on.

Finally, we prepare the notions of state space systems. Let $\Sigma_s = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathbb{R}^{\mathtt{n}},\mathfrak{B}_s)$ denote a system with latent variables. Then Σ_s is said to be a state space system if $\{(w_1,x_1),(w_2,x_2) \in \mathfrak{B}_s$ and $x_1(0) = x_2(0)\} \Longrightarrow \{(w_1,x_1) \wedge_0 (w_2,x_2) \in \mathfrak{B}_s\}$. Here, a latent variable x is said to be a state variable of $\mathfrak{B}$. Let $\Sigma_s = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathbb{R}^{\mathtt{n}},\mathfrak{B}_s)$ denote a state space system whose manifest behavior is $\mathfrak{B}$. For all state space systems $\Sigma_s' = (\mathbb{Z},\mathbb{R}^{\mathtt{q}},\mathbb{R}^{\mathtt{n}'},\mathfrak{B}_s')$ inducing the same manifest behavior $\mathfrak{B}$, Σ_s is said to be minimal if $n \leq n'$ holds.

Consider the behavior $\mathfrak{B}$ and the following map described by $x := X(\sigma)w$ for $w \in \mathfrak{B}$, induced by $X(\xi) \in \mathbb{R}^{\bullet\times\mathtt{d}}[\xi]$. If x is a (minimal) state variable, then $X(\sigma)w$ is called a (minimal) state map for $\mathfrak{B}$. We can also consider a (minimal) state map acting on ℓ.

Let $F(\xi) \in \mathbb{R}^{\bullet\times\mathtt{d}}[\xi]$. Then, $F(\xi)U(\xi)^{-1}$ is strictly proper if and only if there exist $H \in \mathbb{R}^{\bullet\times\mathtt{d}}$ such that $F(\sigma)\ell = Hx$ with a minimal state variable x of $\mathfrak{B}$.

C. Two-variable di-polynomial matrices and quadratic difference forms ([3], [4], [13])

Here, we focus on two-variable di-polynomial matrices. As for the set of two-variable polynomial matrices in discrete time, we omit it here. See the details in [3].

Let $\Phi(\zeta,\eta) = \sum_{k=-M_k,l=-M_l}^{L_k,L_l}\Phi_{kl}\zeta^k\eta^l \in \mathbb{R}^{\mathtt{q}\times\mathtt{p}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$. For $w \in (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}}$ and $v \in (\mathbb{R}^{\mathtt{p}})^{\mathbb{Z}}$, $\Phi(\zeta,\eta)$ induces a bi-linear difference forms $L_\Phi : (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}} \times (\mathbb{R}^{\mathtt{p}})^{\mathbb{Z}} \mapsto (\mathbb{R})^{\mathbb{Z}}$ as

$$L_\Phi(w,v)(t) := \sum_{k=-M_k,l=-M_l}^{L_k,L_l} (\sigma^k w)^{\mathrm{T}}\Phi_{kl}(\sigma^l v)(t). \qquad (1)$$

In the case of $\Phi = I_{\mathtt{q}}$, we use the standard notation $< v,w >$. Next, consider $\sum_{k,l=-M}^{L}\Phi_{kl}\zeta^k\eta^l \in \mathbb{R}_s^{\mathtt{q}\times\mathtt{q}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$ For $w \in (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}}$, $\Phi(\zeta,\eta)$ induces a quadratic difference forms $Q_\Phi : (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}} \times (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}} \mapsto (\mathbb{R})^{\mathbb{Z}}$ as

$$Q_\Phi(w,v)(t) := \sum_{k,l=-M}^{L} (\sigma^k w)^{\mathrm{T}}\Phi_{kl}(\sigma^l v)(t) \qquad (2)$$

If $\Phi = I_{\mathtt{q}}$, we use the standard notation $\|w\|^2$. The nonnegativity of quadratic difference forms induced by $\Phi(\zeta,\eta) \in \mathbb{R}_s^{\mathtt{q}\times\mathtt{q}}[\zeta,\eta]$ are defined by

$$Q_\Phi \geq 0 :\Leftrightarrow Q_\Phi(w)(t) \geq 0, {}^\forall w \in (\mathbb{R}^{\mathtt{q}})^{\mathbb{Z}}, {}^\forall t \in \mathbb{Z}. \qquad (3)$$

As for the ring $\mathbb{R}^{\mathtt{q}\times\mathtt{p}}[\zeta^{-1},\eta]$, $\mathbb{R}^{\mathtt{q}\times\mathtt{p}}[\zeta,\eta^{-1}]$, the bi-linear difference forms are introduced similarly.

Next lemma is a generalization of Lemma 3.1 in [3] for the case of two-variable di-polynomial matrices.

Lemma 1: Let $\Phi(\zeta^{-1},\eta) \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}[\zeta^{-1},\eta]$. The following two statements are equivalent.

1. $\Phi((\xi^{-1})^{-1},\xi) = 0$.
2. There exists $\Psi(\zeta^{-1},\eta) \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}[\zeta^{-1},\eta]$ such that $(\zeta\eta - 1)\Psi(\zeta^{-1},\eta) = \Phi(\zeta^{-1},\eta)$.

D. Discrete time dissipativeness ([3], [4])

Let $\Phi(\zeta^{-1},\zeta,\eta^{-1},\eta) \in \mathbb{R}_s^{\mathtt{d}\times\mathtt{d}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$ induce a Q_Φ, which can be regarded as the quadratic supply rate in the context of the dissipation theory. Q_Φ is said to be dissipative if $\sum_{t=-\infty}^{\infty} Q_\Phi(\ell) \geq 0$ for all $\ell \in l_2(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$. Next, we introduce the notion of storage- and of dissipation function.

1. Q_Ψ induced by $\Psi(\zeta^{-1},\zeta,\eta^{-1},\eta) \in \mathbb{R}_s^{\mathtt{d}\times\mathtt{d}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$ is said to be a storage function with respect to Q_Φ if $Q_\Psi(\ell)(t+1) - Q_\Psi(\ell)(t) \leq Q_\Phi(\ell)(t), {}^\forall t \in \mathbb{Z}$ and ${}^\forall \ell \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$.
2. Q_Δ induced by $\Delta(\zeta^{-1},\zeta,\eta^{-1},\eta) \in \mathbb{R}_s^{\mathtt{d}\times\mathtt{d}}[\zeta^{-1},\zeta,\eta^{-1},\eta]$ is said to be a dissipation function with respect to Q_Φ if $Q_\Delta \geq 0$ and $\sum_{t=-\infty}^{t=\infty} Q_\Phi(\ell)(t) = \sum_{t=-\infty}^{t=\infty} Q_\Delta(\ell)(t)$, ${}^\forall \ell \in l_2(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$.

Dissipativity is characterized in terms of storage- and of dissipation functions as follows (See [4], [13]).

Proposition 2: The following conditions are equivalent:

1) $\sum_{t=-\infty}^{\infty} Q_\Phi(\ell)(t) \geq 0$ for all $\ell \in l_2(\mathbb{Z}, \mathbb{R}^{\mathtt{d}})$;
2) Q_Φ admits a storage function;
3) Q_Φ admits a dissipation function.

Moreover, there is a one-one correspondence between storage and dissipation functions, Q_Ψ and Q_Δ, respectively, defined by $Q_\Psi(\sigma w) - Q_\Psi(w) = Q_\Phi(w) - Q_\Delta(w)$.

These properties holds also in the case of two-variable polynomial matrices in discrete time. See [3] for details.

III. Problem formulation

The problem we attack here is similar to the continuous time case in the section 4 of [1]. Let $\bar{w} \in l_2^T(\mathbb{Z}, \mathbb{R}^{\mathtt{q}})$ be the observed data of the system in $(-\infty, T]$. Let $\mathfrak{B}$ be the behavior of the "nominal" system. Since there are perturbations, noises, and uncertainty in the real measured data $\bar{w}$, note that generally

$$\bar{w} \notin \mathfrak{B}. \tag{4}$$

Assume that $\mathfrak{B}$ of the nominal system has an image representation

$$w = M(\sigma)\ell \tag{5}$$

where $M(\xi) \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}[\xi]$ with full column rank for all complex numbers. Then, we can assume that $M(\xi)$ is described by $M(\xi) = [U(\xi)^{\mathrm{T}} \quad Y(\xi)^{\mathrm{T}}]^{\mathrm{T}}$ such that $U(\xi) \in \mathbb{R}^{\mathtt{d}\times\mathtt{d}}[\xi]$ is nonsingular and $Y(\xi)U(\xi)^{-1}$ is proper after possibly permuting the components of the variable w. Moreover, we assume that

$$\det |U(0)| \neq 0. \tag{6}$$

Under this setting, our aim is to find ℓ such that

$$\sum_{t=-\infty}^{T} \|\bar{w}_{(t)} - (M(\sigma)\ell)_{(t)}\|^2 \tag{7}$$

is minimized for any time $T \in \mathbb{Z}$ deterministically. In this setting, the optimal estimation, say w^*, depends on the past of the observation $\bar{w}$, so we refer to this problem as the deterministic filtering problem. In the following, we focus on how to compute w^* from the past observation $\bar{w}$ [1].

IV. The optimal filtering

A. The new variable for the filter and some preliminaries of dissipation theory

By using the observation $\bar{w}$, consider the behavior $\mathfrak{B}_{fw}$ described by the following kernel representation

$$\begin{bmatrix} M(\sigma^{-1})^{\mathrm{T}} & A(\sigma^{-1})^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \bar{w} \\ f \end{bmatrix} = 0 \tag{8}$$

where, $A(\xi) \in \mathbb{R}^{\mathtt{d}\times\mathtt{d}}[\xi]$ is an anti-Hurwitz spectral factor of

$$M(\xi^{-1})^{\mathrm{T}} M(\xi) = A(\xi^{-1})^{\mathrm{T}} A(\xi). \tag{9}$$

[1] In [1], the smoothing problem in which the whole of the observed data is used for the minimization of the error was also studied as the preparation of the deterministic filter. In discrete time, we can also obtain the corresponding result, however, we omit the detailed discussions here due to the limitations of the space.

At this point, we prepare the following useful lemma for the filtering equations.

Lemma 3: $A(\xi^{-1})^{-\mathrm{T}} M(\xi^{-1})$ is proper.

Proof: First, we obtain

$$\begin{aligned} &A(\xi^{-1})^{-\mathrm{T}} M(\xi^{-1})^{\mathrm{T}} \\ =&A(\xi^{-1})^{-\mathrm{T}} U(\xi^{-1})^{\mathrm{T}} U(\xi^{-1})^{-\mathrm{T}} M(\xi^{-1})^{\mathrm{T}}. \end{aligned} \tag{10}$$

Let $A(\xi) =: A_0 + A_1\xi + \cdots$. Note that we can obtain an anti-Hurwitz spectral factor such that A_0 is non-singular (cf.[7]). Thus, we see that $A(\xi^{-1})^{\mathrm{T}} = A_0^{\mathrm{T}} + \frac{A_1^{\mathrm{T}}}{\xi} + \cdots$ is bi-proper, which implies $A(\xi^{-1})^{-\mathrm{T}}$ is proper. Since we assume that $U(0)$ is non-singular, it follows the same reason from that $U(\xi^{-1})^{\mathrm{T}}$ is proper. On the other hand, it follows from Theorem 5.2 in [5] that $U(\xi^{-1})^{-\mathrm{T}} M(\xi^{-1})^{\mathrm{T}}$ is proper. Thus, the left-hand side of Eq.(10) is also proper. ■

From the above lemma, we see that f is determined from the observation $\bar{w}$ uniquely as

$$f = A(\sigma^{-1})^{-\mathrm{T}} M(\sigma^{-1})^{\mathrm{T}} \bar{w}. \tag{11}$$

Let $\mathtt{n}$ denote the dimension of the minimal state space of $\mathfrak{B}$. Next, we introduce new polynomial matrices introduced from $M(\xi)$ and a minimal state map $X(\xi) \in \mathbb{R}^{\mathtt{n}\times\mathtt{q}}[\xi]$, which corresponds to Proposition 5 in [1] (However, the proof we provide here is different).

Proposition 4: There exist $F(\xi) \in \mathbb{R}^{\mathtt{n}\times\mathtt{d}}[\xi]$ and $W(\xi) \in \mathbb{R}^{\mathtt{n}\times\mathtt{q}}[\xi]$ such that

$$(\zeta\eta - 1)W(\zeta^{-1})^{\mathrm{T}} X(\eta) = M(\zeta^{-1}) - M(\eta) \tag{12}$$

$$(\zeta\eta - 1)F(\zeta^{-1})^{\mathrm{T}} X(\eta) = -A(\zeta^{-1}) + A(\eta) \tag{13}$$

Proof: Note that $M((\xi^{-1})^{-1}) - M(\xi) = 0$. It follows from Lemma 1 that there exists $\Psi(\zeta^{-1}, \eta) \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}[\zeta^{-1}, \eta]$ such that $(\zeta\eta - 1)\Theta(\zeta^{-1}, \eta) = M(\zeta^{-1}) - M(\eta)$. Moreover, we can obtain a canonical factorization $\Theta(\zeta^{-1}, \eta) = W'(\zeta^{-1})^{\mathrm{T}} X(\eta)$ [2], where $W'(\xi) \in \mathbb{R}^{\bullet\times\mathtt{q}}[\xi]$. Since $M(\eta)U(\eta)^{-1}$ is proper with respect to η, $\Theta(\zeta^{-1}, \eta)U(\eta)^{-1}$ is strictly proper, which enables us to rewrite $\Theta(\zeta^{-1}, \eta) = W'(\zeta^{-1})^{\mathrm{T}} S X(\eta)$ by using an appropriate $S \in \mathbb{R}^{\bullet\times\mathtt{n}}$. Defining $W(\xi) := S^{\mathrm{T}} W'(\xi) \in \mathbb{R}^{\mathtt{n}\times\mathtt{q}}[\xi]$ complete the proof of the existence of $W(\xi)$. Similar argument holds for $F(\xi)$. ■

Eq.(12) and Eq.(13) induce

$$\begin{aligned} &\left\langle W(\sigma^{-1})w, X(\sigma)\ell \right\rangle_{(t+1)} - \left\langle W(\sigma^{-1})w, X(\sigma)\ell \right\rangle_{(t)} \\ &\quad = \left\langle M^{\mathrm{T}}(\sigma^{-1})w, \ell \right\rangle_{(t)} - \left\langle w, M(\sigma)\ell \right\rangle_{(t)} \end{aligned} \tag{14}$$

$$\begin{aligned} &\left\langle F(\sigma^{-1})f, X(\sigma)\ell \right\rangle_{(t+1)} - \left\langle F(\sigma^{-1})f, X(\sigma)\ell \right\rangle_{(t+1)} \\ &\quad = -\left\langle A^{\mathrm{T}}(\sigma^{-1})f, \ell \right\rangle_{(t)} + \left\langle f, A(\sigma)\ell \right\rangle_{(t)} \end{aligned} \tag{15}$$

on the latent variable ℓ of $\mathfrak{B}$ and f, respectively.

In the end of this subsection, we also give a proposition on discrete time dissipative systems.

Proposition 5: Let $\Phi(\zeta, \eta) := M(\zeta)^{\mathrm{T}} M(\eta)$ induce the supply rate Q_Φ. Let x be the minimal state of $\mathfrak{B}$. Then, there exists a storage function for Q_Φ, say Q_Ψ induced by

[2] As for canonical factorizations, see [13] and [3].

$\Psi(\zeta,\eta)\in\mathbb{R}^{\mathtt{q}\times\mathtt{q}}[\zeta,\eta]$, such that there exists a $K=K^{\mathrm{T}}>0$ such that $Q_\Psi(w)=x^{\mathrm{T}}Kx$.

Proof: Since $Q_\Phi(\ell)>0$ holds for arbitrary $\ell\in(\mathbb{R}^{\mathtt{d}})^{\mathbb{Z}}$, $\sum_{t=-\infty}^{T}Q_\Phi(\ell)(t)>0$ for any $T\in\mathbb{Z}$ and $\ell\in l_2^T(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$. It follows from Theorem 6.1 of [5] that there exists a non-negative storage function described by quadratic forms of the state of $\mathfrak{B}$.

Next, by using the observability of $M(\xi)$, it is easy to see that $\Phi(e^{-j\omega},e^{j\omega})=M(e^{-j\omega})^{\mathrm{T}}M(e^{j\omega})>0$ on the unit circle, which implies that the existence of an anti-Hurwitz spectral factor, say $A(\xi)$, of $\Phi(e^{-j\omega},e^{j\omega})$. Here, we can also assume that $A(0)$ is nonsingular. Noting that $A(\zeta)^{\mathrm{T}}A(\eta)$ induces a dissipation rate, it follows from Theorem 4.2 of [5] that the corresponding storage function is the maximum storage function, say Q_{Ψ_+} induced by $\Psi_+(\zeta,\eta)$. Moreover, it follows from Theorem 5.2 of [5] that $Q_{\Psi_+}(\ell)$ can be described by the quadratic non-negative function of the minimal state of $\mathfrak{B}$ as $Q_{\Psi_+}(\ell)=x^{\mathrm{T}}Kx$, where $K=K^{\mathrm{T}}\geq 0$. At this point, the dissipation relation with respect to $Q_{\Psi_+}(\ell)=x^{\mathrm{T}}Kx$ can be described by

$$(x^{\mathrm{T}}Kx)_{(t+1)}-(x^{\mathrm{T}}Kx)_{(t)}=\|M(\sigma)\ell\|_{(t)}^2-\|A(\sigma)\ell\|_{(t)}^2,\quad(16)$$

for all $t\in\mathbb{Z}$. In the following, we show the positivity of K. Suppose that there exists $T\in\mathbb{Z}$ and nonzero $a\in\mathbb{R}^{\mathtt{n}}$ such that $a^{\mathrm{T}}Ka=0$ at time T. Note that a can be described by using the minimal state map $X(\xi)=\sum_{i=0}^{\mu}X_i\xi^i\in\mathbb{R}^{\mathtt{n}\times\mathtt{d}}[\xi]$ as $a=\sum_{i=0}^{\mu}X_i\ell_{(T+1+i)}$ for appropriate finite time series $\ell_{(T+1)},\cdots,\ell_{(T+\mu+1)}$. By the way, due to the nonsingularity of A_0, we can find $\ell_{(0)}$ such that $\ell_{(0)}=-A_0^{-1}(A_1\ell_{(T+1)}+\cdots+A_\nu\ell_{(T+\nu)})$ for arbitrary $\ell_{(T+1)},\cdots,\ell_{(T+\nu)}$. Repeating this procedure until $t=-\infty$ enables us to obtain the solution ℓ in $l_2^{T+\nu}(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$ of $(A(\sigma)\ell)=0$ for arbitrary $\ell_{(T+1)},\cdots,\ell_{(T+\nu)}$. In the case of $\nu\leq\mu+1$, we can also take arbitrary $\ell_{(T+1)},\cdots,\ell_{(T+\mu+1)}$. Define $h:=\mathtt{max}(\nu,\mu+1)$. Thus, for a, there exists $\ell\in l_2^{T+h}(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$ such that $A(\sigma)\ell_{(t)}=0$ in $t\in(-\infty,T+h]$ and $a=\sum_{i=0}^{\mu}X_i\ell_{(T+1+i)}$. By using this $\ell\in l_2^{T+h}(\mathbb{Z},\mathbb{R}^{\mathtt{d}})$, and summing Eq.(16) from T to $-\infty$, we obtain $x^{\mathrm{T}}Kx_{(T+1)}=a^{\mathrm{T}}Ka=\sum_{t=-\infty}^{T}\|M(\sigma)\ell\|_{(t)}^2=0$, which implies

$$w_{(t)}=\|M(\sigma)\ell\|_{(t)}=0,^\forall t\in(-\infty,T].\quad(17)$$

Since $X(\xi)$ induces one of the minimal state maps for $\mathfrak{B}$, $X(\xi)U(\xi)^{-1}$ is strictly proper. This enables us to describe this rational function as the following Laurent expansion $X(\xi)U(\xi)^{-1}=\sum_{i=1}^{\infty}\frac{H_{-i}}{\xi^i}$, $H_{-i}\in\mathbb{R}^{\mathtt{n}\times\mathtt{d}}$. This means that the state $x_{(T+1)}$ is determined uniquely from the input $u_{(t)}:=(U(\sigma)\ell)_{(t)}$ in $t\in(-\infty,T]$, in other words,

$$x_{(T+1)}=\sum_{t=-\infty}^{T}H_{-i}u_{(t)}\quad(18)$$

On the other hand, Eq.(17) implies that the right hand side of Eq.(18) is equal to zero, which yields $x_{(T+1)}=a=0$. Hence, we conclude $K>0$. ∎

From the above theorem, we can assume that the maximum storage function is described by $\|X(\sigma)\ell\|^2$ (i.e., we can assume $K=I$) without loss of generality.

B. A sufficient condition for the optimal filter

By introducing the new variable f as in Eq.(8) or Eq.(11), we give a sufficient condition for the latent variables of the nominal model to minimize the cost function as follows.

Theorem 6: Let f be Eq.(8) or Eq.(11) Then, if ℓ in the nominal model $w=M(\sigma)\ell$ satisfies

$$f_{(t)}=(A(\sigma)\ell)_{(t)},\quad ^\forall t\in(-\infty,T]\quad(19)$$
$$(W(\sigma^{-1})\bar{w}+F(\sigma^{-1})f+X(\sigma)\ell)_{(T+1)}=0,\quad(20)$$

Eq.(7) is minimized.

Proof: First, calculate Eq.(7) as follows.

$$\begin{aligned}
&\sum_{t=-\infty}^{T}\|w-M(\sigma)\ell\|_{(t)}^2\\
&=\sum_{t=-\infty}^{T}\left(\|w\|^2-\|f\|^2+\|f\|^2-2\langle w,M(\sigma)\ell\rangle+\|M(\sigma)\ell\|^2\right)_{(t)}\\
&=\sum_{t=-\infty}^{T}\left(\|w\|^2-\|f\|^2+\|f\|^2\right)_{(t)}+2\left\langle W(\sigma^{-1})w,X(\sigma)\ell\right\rangle_{(T+1)}\\
&\quad-2\sum_{t=-\infty}^{T}\left\langle M(\sigma^{-1})^{\mathrm{T}}w,\ell\right\rangle_{(t)}+\sum_{t=-\infty}^{T}\|A(\sigma)\ell\|_{(t)}^2+\|X(\sigma)\ell\|_{(T+1)}^2\\
&=\sum_{t=-\infty}^{T}(\|w\|^2-\|f\|^2)_{(t)}+\sum_{t=-\infty}^{T}\|f-A(\sigma)\ell\|_{(t)}^2\\
&\quad+2\left\langle W(\sigma^{-1})w,X(\sigma)\ell\right\rangle_{(T+1)}-2\sum_{t=-\infty}^{T}\left\langle M(\sigma^{-1})^{\mathrm{T}}w,\ell\right\rangle_{(t)}\\
&\quad+2\sum_{t=-\infty}^{T}\langle f,A(\sigma)\ell\rangle_{(t)}+\|(X(\sigma)\ell)\|_{(T+1)}^2\\
&=\sum_{t=-\infty}^{T}(\|w\|^2-\|f\|^2)_{(t)}+\sum_{t=-\infty}^{T}\|f-A(\sigma)\ell\|_{(t)}^2\\
&\quad+2\left\langle W(\sigma^{-1})w,X(\sigma)\ell\right\rangle_{(T+1)}-2\sum_{t=-\infty}^{T}\left\langle A(\sigma^{-1})^{T}f,\ell\right\rangle_{(t)}\\
&\quad+2\sum_{t=-\infty}^{T}\langle f,A(\sigma)\ell\rangle_{(t)}+\|(X(\sigma)\ell)\|_{(T+1)}^2\\
&=\sum_{t=-\infty}^{T}(\|w\|^2-\|f\|^2)_{(t)}+\sum_{t=-\infty}^{T}\|f-A(\sigma)\ell\|_{(t)}^2\\
&\quad+2\left\langle W(\sigma^{-1})w,X(\sigma)\ell\right\rangle_{(T+1)}+\|X(\sigma)\ell\|_{(T+1)}^2\\
&\quad+2\left\langle F(\sigma^{-1})f,X(\sigma)\ell)\right\rangle_{(T+1)}\\
&=\sum_{t=\infty}^{T}(\|w\|^2-\|f\|^2)_{(t)}+\sum_{t=-\infty}^{T}\|f-A(\sigma)\ell\|_{(t)}^2\\
&\quad+\|W(\sigma^{-1})w+F(\sigma^{-1})f+X(\sigma)\ell\|_{(T+1)}^2\\
&\quad-\|W(\sigma^{-1})w+F(\sigma^{-1})f\|_{(T+1)}^2
\end{aligned}\quad(21)$$

In the above manipulations, we use Eq.(14) and Eq.(16) in the second equation, complete the square of $\sum_{t=-\infty}^{T}\|f-A(\sigma)\ell\|_{(t)}^2$ in the third equation, use Eq.(8) and Eq.(15) in the fourth and fifth equation respectively, and finally complete the square of $\|W(\sigma^{-1})\bar{w}+F(\sigma^{-1})f+X(\sigma)\ell\|_{(T+1)}^2$ in the last equation. Since the problem is to minimize the cost function described by Eq.(7) with respect to the latent variable ℓ, we see that Eq.(7) is minimized if Eq.(19) and Eq.(20) holds . ■

From Theorem 6, the optimal estimated variable w^* is described by using the latent variable ℓ^* satisfying Eq.(19) and Eq.(20) as

$$w^*_{(T)} = (M(\sigma)\ell^*)_{(T)}. \tag{22}$$

V. The structure of the optimal filter

Focusing on Eq.(20) enables us to obtain the following observation which plays a crucial roles in the implementation of the deterministic filter (Of course, this theorem corresponds to Proposition 7 of [1] in the continuous time, however the proof here in the discrete time case is quite different from the continuous time).

Theorem 7: Let $W(\xi^{-1})$ and $F(\xi^{-1})$ be di-polynomial matrices induced by (12) and (13), respectively. Then

$$z := -W(\sigma^{-1})\bar{w} - F(\sigma^{-1})f \tag{23}$$

is the minimal state variable of the behavior $\mathfrak{B}_{fw}$ described by Eq.(8).

Proof: First, it is easy to see that

$$X'(\xi) := \begin{bmatrix} I_{\mathtt{d}} & \xi I_{\mathtt{d}} & \cdots & \xi^{L-1} I_{\mathtt{d}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{\mathtt{Ld}\times\mathtt{d}}[\xi] \tag{24}$$

is one of the (not necessarily minimal) state maps of $\mathfrak{B}$. Let $X(\xi) \in \mathbb{R}^{\mathtt{n}\times\mathtt{d}}[\xi]$ be the minimal state map for $\mathfrak{B}$ acting on ℓ. Then, it is easy to show that there exists a full row rank matrix $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{Ld}}$ such that $X(\xi) = VX'(\xi)$. Applying this relation to Eq.(12) and Eq.(13) yields

$$(\zeta\eta - 1)W(\zeta^{-1})^{\mathrm{T}} VX'(\eta) = M(\zeta^{-1}) - M(\eta) \tag{25}$$

$$(\zeta\eta - 1)F(\zeta^{-1})^{\mathrm{T}} VX'(\eta) = -A(\zeta^{-1}) + A(\eta) \tag{26}$$

Substituting Eq. (24) into Eq.(25) and Eq.(26) yields

$$V^{\mathrm{T}}W(\xi^{-1}) = -\begin{bmatrix} M_L^{\mathrm{T}}\xi^{-L} + \cdots + M_1^{\mathrm{T}}\xi^{-1} \\ M_L^{\mathrm{T}}\xi^{-L+1} + \cdots + M_2^{\mathrm{T}}\xi^{-1} \\ \vdots \\ M_L^{\mathrm{T}}\xi^{-1} \end{bmatrix} \tag{27}$$

$$V^{\mathrm{T}}F(\xi^{-1}) = \begin{bmatrix} A_L^{\mathrm{T}}\xi^{-L} + \cdots + A_1^{\mathrm{T}}\xi^{-1} \\ A_L^{\mathrm{T}}\xi^{-L+1} + \cdots + A_2^{\mathrm{T}}\xi^{-1} \\ \vdots \\ A_L^{\mathrm{T}}\xi^{-1} \end{bmatrix}. \tag{28}$$

From Eq. (27) and Eq.(28), we obtain

$$\begin{aligned} &V^{\mathrm{T}} \begin{bmatrix} W(\sigma^{-1}) & F(\sigma^{-1}) \end{bmatrix} \begin{bmatrix} \bar{w} \\ f \end{bmatrix}_{(t)} \\ &= \begin{bmatrix} M_0^{\mathrm{T}}\bar{w}_{(t)} \\ M_0^{\mathrm{T}}\bar{w}_{(t+1)} + M_1^{\mathrm{T}}\bar{w}_{(t)} \\ \vdots \\ M_0^{\mathrm{T}}\bar{w}_{(t+L-1)} + \cdots + M_{L-1}^{\mathrm{T}}\bar{w}_{(t)} \end{bmatrix} \\ &\quad - \begin{bmatrix} A_0^{\mathrm{T}} f_{(t)} \\ A_0^{\mathrm{T}} f_{(t+1)} + A_1^{\mathrm{T}} f_{(t)} \\ \vdots \\ A_0^{\mathrm{T}} f_{(t+L-1)} + \cdots + A_{L-1}^{\mathrm{T}} f_{(t)} \end{bmatrix} \end{aligned} \tag{29}$$

along $(\bar{w}, f) \in \mathfrak{B}_{fw}$ for any time $t \in \mathbb{Z}$. Now, by using the result on discrete time state map of [8], we see that the right hand side of Eq.(29) is the state map of the behavior $\mathfrak{B}_{fw}$. Moreover, by applying Theorem 6.2 of [9] to the discrete time case and noting that V^{T} is full column rank, we also see that the $\begin{bmatrix} W(\xi^{-1}) & F(\xi^{-1}) \end{bmatrix}$ induces the minimal state map on $\mathfrak{B}_{fw}$. ■

From Theorem 6, the latent variable ℓ satisfying Eq.(20) must also satisfy Eq.(19) and (22), that is,

$$\begin{bmatrix} w^* \\ f \end{bmatrix} = \begin{bmatrix} M(\sigma) \\ A(\sigma) \end{bmatrix} \ell. \tag{30}$$

Let $\mathfrak{B}^*_{wf}$ be the behavior of a system described by Eq.(30). Then, we also obtain the following theorem, which also plays a crucial role in the implementation of the optimal filter deterministically.

Theorem 8: The minimal state map $X(\xi)$ for $\mathfrak{B}$ is also the minimal state map of the behavior $\mathfrak{B}^*_{wf}$ described by Eq.(30).

Proof: First, from the spectral factorization, we obtain

$$U(\xi^{-1})^{\mathrm{T}}U(\xi) + Y(\xi^{-1})^{\mathrm{T}}Y(\xi) = A(\xi^{-1})^{\mathrm{T}}A(\xi). \tag{31}$$

Note the assumption that $Y(\xi)U(\xi)^{-1}$ is proper and $U(0)$ is nonsingular. Rewrite Eq.(31) as

$$U(\xi^{-1})^{\mathrm{T}} + Y(\xi^{-1})^{\mathrm{T}}Y(\xi)U(\xi)^{-1} = A(\xi^{-1})^{\mathrm{T}}A(\xi)U(\xi)^{-1}. \tag{32}$$

From the properness of $Y(\xi)U(\xi)^{-1}$ and $Y(\xi^{-1})$, we see $Y(\xi^{-1})^{\mathrm{T}}Y(\xi)U(\xi)^{-1}$ is also proper. Moreover, it follows from the nonsingularities of $A(0)$ and $U(0)$ that the left hand side of Eq.(32) is proper and $A(\xi^{-1})^{\mathrm{T}}$ is bi-proper. This implies that $A(\xi)U(\xi)^{-1}$ is also proper. By using this fact and applying Section 8 in [9] to the discrete time case, $X(\xi)$ is also the minimal state map for $\mathfrak{B}^*_{wf}$. (We omit the detailed proof here). ■

Combined with Theorem 7 and Theorem 8, the optimal condition described by Eq.(20) shows that the equivalence between the minimal state variable for $\mathfrak{B}_{fw}$ described by Eq.(8) and that for $\mathfrak{B}^*_{wf}$ described by Eq.(30) is required.

VI. Implementation of the optimal filter

Similarly to [1], Theorem 7 and Theorem 8 enables us to obtain one of the minimal state space representation of $\mathfrak{B}_{fw}$ by using Eq.(20)

$$\begin{aligned} z_{(T+1)} = (X(\sigma)\ell)_{(T+1)} &= -(W(\sigma^{-1}\bar{w} + F(\sigma^{-1})f)_{(T+1)} \\ &= A_{fw}z_{(T)} + B_{fw}\bar{w}_{(T)} \end{aligned} \tag{33}$$

$$f_{(T)} = C_{fw}z_{(T)} + D'\bar{w}_{(T)} \tag{34}$$

where $A_{fw} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $B_{fw} \in \mathbb{R}^{\mathtt{n}\times\mathtt{q}}$, $C_{fw} \in \mathbb{R}^{\mathtt{d}\times\mathtt{n}}$ and $C_{fw} \in \mathbb{R}^{\mathtt{d}\times\mathtt{q}}$ are appropriate matrices. From Theorem 8, $X(\xi)$ is also a minimal state map for $\mathfrak{B}^*_{wf}$, so we see that there exist appropriate constant matrices $C^*_{wf} \in \mathbb{R}^{\mathtt{q}\times\mathtt{n}}$, $D^*_{wf} \in \mathbb{R}^{\mathtt{q}\times\mathtt{d}}$ such that

$$w^*_{(T)} = C^*_{wf}(X(\sigma)\ell)_{(T)} + D^*_{wf}f_{(T)} \tag{35}$$

obeying Eq.(30).

Summing up the above discussions, the optimal filter can be implemented by using the recursive state space representation Eq.(33) and Eq.(34) in order to estimate the minimal state of the system and f. Then, substituting this $f_{(T)}$ and $x_{(T)} := (X(\sigma)\ell)_{(T)}$ into Eq.(35) yields the optimal estimated variable $w^*_{(T)}$ at time T.

VII. Example

Consider $\mathfrak{B}$ described by

$$w := \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} := \begin{bmatrix} \sigma^2 + \frac{5}{6}\sigma + \frac{1}{6} \\ \sigma^2 + \frac{5}{2}\sigma + 1 \end{bmatrix} \ell. \tag{36}$$

Suppose that the observation $\bar{w} = [w_1 + n_1 \ \ w_2 + n_2]$, where $n_{1(t)} = \sin(0.01t) + \sin(0.5t) + 0.05v_{1(t)}$, $n_{2(t)} = 2\sin(0.1t) + 0.02v_{2(t)}$, and v_1 and v_2 are white noise with the maximal amplitude 0.05 (they has no correlation each other). Thus, the observation $\bar{w}$ does not obey Eq.(36). For $\mathfrak{B}$ described by Eq.(36), we choose one of the minimal state maps for $\mathfrak{B}$ as $X(\xi) = \begin{bmatrix} 1 & \xi \end{bmatrix}^{\mathrm{T}}$ which implies that the minimal state is described by

$$x_{(t)} := \begin{bmatrix} x_{1(t)} \\ x_{2(t)} \end{bmatrix} := (X(\sigma)\ell)_{(t)} = \begin{bmatrix} \ell_{(t)} \\ \ell_{(t+1)} \end{bmatrix} \tag{37}$$

by using the latent variable ℓ of Eq.(36). Next, we compute the spectral factorization and then we obtain the anti-Hurwitz spectral factor $A(\xi)$ of $M(\xi^{-1})^{\mathrm{T}}M(\xi)$ as $A(\xi) = 0.5203\xi^2 + 2.1618\xi + 2.2424$. By applying $M(\xi)$ and $A(\xi)$ to Eq.(12) and Eq.(13) respectively, we obtain $W(\xi^{-1}) \in \mathbb{R}^{2\times 2}[\xi^{-1}]$ $F(\xi^{-1}) \in \mathbb{R}^{2\times 1}[\xi^{-1}]$, respectively. In this example, note that the projection matrix V used in the previous section is I. By using the above polynomial matrices, we can calculate constant matrices required in Eq.(33), Eq.(34), and Eq.(35). In Figure 1, we show the observation $\bar{w} := [w_1 \ w_2]^{\mathrm{T}}$ and the estimation w^*. As shown in these figures, the deterministic filtering proposed here can be *almost completely* estimate the manifest variable of the system via the observation. Next, in Figure 2, we also show the real state variables and the estimated state variables implemented by using Eq.(33) recursively. This figure also illustrates that the filter estimates optimally the state variables.

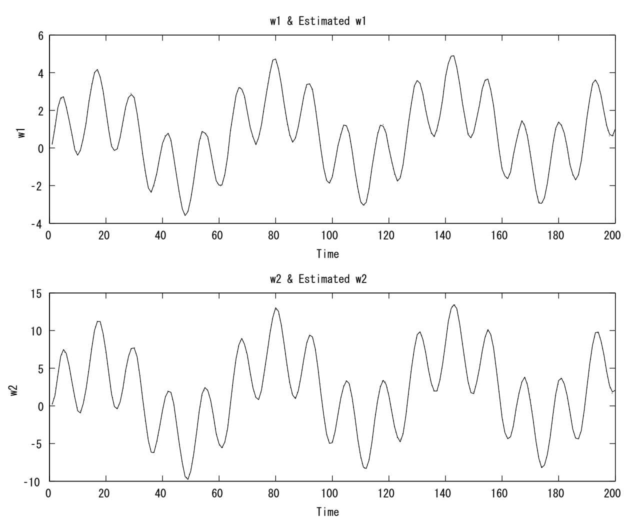

Fig. 1. The estimated manifest variables (the solid lines) and the real ones (the broken lines)

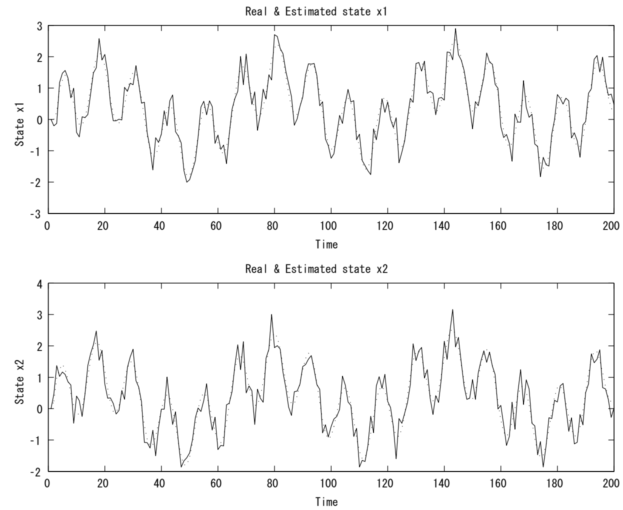

Fig. 2. The estimated state variables (the solid lines) and the real ones (the broken lines)

VIII. Conclusions

In this paper, we have studied the discrete-time deterministic filtering in the behavioral setting . Future works are to eliminate the assumptions imposed to derive our results in this paper and to exploit the $\mathcal{H}_\infty$ filtering problem.

References

[1] F. Fagnani and J. C. Willems: Deterministic Kalman filtering in a behavioral framework; *Syst. and Contr. Lett.*, Vol. 32, pp. 302–312 (1997)

[2] C. Heij: *Deterministic Identification of Dynamical Systems*, Springer-Verlag, (1989)

[3] O. Kaneko and T. Fujii: Discrete time average positivity and spectral factorization in a behavioral framework; *Syst. and Contr. Lett.*, Vol. 39, pp. 31–44 (2000)

[4] O. Kaneko and T. Fujii: Discrete time dissipativeness on the dipolynomial ring; *Proc. of 15th IFAC World Congress*, CD-ROM (2002)

[5] O. Kaneko and T. Fujii: When is a storage function a state function in discrete time? ; *SIAM Journal on Control and Optimization*, Vol. 42, pp. 1374–1394 (2003)

[6] O. Kaneko: The behavioral approach to the discrete time deterministic Kalman Filtering on two-variable di-polynomial matrices; *Syst. and Contr. Lett.*, submitted (2005)

[7] V. M. Popov: *Hyperstability of control systems*, Springer-Verlag, (1973)

[8] C. Praagman. Inputs, output and the states in the representation of time series. *Analysis and Optimization of Systems (A. Bensoussan and J.Lions eds.)*, pp. 1069–1078, Springer Verlag, Berlin, (1988)

[9] P. Rapisarda and J. C. Willems: State maps for linear systems. *SIAM J. on Contr. and Optimiz.*, Vol. 35, pp. 1053–1091, (1997)

[10] H. L. Trentelman, and J. C. Willems; Every storage function is a state function: *Systems and Control Letters*, Vol. 39, pp. 31–44, (1997)

[11] J. C. Willems: Paradigms and puzzles in the theory of dynamical systems; *IEEE. Trans. Automat. Cont.*, Vol. 36, No. 3, pp. 259–294 (1991)

[12] J. C. Willems and F. Fagnani: Deterministic Kalman filtering in a behavioral setting; *Proc. of 3rd European, Contr. Conf.*, Vol. 32, pp. 2525–2533, (1995)

[13] J. C. Willems and H. L. Tremtelman: On quadratic differential forms; *SIAM J. on Contr. and Optimiz.*, Vol.36, Vol.5, pp. 1703–1749 (1998)