Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Subspace identification algorithm for stochastic systems equipped with zeros close to unit circle

Hideyuki Tanaka * **Kenji Ikeda** **

* Graduate School of Education, Hiroshima University, Hiroshima, Japan (e-mail: tanakalpha@hiroshima-u.ac.jp)
** Graduate School of Technology and Science, Tokushima University, Tokushima, Japan (e-mail: ikeda@is.tokushima-u.ac.jp)

Abstract: In identifying a stochastic system possessing zeros close to the unit circle, the effect of the initial state appears in the estimates. This paper derives a stochastic subspace identification algorithm for such a system. A new stochastic realization algorithm is developed based on the covariance matrices of the state and the white-noise input, by taking the initial state and positive realness into account. A subspace identification algorithm is obtained by applying the realization algorithm to a finite string of data. Numerical simulation results show that the proposed algorithm provides favorable results compared with the conventional ones.

Keywords: stochastic realization, stochastic subspace identification, positive realness, initial state, Kalman gain

1. INTRODUCTION

Stochastic subspace identification algorithms identify linear stochastic systems directly from a finite string of time-series data (Van Overschee and De Moor, 1993, 1996). Stochastic realization algorithms (e.g. Faurre, 1976; Akaike, 1975) have formed a theoretical basis for stochastic subspace identification algorithms. Stochastic realization theory is constructed based on stochastic variables or the exact covariance data, whereas stochastic subspace identification is on the finite string of data.

Positive realness is one of important problems in stochastic systems identification (Vaccaro and Vukina, 1993; Lindquist and Picci, 1996; Dahlén et al., 1998), and the problem stems from inexact covariance matrices based on a finite string of data. Although the stochastic subspace identification algorithm given by Van Overschee and De Moor (1993) did not guarantee positive realness, they (1996) developed an algorithm for ensuring the property based on the residual of the state estimate. Mari et al. (2000) and Goethals et al. (2003) moreover developed algorithms via semi-definite programming and regularization, respectively. Akçay and Türkay (2015) proposed a regularized and reweighted nuclear norm minimization approach.

Subspace methods have often ignored the presence of the initial effect due to a finite string of data. However, Bauer (2005) pointed out that the effect might be problematic, if careful modeling of initial conditions is crucial. The effect of the initial state becomes large, if the stochastic part of the system has zeros close to the unit circle. We have studied subspace identification, deleting effects of the initial state and using the observability matrix (Ikeda, 2015; Ikeda and Tanaka, 2017; Tanaka and Ikeda, 2018), though the settins are different from this paper.

The main purpose of this paper is to develop a numerically sound algorithm for identifying stochastic systems equipped with zeros close to the unit circle. We therefore develop a stochastic realization algorithm in order to take the initial state and positive realness into account, by using the observability matrix. We then derive a stochastic subspace identification algorithm.

2. PROBLEM SETTING AND POSITIVE REALNESS

We state the problem setting, reviewing positive realness.

2.1 Stochastic system

Consider a linear-time-invariant stochastic system:

$$x_{t+1} = Ax_t + w_t, \tag{1a}$$
$$y_t = Cx_t + v_t, \tag{1b}$$

where $x_t \in \mathbb{R}^n$ and $y_t \in \mathbb{R}^p$ are respectively the state and output of the system, and A is stable. The signals $w_t \in \mathbb{R}^n$ and $v_t \in \mathbb{R}^p$ are white noises with zero mean satisfying

$$\mathrm{E}\left\{ \begin{bmatrix} w_s \\ v_s \end{bmatrix} \begin{bmatrix} w_t \\ v_t \end{bmatrix}^\top \right\} = \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \delta_{st} \geq 0, \tag{2}$$

where δ_{ij} is the Kronecker delta satisfying $\delta_{ii} = 1$ and $\delta_{ij} = 0$ $(i \neq j)$, and where we suppose that $R > 0$ holds. Let us describe the covariance matrix of y_t as

$$\Lambda_k = \mathrm{E}\left\{ y_{t+k} y_t^\top \right\}$$

and express the covariance between x_{t+1} and y_t by $G = \mathrm{E}\left\{ x_{t+1} y_t^\top \right\}$. We then have a decomposition

$$\Lambda_k = CA^{k-1}G \quad (k \geq 1) \tag{3}$$

and suppose that $(A,\, G,\, C)$ is a minimal realization. The spectral density function of y_t is given by

$$\Phi(z) = \sum_{k=-\infty}^{\infty} \Lambda_k z^{-k}$$

and satisfies the following equations from (1) and (2):

$$\Phi(z) = \begin{bmatrix} (C(zI-A)^{-1})^\top \\ I_p \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} (C(z^{-1}I-A)^{-1})^\top \\ I_p \end{bmatrix} \quad (4)$$
$$= \Psi(z) + \Psi^\top(z^{-1}),$$

where $\Psi(z) := C(zI_n - A)^{-1}G + \frac{1}{2}\Lambda_0$ is the causal part of $\Phi(z)$. We suppose that $\Psi(z)$ is positive real, i.e. $\Phi(z) > 0$ for $|z| = 1$. Consider the following Riccati equation:

$$P = APA^\top + (G - APC^\top)(\Lambda_0 - CPC^\top)^{-1}(G - APC^\top)^\top. \quad (5)$$

Using the stabilizing solution $\hat{P}$ to (5), we define

$$\hat{K} = (G - A\hat{P}C^\top)(\Lambda_0 - C\hat{P}C^\top)^{-1}, \quad (6a)$$
$$\hat{\Omega} = \Lambda_0 - C\hat{P}C^\top, \quad (6b)$$

where $A - \hat{K}C$ is stable, and we call $\hat{K}$ and $\hat{\Omega}$ respectively the Kalman gain and the innovation covariance. The innovation form of y_t is as follows:

$$\hat{x}_{t+1} = A\hat{x}_t + \hat{K}e_t, \quad (7a)$$
$$y_t = C\hat{x}_t + e_t, \quad (7b)$$
$$\mathrm{E}\{e_s e_t^\top\} = \hat{\Omega}\delta_{st}, \quad (7c)$$

where e_t is the innovation of y_t and is a white noise. The spectral density function of $\Phi(z)$ has the factorization

$$\Phi(z) = \hat{W}(z)\hat{W}^\top(z^{-1}), \quad (8)$$

where $\hat{W}(z)$ is given by

$$\hat{W}(z) = \left(C(zI_n - A)^{-1}\hat{K} + I_p\right)\hat{\Omega}^{\frac{1}{2}}. \quad (9)$$

The spectral function $\Psi(z)$ should be positive real to have the spectral factor $\hat{W}(z)$, and positive realness stems from the positive semi-definiteness of the covariance matrix in (2) because of (4).

Subspace identification algorithms have often ignored the effect of the initial state in estimating the state, and the effect might be problematic (Bauer, 2005). Let us explain this fact using stochastic system (7). Defining a matrix

$$F := A - \hat{K}C, \quad (10)$$

we have $\hat{x}_{t+1} = F\hat{x}_t + \hat{K}y_t$ from (7) and (10), and we obtain the following equation:

$$\hat{x}_\tau = F^\tau \hat{x}_0 + \begin{bmatrix} F^{\tau-1}\hat{K} & F^{\tau-2}\hat{K} & \cdots & \hat{K} \end{bmatrix} Y_{0:\tau-1},$$

where $Y_{i:j}$ is defined as

$$Y_{i:j} := \begin{bmatrix} y_i^\top & y_{i+1}^\top & \cdots & y_j^\top \end{bmatrix}^\top.$$

If $F^\tau \approx 0$, we have an approximation to the state $\hat{x}_\tau$ by means of a finite number of the linear combination of $y_0, y_1, \ldots, y_{\tau-1}$. However, if $F^\tau \approx 0$ does *not* hold, or if $\hat{W}(z)$ has zeros close to the unit circle, $F^\tau\hat{x}_0$ does not disappear and leads to biased estimates. This type of approximation has been made by Van Overschee and De Moor (1993, 1994); Jansson (2003); Chiuso and Picci (2005) in developing subspace identification algorithms.

Given a finite string of y_t, we will derive a numerically sound algorithm for identifying the system (7) that has a property $F^\tau \not\approx 0$. We newly develop a stochastic realization algorithm to this end, given a finite interval of data $\{\Lambda_0, \ldots, \Lambda_{2\tau-1}\}$, and supposing that the following matrices are of full rank for $k = \tau - 1$:

$$\mathcal{O}_k := \begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{k-1})^\top \end{bmatrix}^\top \in \mathbb{R}^{kp\times n},$$
$$\mathcal{C}_k := \begin{bmatrix} A^{k-1}G & A^{k-2}G & \cdots & G \end{bmatrix} \in \mathbb{R}^{n\times kp},$$

where $\mathcal{O}_\tau$ and $\mathcal{C}_\tau$ are respectively the observability and reachabilty matrices. We also take positive realness into account to derive a stochastic subspace identification method. In the rest of section 2, we will explain the differences between conventional algorithms and our approach, by reviewing the positive real problem.

2.2 Review of positive real problem

We review stochastic realization (Faurre, 1976), stochastic subspace identification (Van Overschee and De Moor, 1993), and the positive real problem (Lindquist and Picci, 1996); interested readers may see Katayama (2005) for stochastic realization. Let us describe

$$\Theta_k := \mathrm{E}\{Y_{0:k-1}Y_{0:k-1}^\top\},$$
$$\mathcal{H}_\tau := \mathrm{E}\{Y_{\tau:2\tau-1}Y_{0:\tau-1}^\top\}.$$

In case of stochastic realization, we have the exact covariance data Λ_k and obtain the decomposition from (3):

$$\mathcal{H}_\tau = \mathcal{O}_\tau\mathcal{C}_\tau \in \mathbb{R}^{\tau p\times\tau p} \quad (11)$$

for large enough $\tau > n$ and the realization (A, G, C, Λ_0) from $\mathcal{O}_\tau$ and $\mathcal{C}_\tau$ (Ho and Kalman, 1966). Defining

$$X = \mathrm{E}\{x_t x_t^\top\} \quad (12)$$

for (1), we have $X > 0$ satisfying

$$X = AXA^\top + Q, \quad (13a)$$
$$G = AXC^\top + S, \quad (13b)$$
$$\Lambda_0 = CXC^\top + R, \quad (13c)$$

and hence the LMI (Linear Matrix Inequality) from (2):

$$\begin{bmatrix} X - AXA^\top & G - AXC^\top \\ (G - AXC^\top)^\top & \Lambda_0 - CXC^\top \end{bmatrix} \geq 0, \quad (14)$$

where $R = \Lambda_0 - CXC^\top > 0$. It is well known that $\Psi(z)$ is positive real, if and only if there exists a solution to (14) and that the stabilizing solution to (5) is the minimum solution to (14): $\hat{P} \leq X$. In case of stochastic subspace identification (Van Overschee and De Moor, 1993), we have only approximations $\tilde{\Lambda}_k$ and $(\tilde{A}, \tilde{G}, \tilde{C}, \tilde{\Lambda}_0)$ respectively to the covariance matrices Λ_k and the realization (A, G, C, Λ_0), because only a finite string of y_t is available. The approximation does not necessarily ensure existence of the solution to the LMI:

$$\begin{bmatrix} X - \tilde{A}X\tilde{A}^\top & \tilde{G} - \tilde{A}X\tilde{C}^\top \\ (\tilde{G} - \tilde{A}X\tilde{C}^\top)^\top & \tilde{\Lambda}_0 - \tilde{C}X\tilde{C}^\top \end{bmatrix} \geq 0, \quad (15)$$

meaning that $\tilde{\Psi}(z) := \tilde{C}(zI - \tilde{A})^{-1}\tilde{G} + \frac{1}{2}\tilde{\Lambda}_0$ may not be positive real and that the Riccati equation

$$P = \tilde{A}P\tilde{A}^\top + (\tilde{G} - \tilde{A}P\tilde{C}^\top)(\tilde{\Lambda}_0 - \tilde{C}P\tilde{C}^\top)^{-1}(\tilde{G} - \tilde{A}P\tilde{C}^\top)^\top$$

may not have the stabilizing solution. The statistical problem of stochastic modeling from estimated covarianced is phrased in Lindquist and Picci (1996).

Van Overschee and De Moor (1996) proposed an algorithm for ensuring positive realness based on positive semi-definiteness of the residual of the estimated state. The stochastic subspace identification algorithm (Van Overschee and De Moor, 1996) estimates the state $\tilde{\boldsymbol{x}}_\tau \in \mathbb{R}^{\tilde{n}\times N}$ via the orthogonal projection and solves the set of linear equations for A and C:

$$\begin{bmatrix} \tilde{\boldsymbol{x}}_{\tau+1} \\ \boldsymbol{y}_\tau \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} \tilde{\boldsymbol{x}}_\tau + \begin{bmatrix} \boldsymbol{\rho}_w \\ \boldsymbol{\rho}_v \end{bmatrix}, \quad (16)$$

where $\boldsymbol{y}_t = [\, y_t \cdots y_{t+N-1} \,] \in \mathbb{R}^{p\times N}$. The covariance matrices of the residuals $\boldsymbol{\rho}_w \in \mathbb{R}^{\tilde{n}\times N}$ and $\boldsymbol{\rho}_v \in \mathbb{R}^{p\times N}$ are then estimated as follows:

$$\begin{bmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^\top & \tilde{R} \end{bmatrix} := \frac{1}{N} \begin{bmatrix} \boldsymbol{\rho}_w \\ \boldsymbol{\rho}_v \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_w \\ \boldsymbol{\rho}_v \end{bmatrix}^\top \geq 0. \tag{17}$$

Describing the solutions A and C to (16) respectively as $\tilde{A}$ and $\tilde{C}$, we solve the Lyapunov equation (18a) and estimate G and Λ_0 respectively as (18b) and (18c) based on (13):

$$X = \tilde{A}X\tilde{A}^\top + \tilde{Q}, \tag{18a}$$

$$\tilde{G} := AXC^\top + \tilde{S}, \tag{18b}$$

$$\tilde{\Lambda}_0 := \tilde{C}X\tilde{C}^\top + \tilde{R}. \tag{18c}$$

Since $\tilde{Q}$, $\tilde{S}$, and $\tilde{R}$ satisfy positive semi-definiteness (17), X satisfies the LMI (15). The stochastic subspace identification algorithm developed by Van Overschee and De Moor (1996) carries out numerically stable computations in estimating the spectral factor. Unfortunately, it gives biased estimates in identifying the stochastic system for $F^\tau \not\approx 0$, because the estimates $\tilde{Q}$, $\tilde{S}$ and $\tilde{R}$ are biased due to the approximation error of the state $\tilde{\boldsymbol{x}}_\tau$.

Conventional algorithms ensure positive realness by modifying matrices $(\tilde{A}, \tilde{G}, \tilde{C}, \tilde{\Lambda}_0)$. Vaccaro and Vukina (1993) showed how to modify the covariance model to guarantee positive realness. Mari et al. (2000) developed a convex optimization algorithm for guaranteeing positive realness. Mari (2000) proposed an algorithm for obtaining a stable model for $\Psi(z)$ by modifying $\tilde{A}$, and summarized three types of LMI algorithms for guaranteeing positive realness. Goethals et al. (2003) presented an algorithm for imposing positive realness by adding a regularization term. Akçay and Türkay (2015) proposed a method for transforming a non-positive real transfer function matrix into a positive real one. The algorithms (Mari, 2000; Mari et al., 2000; Goethals et al., 2003) for modifying rational transfer matrices $(\tilde{A}, \tilde{G}, \tilde{C}, \tilde{\Lambda}_0)$ do *not* suffer from the approximation error due to $F^\tau \not\approx 0$, since they do not need to make the approximation $F^\tau \approx 0$ in estimating $(\tilde{A}, \tilde{G}, \tilde{C})$. However, the algorithms for modifying rational transfer matrices possibly demand numerically severe computations, because they seek modification of the matrices of $\tilde{\Psi}(z)$ by checking whether it is positive real or not.

In this paper, we develop a stochastic realization algorithm avoiding explicit use of the LMI (14) to guarantee positive realness. We rather find (Q, S, R) satisfying the positive semi-definite condition in (2), using the observability matrix $\mathcal{O}_\tau$ and the covariance data $\{\Lambda_0, \ldots, \Lambda_{2\tau-1}\}$. We thus derive a numerically sound algorithm for stochastic realization without making the approximation $F^\tau \approx 0$. By applying the realization algorithm to a finite string of data y_t, we moreover derive a stochastic subspace identification algorithm guaranteeing positive realness.

3. STOCHASTIC REALIZATION

Suppose that a finite interval of the exact covariance matrices $\{\Lambda_0, \ldots, \Lambda_{2\tau-1}\}$ or $\Theta_{2\tau}$ is given. We introduce stochastic balancing (e.g. Desai and Pal, 1984) and the Kalman filter as preliminaries. We then derive a new stochastic realization algorithm.

3.1 Preliminaries

We introduce stochastic balancing or the CCA (Canonical Correlation Analysis) weighting, since it leads optimal accuracy in the asymptotic variance (Bauer and Ljung, 2002). We partition $\Theta_{2\tau} \in \mathbb{R}^{2\tau p\times 2\tau p}$ as follows:

$$\Theta_{2\tau} = \begin{bmatrix} \Theta_\tau & \mathcal{H}_\tau^\top \\ \mathcal{H}_\tau & \Theta_\tau \end{bmatrix}.$$

Compute the Choleskey factorization:

$$\begin{bmatrix} \Theta_\tau & \mathcal{H}_\tau^\top \\ \mathcal{H}_\tau & \Theta_\tau \end{bmatrix} = \begin{bmatrix} L_{pp} & 0 \\ L_{fp} & L_{ff} \end{bmatrix} \begin{bmatrix} L_{pp} & 0 \\ L_{fp} & L_{ff} \end{bmatrix}^\top. \tag{19}$$

We then have $\Theta_\tau^{-\frac{1}{2}}\mathcal{H}_\tau\Theta_\tau^{-\frac{1}{2}} = \Theta_\tau^{-\frac{1}{2}}L_{fp}L_{pp}^\top\Theta_\tau^{-\frac{1}{2}}$ and calculate the singular value decomposition (SVD):

$$\begin{aligned} \Theta_\tau^{-\frac{1}{2}}L_{fp}L_{pp}^\top\Theta_\tau^{-\frac{1}{2}} &= [\, U_1 \; U_2 \,] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix} \\ &= U_1\Sigma_1V_1^\top \quad (\Sigma_1 \in \mathbb{R}^{n\times n}). \end{aligned} \tag{20}$$

From (11), we thus have $\mathcal{C}_\tau = \Sigma_1^{\frac{1}{2}}V_1^\top\Theta_\tau^{\frac{1}{2}}$,

$$\mathcal{O}_\tau = \Theta_\tau^{\frac{1}{2}}U_1\Sigma_1^{\frac{1}{2}} \in \mathbb{R}^{\tau p\times n}, \tag{21}$$

and hence $(A,\, C)$ via shift invariance. It should be noted that (20) leads to a finite-interval stochastically balanced realization (Lindquist and Picci, 1996) and that the stochastic balancing (Desai and Pal, 1984) is obtained for $\tau \to \infty$.

Given the system (1) with $(A,\, C)$ and $(Q,\, S,\, R)$, we can construct the Kalman filter. Compute the stabilizing solution $\hat{Y}$ to the following Riccati equation:

$$\begin{aligned} Y = AYA^\top &- (AYC^\top + S) \\ &(CYC^\top + R)^{-1}(AYC^\top + S)^\top + Q. \end{aligned} \tag{22}$$

Determine the Kalman gain and the innovation covariance as follows:

$$\hat{K} = (A\hat{Y}C^\top + S)(C\hat{Y}C^\top + R)^{-1}, \tag{23a}$$

$$\hat{\Omega} = C\hat{Y}C^\top + R, \tag{23b}$$

where $\hat{K}$ and $\hat{\Omega}$ in (23) are respectively the same as (6), from (5), (13), (22), and $P = X - Y$. The Kalman filter

$$\hat{x}_{t+1} = A\hat{x}_t + \hat{K}(y_t - C\hat{x}_t)$$

minimizes $\mathrm{Trace}\,(\mathrm{E}(x_t - \hat{x}_t)(x_t - \hat{x}_t)^\top)$ and gives the innovation $e_t = y_t - C\hat{x}_t$. We thus have the innovation form (7), if we can compute $(A,\, C)$ and $(Q,\, S,\, R)$.

3.2 Derivation of new stochastic realization algorithm

We derive a new stochastic realization algorithm by computing $(Q,\, S,\, R)$. Let us define the followings:

$$\begin{bmatrix} \Gamma \\ \hline H \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ \hline 0 & I_p \end{bmatrix}, \quad \xi_t = \begin{bmatrix} w_t \\ v_t \end{bmatrix}. \tag{24}$$

Describing $\Pi := \mathrm{E}\{\xi_t\xi_t^\top\}$, we have

$$\Pi = \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} = \begin{bmatrix} \Gamma \\ H \end{bmatrix} \Pi \begin{bmatrix} \Gamma \\ H \end{bmatrix}^\top \geq 0 \tag{25}$$

from (2) and (24). We define a matrix-valued function $\mathcal{T}_\tau(A,\, \Gamma,\, C,\, H)$ that has the block-Toeplitz structure

$$\mathcal{T}_\tau(A,\, \Gamma,\, C,\, H) := \begin{bmatrix} H & & 0 \\ C\Gamma & H & \\ \vdots & \ddots & \ddots \\ CA^{\tau-2}\Gamma & \cdots & C\Gamma \;\; H \end{bmatrix} \tag{26}$$

and the following matrix for a simple notation:

$$\bar{\mathcal{T}}_\tau := \mathcal{T}_\tau(A,\, \Gamma,\, C,\, H) \in \mathbb{R}^{\tau p \times \tau(n+p)}.$$

We have the following equation from (1), (24), and (26):

$$Y_{\tau:2\tau-1} = \mathcal{O}_\tau x_\tau + \bar{\mathcal{T}}_\tau \Xi_{\tau:2\tau-1}, \tag{27}$$

where $\Xi_{i:j} := \begin{bmatrix} \xi_i^\top & \xi_{i+1}^\top & \cdots & \xi_j^\top \end{bmatrix}^\top$. Define moreover a matrix-valued function for $M \in \mathbb{R}^{(n+p)\times(n+p)}$ as

$$\mathcal{D}_\tau(M) := \text{block-diag}(M,\, \ldots,\, M) \in \mathbb{R}^{\tau(n+p)\times\tau(n+p)}.$$

We then have the following equations for $\Xi_{\tau:2\tau-1}$:

$$\mathrm{E}\left\{x_\tau \Xi_{\tau:2\tau-1}^\top\right\} = 0, \tag{28a}$$

$$\mathrm{E}\left\{\Xi_{\tau:2\tau-1}\Xi_{\tau:2\tau-1}^\top\right\} = \mathcal{D}_\tau(\Pi). \tag{28b}$$

We obtain the following Theorem.

Theorem 1. The covariance matrices Θ_τ, X, and Π satisfy

$$\Theta_\tau = \mathcal{O}_\tau X \mathcal{O}_\tau^\top + \bar{\mathcal{T}}_\tau \mathcal{D}_\tau(\Pi)\bar{\mathcal{T}}_\tau^\top. \tag{29}$$

Proof . We have (29) from (12), (27), and (28). ■

We compute Π by solving (29). To this end, we minimize $\gamma \in \mathbb{R}$ under the following LMI constraints:

$$\gamma\Theta_\tau \geq \Theta_\tau - \left(\mathcal{O}_\tau X \mathcal{O}_\tau^\top + \bar{\mathcal{T}}_\tau \mathcal{D}_\tau(\Pi)\bar{\mathcal{T}}_\tau^\top\right) \geq 0 \tag{30a}$$

$$X \geq 0, \quad \Pi \geq 0, \tag{30b}$$

where variables are γ, $X \in \mathbb{R}^{n\times n}$, and $\Pi \in \mathbb{R}^{(n+p)\times(n+p)}$.

We propose a stochastic realization algorithm as follows.

Stochastic realization algorithm:

Step 1: Calculate the SVD (20) and $\mathcal{O}_\tau$ in (21).

Step 2: Obtain A and C using $\mathcal{O}_\tau$ and shift invariance and construct $\bar{\mathcal{T}}_\tau = \mathcal{T}_\tau(A,\, \Gamma,\, C,\, H)$ in (26).

Step 3: Compute Π in (30) by minimizing γ.

Step 4: Obtain $(Q,\, S,\, R)$ from (25). Find the stabilizing solution $\hat{Y}$ to (22) and obtain $\hat{K}$ and $\hat{\Omega}$ in (23).

The traditional stochastic realization algorithm computes (A, G, C, Λ_0) and solves (5), where $\Psi(z)$ should be positive real or there should exist solutions X to (14). On the other hand, the proposed algorithm finds Π in (29) and solves (22) subject to $X \geq 0$ and $\Pi \geq 0$. It should be noted that Π is computed from $\mathcal{O}_\tau$ and $\{\Lambda_0,\, \ldots,\, \Lambda_{2\tau-1}\}$ and that there exists the Kalman filter because of $\Pi \geq 0$.

4. STOCHASTIC SUBSPACE IDENTIFICATION

Suppose that a finite string of time-series data $\{y_0,\, y_1,\, \ldots,\, y_{2\tau+N-1}\}$ is given. We apply the proposed stochastic realization algorithm to the data and derive a stochastic subspace identification method. Let us construct the block-Hankel matrix:

$$\begin{bmatrix} Y_p \\ Y_f \end{bmatrix} = \begin{bmatrix} y_0 & \cdots & y_{N-1} \\ \vdots & & \vdots \\ y_{\tau-1} & \cdots & y_{\tau+N-2} \\ \hline y_\tau & \cdots & y_{\tau+N-1} \\ \vdots & & \vdots \\ y_{2\tau-1} & \cdots & y_{2\tau+N-2} \end{bmatrix} \in \mathbb{R}^{2\tau p \times N} \tag{31}$$

and define $Y_a := \begin{bmatrix} Y_p^\top & Y_f^\top \end{bmatrix}^\top$. We suppose that an estimate $\tilde{\Theta}_{2\tau} := \frac{1}{N} Y_a Y_a^\top$ of $\Theta_{2\tau}$ satisfies $\tilde{\Theta}_{2\tau} > 0$. We moreover calculate $\tilde{\Theta}_{\tau(p)} = \frac{1}{N} Y_p Y_p^\top$ and $\tilde{\Theta}_{\tau(f)} = \frac{1}{N} Y_f Y_f^\top$ for estimating Θ_τ. We then compute the LQ decomposition

$$\frac{1}{\sqrt{N}} \begin{bmatrix} Y_p \\ Y_f \end{bmatrix} = \begin{bmatrix} \tilde{L}_{pp} & 0 \\ \tilde{L}_{fp} & \tilde{L}_{ff} \end{bmatrix} \begin{bmatrix} \tilde{Q}_p^\top \\ \tilde{Q}_f^\top \end{bmatrix}, \tag{32}$$

where $\tilde{L}_{pp}$, $\tilde{L}_{fp}$, and $\tilde{L}_{ff}$ are estimates of L_{pp}, L_{fp}, and L_{ff} respectively from (19). According to (20), we compute

$$\tilde{\Theta}_{\tau(f)}^{-\frac{1}{2}} \tilde{L}_{fp} \tilde{L}_{pp}^\top \tilde{\Theta}_{\tau(p)}^{-\frac{1}{2}} = \begin{bmatrix} \tilde{U}_1 & \tilde{U}_2 \end{bmatrix} \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \tilde{V}_1^\top \\ \tilde{V}_2^\top \end{bmatrix} \approx \tilde{U}_1 \tilde{\Sigma}_1 \tilde{V}_1^\top \quad (\tilde{\Sigma}_1 \in \mathbb{R}^{\tilde{n}\times\tilde{n}}), \tag{33}$$

where the size of $\tilde{\Sigma}_1$ may not be $n \times n$, different from (20). Based on (21), we estimate the observability matrix $\mathcal{O}_\tau$:

$$\tilde{\mathcal{O}}_\tau = \tilde{\Theta}_{\tau(f)}^{\frac{1}{2}} \tilde{U}_1 \tilde{\Sigma}_1^{\frac{1}{2}} \in \mathbb{R}^{\tau p \times \tilde{n}}. \tag{34}$$

Being aware that $\bar{\mathcal{T}}_\tau$ has the structure

$$\bar{\mathcal{T}}_\tau = \left[\begin{array}{c|c|c|c} H & 0 & \cdots & 0 \\ & H & \ddots & \vdots \\ \mathcal{O}_{\tau-1}\Gamma & \mathcal{O}_{\tau-2}\Gamma & \ddots & 0 \\ & & \cdots & H \end{array}\right], \tag{35}$$

we have an estimate $\tilde{\mathcal{T}}_\tau$ for the block-Toeplitz matrix $\bar{\mathcal{T}}_\tau$.

We minimize γ under the constraints of LMIs

$$\gamma\tilde{\Theta}_{\tau(f)} \geq \tilde{\Theta}_{\tau(f)} - \left(\tilde{\mathcal{O}}_\tau X \tilde{\mathcal{O}}_\tau^\top + \tilde{\mathcal{T}}_\tau \mathcal{D}_\tau(\Pi)\tilde{\mathcal{T}}_\tau^\top\right) \geq 0, \tag{36a}$$

$$X \geq 0, \quad \Pi \geq 0, \tag{36b}$$

where the variables of the LMI constraints are $\gamma \in \mathbb{R}$, $X \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$ and $\Pi \in \mathbb{R}^{(\tilde{n}+p)\times(\tilde{n}+p)}$. Partition Π as follows

$$\Pi = \begin{bmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^\top & \tilde{R} \end{bmatrix}, \quad (\tilde{Q} \in \mathbb{R}^{\tilde{n}\times\tilde{n}},\ \tilde{R} \in \mathbb{R}^{p\times p}). \tag{37}$$

We estimate A and C from the following calculation:

$$\tilde{A} = (\tilde{\mathcal{O}}_\tau^\downarrow)^\dagger \tilde{\mathcal{O}}_\tau(1+p:\tau p,:), \tag{38a}$$

$$\tilde{C} = \tilde{\mathcal{O}}_\tau(1:p,:), \tag{38b}$$

where $(\cdot)^\dagger$ is the pseudo-inverse and $\tilde{\mathcal{O}}_\tau^\downarrow := \tilde{\mathcal{O}}_\tau(1:(\tau-1)p,:)$. We thus have $(\tilde{A},\, \tilde{C},\, \tilde{Q},\, \tilde{S},\, \tilde{R})$ for an estimate of $(A,\, C,\, Q,\, S,\, R)$. According to (22) and (23), we obtain estimates $\tilde{K}$ and $\tilde{\Omega}$ respectively for $\hat{K}$ and $\hat{\Omega}$. If $\tilde{A}$ is stable,

$$\tilde{W}(z) := (\tilde{C}(zI_{\tilde{n}} - \tilde{A})^{-1}\tilde{K} + I_p)\tilde{\Omega}^{\frac{1}{2}}, \tag{39a}$$

$$\tilde{\Phi}(z) = \tilde{W}(z)\tilde{W}^\top(z^{-1}) \tag{39b}$$

are respectively estimates of $W(z)$ and $\Phi(z)$ based on (9) and (8), and $\tilde{W}(z)$ is of minimum phase.

The spectral factor $\tilde{W}(z)$ may not be stable, or $\tilde{A}$ may be unstable. We then consider spectral factororization of $\tilde{\Phi}^{-1}(z^{-1})$ for $|z| = 1$ (Tanaka and Katayama, 2005):

$$\tilde{\Phi}^{-1}(z^{-1}) = \acute{W}^{-\top}(z)\acute{W}^{-1}(z^{-1}). \tag{40}$$

Defining $\tilde{F} = \tilde{A} - \tilde{K}\tilde{C}$, we have

$$\tilde{W}^{-\top}(z) = \left(-\tilde{K}^\top(zI_{\tilde{n}} - \tilde{F}^\top)^{-1}\tilde{C}^\top + I_p\right)\tilde{\Omega}^{-\frac{\top}{2}}.$$

We thus determine the following matrices

$$\begin{bmatrix} \acute{Q} & \acute{S} \\ \acute{S}^\top & \acute{R} \end{bmatrix} = \begin{bmatrix} \tilde{C}^\top \\ I_p \end{bmatrix} \tilde{\Omega}^{-1} \begin{bmatrix} \tilde{C}^\top \\ I_p \end{bmatrix}^\top.$$

We find the stailizing solution $\acute{Z}$ to the Riccati equation

$$Z = \tilde{F}^\top Z \tilde{F} - (-\tilde{F}^\top Z\tilde{K} + \acute{S})(\tilde{K}^\top Z\tilde{K} + \acute{R})^{-1}(-\tilde{F}^\top Z\tilde{K} + \acute{S})^\top + \acute{Q}$$

and then define

$$\acute{C} := \left((-\tilde{F}^\top \acute{Z}\tilde{K} + \acute{S})(\tilde{K}^\top \acute{Z}\tilde{K} + \acute{R})^{-1}\right)^\top,$$

$$\acute{\Omega} := (\tilde{K}^\top \acute{Z}\tilde{K} + \acute{R})^{-1},$$

and $\acute{A} := \tilde{F} + \tilde{K}\acute{C}$. We then have an estimate of $W(z)$:

$$\acute{W}(z) = \left(\acute{C}(zI_{\tilde{n}} - \acute{A})^{-1}\tilde{K} + I_p \right) \acute{\varOmega}^{\frac{1}{2}}.$$

Since $\acute{A}$ is stable, we have a stable $\acute{W}(z)$ satisfying (40). We summarize a subspace identification algorithm.

Stochastic subspace identification algorithm:

Step 1: Construct the block-Hankel matrices Y_p and Y_f as (31). Compute the LQ decomposition (32).

Step 2: Calculate the SVD (33) and $\tilde{\mathcal{O}}_\tau$ in (34). Based on (35), construct $\tilde{\mathcal{T}}_\tau$ using $\tilde{\mathcal{O}}_\tau$.

Step 3: Obtain $\varPi$ in (36) by minimizing γ.

Step 4: Obtain $(\tilde{Q},\ \tilde{S},\ \tilde{R})$ in (37) and $\tilde{A}$ and $\tilde{C}$ in (38). According to (23), obtain estimates $\tilde{K}$ and $\tilde{\varOmega}$ respectively for $\hat{K}$ and $\hat{\varOmega}$. If $\tilde{W}(z)$ in (39a) is stable, it is an estimate of $W(z)$.

Step 5: If $\tilde{A}$ is unstable, compute the spectral factor in (40) and obtain stable $\acute{W}(z)$ for an estimate of $W(z)$.

One of major differences between the conventional algorithms (e.g. Mari, 2000; Mari et al., 2000; Goethals et al., 2003) for modifying the rational transfer matrices of $\tilde{\varPsi}(z)$ and the proposed one appears in Step 3 of the proposed algorithm. They use the LMI in (15) for guaranteeing positive realness, whereas we use the LMIs in (36). In other words, they seek $(\tilde{Q},\ \tilde{R},\ \tilde{S})$ from the system $\tilde{\varPsi}(z)$ under the constraint that the modified $\tilde{\varPsi}(z)$ is positive real. On the other hand, we compute $(\tilde{Q},\ \tilde{R},\ \tilde{S})$ from the finite number of covariance matrices $\tilde{\varLambda}_k$ and the observability matrix $\tilde{\mathcal{O}}_\tau$ under the constraints $X \geq 0$ and $\varPi \geq 0$. We hence avoid explicit use of the LMI (14).

The proposed algorithm does not suffer from the approximation error caused by $F^\tau \not\approx 0$, because we do not have to make the approximation $F^\tau \approx 0$ in estimating $\tilde{\varTheta}_{2\tau}$ and $\tilde{\mathcal{O}}_\tau$ used in Step 3. We can therefore use the proposed algorithm for identifying the stochastic system equipped with zeros close to the unit circle.

5. NUMERICAL SIMULATION RESULTS

In this section, we show numerical simulation results. Suppose that $\hat{W}(z)$ is given by $(n = 6)$:

$$\hat{W}(z) = \frac{1}{D(z)} \begin{bmatrix} N_{11}(z) & N_{12}(z) \\ N_{21}(z) & N_{22}(z) \end{bmatrix},$$

where the numerator $N_{ij}(z)$ (i, j=1, 2) and the denominator $D(z)$ are as follows: $N_{11}(z) = z^6 - 2.8900z^5 + 3.9111z^4 - 3.1190z^3 + 1.2580z^2 - 0.2267z + 0.0136$, $N_{12}(z) = 0.1000z^4 - 0.0619z^3 + 0.0083z^2 + 0.0003z - 0.0001$, $N_{21}(z) = -2.7028z^5 + 7.2763z^4 - 9.3407z^3 + 6.9248z^2 - 2.3914z + 0.2940$, $N_{22}(z) = z^6 - 2.6996z^5 + 3.1487z^4 - 2.1931z^3 + 0.7857z^2 - 0.1212z + 0.0061$, $D(z) = z^6 - 3.6000z^5 + 5.3799z^4 - 4.2719z^3 + 1.9008z^2 - 0.4494z + 0.0441$. It should be noted that $\hat{W}(z)$ is stable and of minimum phase and that the eigen-values of the matrix F in (10) are $\{0.0961,\ 0.2034,\ 0.2999,\ 0.4001,\ 0.4950 \pm 0.8574j\}$ and the absolute value of $0.4950 + 0.8574j$ is 0.9901.

We estimate the system, using the algorithm of Van Overschee and De Moor (1996) ($\tau = 10$, $\tilde{n} = 5$) and simulating the system for different 30 realizations and supposing that $\nu = 2\tau + N - 1 = 5{,}000$ of data are given. Fig. 1 shows estimates of $\hat{W}_{ij}(z) := N_{ij}/D(z)$ (i, $j = 1, 2$), indicating that estimates of $\hat{W}_{11}(z)$ and $\hat{W}_{21}(z)$ do not capture the valleys at the frequency around 1.

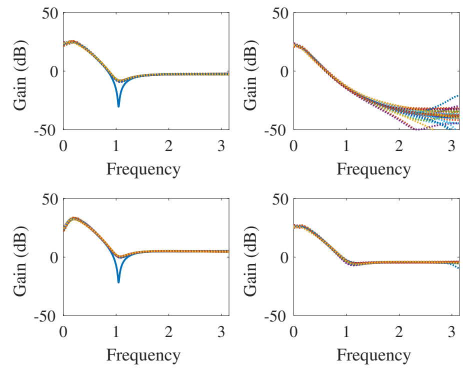

Fig. 1. Gain plots of systems estimated by the algorithm of Van Overschee and De Moor (1996) for $\tau = 10$ and $\tilde{n} = 5$, where the solid blue lines are true. The plots at the top left, top right, bottom left and bottom right respectively show estimates of $\hat{W}_{11}(z)$, $\hat{W}_{12}(z)$, $\hat{W}_{21}(z)$ and $\hat{W}_{22}(z)$.

We estimate the system by using the algorithm of Mari (2000). For unstable $\tilde{A} \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$, we compute

$$\begin{bmatrix} T^{-1} & 0 \\ 0 & I_p \end{bmatrix} \left[\begin{array}{c:c} \tilde{A} & \tilde{G} \\ \hdashline \tilde{C} & 0 \end{array}\right] \begin{bmatrix} T & 0 \\ 0 & I_p \end{bmatrix} = \left[\begin{array}{cc:c} \tilde{A}_u & 0 & \tilde{G}_u \\ 0 & \tilde{A}_s & \tilde{G}_s \\ \hdashline \tilde{C}_u & \tilde{C}_s & 0 \end{array}\right],$$

where $T \in \mathbb{R}^{\tilde{n}\times\tilde{n}}$, and $\tilde{A}_u$ and $\tilde{A}_s$ are respectively anti-stable and stable. We then have

$$\tilde{\varPsi}(z) = \tilde{\varPsi}_u(z) + \tilde{\varPsi}_s(z) + \frac{1}{2}\tilde{\varLambda}_0,$$
$$\tilde{\varPsi}_u(z) := \tilde{C}_u(zI - \tilde{A}_u)^{-1}\tilde{G}_u,$$
$$\tilde{\varPsi}_s(z) := \tilde{C}_s(zI - \tilde{A}_s)^{-1}\tilde{G}_s.$$

Mari (2000) proposed to use

$$\breve{\varPsi}(z) = \tilde{\varPsi}_u^\top(z^{-1}) + \tilde{\varPsi}_z(z) + \frac{1}{2}\tilde{\varLambda}_0,$$

instead of $\tilde{\varPsi}(z)$ for unstable $\tilde{A}$, since the ultimate goal of stochastic system identification is producing a spectral factor for $\varPhi(z) = \varPsi(z) + \varPsi^\top(z^{-1})$. We have indeed $\tilde{\varPsi}(z) + \tilde{\varPsi}^\top(z^{-1}) = \breve{\varPsi}(z) + \breve{\varPsi}^\top(z^{-1})$. We therefore use

$$\tilde{\varPsi}_u^\top(z^{-1}) \approx -\tilde{G}_u^\top(zI - \tilde{A}_u^{-\top})^{-1}\tilde{A}_u^{-\top}\tilde{C}_u^\top - \frac{1}{2}\left((\tilde{G}_u^\top \tilde{A}_u^{-\top}\tilde{C}_u^\top) + (\tilde{G}_u^\top \tilde{A}_u^{-\top}\tilde{C}_u^\top)^\top \right).$$

We modify both $\tilde{G}$ and $\tilde{\varLambda}_0$ for not satisfying the positive real constraint (15). We show the gain plots of estimates in Fig. 2. Comparing Figs. 1 and 2, we observe that the algorithm of Mari (2000) captures the valleys around the frequency 1, though the gain plots are scattered around the true system.

We finally show estimates given by the proposed algorithm ($\tau = 10$, $\tilde{n} = 5$) in Fig. 3. We see that it captures the valleys and that the gain plots are less scattered.

6. CONCLUSIONS

In this paper, we have developed a subspace identification algorithm for stochastic systems equipped with zeros close

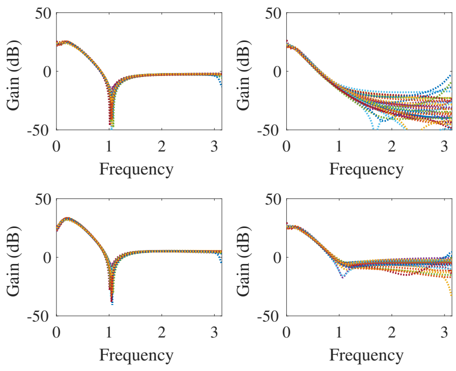

Fig. 2. Gain plots of systems estimated by the algorithm by Mari (2000) for $\tau = 10$ and $\tilde{n} = 5$.

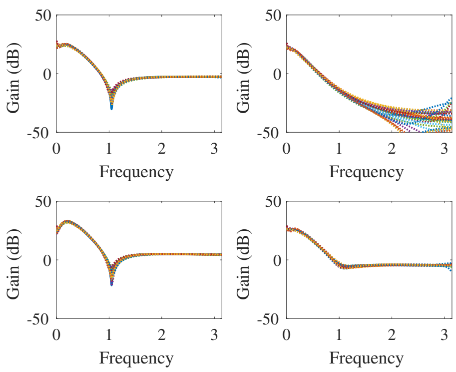

Fig. 3. Gain plots of the systems estimated by the proposed algorithm ($\tau = 10$, $\tilde{n} = 5$).

to the unit circle. By reviewing the methods for ensuring positive realness in the conventional algorithms, we developed a new stochastic realization algorithm. Conventional algorithms have ensured positive realness of the causal part of the spectral density function by modifying rational transfer matrices. On the other hand, we have derived an equation for the observability matrix and the covariance matrices of the output, state, and white noise, and we guaranteed positive realness by means of positive semi-definiteness of the covariance of the white noise. We derived a stochastic subspace identification algorithm by applying the realization algorithm to a finite string of data. Numerical simulation results indicated that the present algorithm successfully estimates the system. It is a future problem to make consistency analysis and to study uniqueness of the solution.

REFERENCES

Akaike, H. (1975). Markovian representation of stochastic processes by canonical variables. *SIAM Journal on Control*, 13(1), 162–173.

Akçay, H. and Türkay, S. (2015). Positive realness in stochastic subspace identification: A regularized and reweighted nuclear norm minimization approach. *2015 European Control Conference (ECC)*, 1760–1765.

Bauer, D. (2005). Asymptotic properties of subspace estimators. *Automatica*, 41(3), 359–376.

Bauer, D. and Ljung, L. (2002). Some facts about the choice of the weighting matrices in Larimore type of subspace algorithms. *Automatica*, 38(5), 763–773.

Chiuso, A. and Picci, G. (2005). Consistency analysis of some closed-loop subspace identification methods. *Automatica*, 41(3), 377–391.

Dahlén, A., Lindquist, A., and Mari, J. (1998). Experimental evidence showing that stochastic subspace identification methods may fail. *Systems & Control Letters*, 34(5), 303–312.

Desai, U.B. and Pal, D. (1984). A transformation approach to stochastic model reduction. *IEEE Transactions on Automatic Control*, 29(12), 1097–1100.

Faurre, P.L. (1976). Stochastic realization algorithms. In R.K. Mehra and D.G. Lainiotis (eds.), *System Identification: Advances and Case Studies*, 1–25. Academic Press.

Goethals, I., Van Gestel, T., Suykens, J., Van Dooren, P., and De Moor, B. (2003). Identification of positive real models in subspace identification by using regularization. *IEEE Transactions on Automatic Control*, 48(10), 1843–1847.

Ho, B.L. and Kalman, R.E. (1966). Effective construction of linear state-variable models from input/output functions. *Regelungstechnik*, 14(1-12), 545–548.

Ikeda, K. (2015). Consistent estimate of Kalman gain in subspace identification method. *IEEE Conference on Control Applications*, 151–156.

Ikeda, K. and Tanaka, H. (2017). An SDP formulation for consistent estimate of innovations model. *Proc. of the 2017 Asian Control Conference*, 1772–1777.

Jansson, M. (2003). Subspace identification and ARX modeling. *IFAC Proceedings Volumes*, 36(16), 1585–1590.

Katayama, T. (2005). *Subspace Methods for System Identification*. Springer-Verlag London.

Lindquist, A. and Picci, G. (1996). Canonical correlation analysis, approximate covariance extension, and identification of stationary time series. *Automatica*, 32(5), 709–733.

Mari, J. (2000). Modifications of rational transfer matrices to achieve positive realness. *Signal Processing*, 80(4), 615–635.

Mari, J., Stoica, P., and McKelvey, T. (2000). Vector ARMA estimation: A reliable subspace approach. *IEEE Transactions on Signal Processing*, 48(7), 2092–2104.

Tanaka, H. and Ikeda, K. (2018). Closed-loop subspace identification taking initial state into account. *IFAC-PapersOnLine*, 51(15), 604–609.

Tanaka, H. and Katayama, T. (2005). Stochastic subspace identification guaranteeing stability and minimum phase. *Preprints of the 16th IFAC World Congress (IFAC 2005)*.

Vaccaro, R.J. and Vukina, T. (1993). A solution to the positivity problem in the state-space approach to modeling vector-valued time series. *Journal of Economic Dynamics and Control*, 17(3), 401–421.

Van Overschee, P. and De Moor, B. (1993). Subspace algorithms for the stochastic identification problem. *Automatica*, 29(3), 649–660.

Van Overschee, P. and De Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1), 75–93.

Van Overschee, P. and De Moor, B. (1996). *Subspace Identification for Linear Systems — Theory – Implementation – Applications —*. Kluwer Academic Publishers.