Available online at www.sciencedirect.com
SCIENCE DIRECT®
Automatica 40 (2004) 1333–1347
automatica
www.elsevier.com/locate/automatica

Frequency domain subspace-based identification of discrete-time power spectra from nonuniformly spaced measurements☆

Hüseyin Akçay*, Semiha Türkay

Department of Electrical and Electronics Engineering, Anadolu University, Eskişehir 26470, Turkey

Received 1 June 2003; received in revised form 29 February 2004; accepted 18 March 2004

Abstract

In this paper, we present a new subspace-based algorithm for the identification of multi-input/multi-output, square, discrete-time, linear-time invariant systems from nonuniformly spaced power spectrum measurements. The algorithm is strongly consistent and it is illustrated with one practical example that solves a stochastic road modeling problem.

Keywords: System identification; Power spectrum; Subspace methods; State-space models; Nonuniformly spaced measurements; Shape filter

1. Introduction

Identification of multi-input/multi-output systems from a measured power spectrum is a problem arising in certain applications; for example, the design of linear shaping filters for noise processes. A practical application is the modeling of stochastic road disturbances experienced by a vehicle moving forward. The goal here is to model road spectrum by a rational transfer function of reasonably low order and to use this approximation for a design of a *linear shaping filter* with a white noise input. Once such an approximation is made, the vehicle control problem can be formulated in standard form. The algorithm of this paper determines a state-space realization of road spectrum. Applications to the modeling of acoustic power spectra and the modeling of passenger sensitivity for car accelerations are presented in Van Overschee, De Moor, Dehandschutter, and Swevers (1997).

In this paper, we study the problem of fitting a linear discrete-time power spectrum to given measured power spectrum samples. A *parametric* or *model-based* approach to this problem uses a nonlinear least-squares criterion, which is optimized by an iterative nonlinear search in the parameter space. Discussion of parametric as well as *nonparametric* methods, which mostly use time-domain data, can be found in Kay (1988), Priestley (1989) and Stoica and Moses (1997). Drawbacks of this approach are convergence problems and difficulty of parameterizing multi-input/multi-output systems. There has been an extensive amount of research to determine the *so-called* canonical models (Glover & Willems, 1974; Guidorzi, 1974, 1981; Van Overbeek & Ljung, 1982).

The subspace approach, on the other hand, does not suffer from any of these inconveniences. In subspace identification algorithms, there is no explicit need for parameterization since full state-space models are used and the only parameter is the order of the system. The major advantage of subspace identification algorithms over the classical prediction error methods (Ljung, 2000) is the absence of nonlinear parametric optimization problems. Subspace identification algorithms are noniterative and therefore do not suffer from convergence problems. They always produce results, which are often good for practical data.

Given time domain measurements, there are many state-space subspace identification algorithms available (Larimore, 1990; Verhaegen, 1994; Viberg, 1995; Van Overschee & De Moor, 1996a). Frequency domain subspace identification algorithms have already appeared in the literature (Liu, Jacques, & Miller, 1994; McKelvey, Akçay, & Ljung, 1996a, b; Van Overschee & De Moor, 1996b). They can be described as direct frequency domain formulations of time-domain subspace algorithms. If the excitation of the system is well-designed, each

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of editor Torsten Söderström.

* Corresponding author. Tel.: +90-222-335-0580; fax: +90-222-323-9501.

E-mail address: huakcay@anadolu.edu.tr (H. Akçay).

0005-1098/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2004.03.010

measurement in the frequency domain compiled from a large number of time-domain measurement is of high quality. Moreover, data originating from different experiments can easily be combined in the frequency domain (Schoukens & Pintelon, 1991). However, these algorithms are not directly applicable for the identification of frequency domain power spectra since rational spectrum models are constrained to have *positive real* transfer functions.

In Van Overschee et al. (1997), a subspace algorithm which uses spectrum samples obtained at uniformly spaced frequencies was presented. The algorithm in Van Overschee et al. (1997) is based on McKelvey et al. (1996a); and it uses biased impulse response coefficients expressed as functions of the system matrices. However, this algorithm generates *strongly consistent* power spectrum estimates. A related work is Verhaegen (1996). In this work, a subspace algorithm for the time domain identification of mixed causal and anti-causal systems was proposed. The frequency domain extension of this algorithm was given in Fraanje, Verhaegen, Verdult, and Pintelon (2003).

The objective of this paper is to remove the restriction on the frequencies. This problem is precisely formulated in Section 2. In Section 3, we present our subspace-based algorithm to identify multi-input/multi-output systems from power spectrum samples measured at nonuniformly spaced frequencies and show that this algorithm is not only strongly consistent but also recovers finite-dimensional rational spectra given a finite number of noise-free data (depending on the model order). The proposed algorithm is based on the results in Van Overschee et al. (1997), Verhaegen (1996) and Fraanje et al. (2003). The proofs are given in the Appendices. In Section 4, the properties of the new algorithm are studied by means of two examples. In the first example, we simulate a system that has a power spectrum with sharp peaks. In the second example, we illustrate the practical relevance of the problem treated in this paper by solving a stochastic road modeling problem.

2. Problem formulation

Consider a multi-input/multi-output square linear-time invariant discrete-time system represented by the state-space equations:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

where $x(t) \in \mathbf{R}^n$ is the state, $u(t) \in \mathbf{R}^m$ and $y(t) \in \mathbf{R}^m$ are, respectively, the input and the output of the system. The transfer function of system (1) denoted by $G(z)$ is computed as

$$G(z) = D + C(zI_n - A)^{-1}B, \tag{2}$$

where I_n is the $n \times n$ identity matrix. We summarize the requirements on $G(z)$ in the following:

Assumption 2.1. System (1) is stable and strictly minimum phase: all eigenvalues of A and $A - BD^{-1}C$ lie strictly inside the unit circle. The pairs $\{A, B\}$ and $\{A, C\}$ are controllable and observable, respectively. All eigenvalues of A are nonzero and distinct.

Thus, system (1) is a minimal stochastic system. Note that since the *Jordan canonical form* is not numerically stable, a slight perturbation of A will lead to distinct eigenvalues if A has repeated eigenvalues.

Assuming that $u(t)$ is zero mean unity variance white noise process, the power spectrum associated with (1) denoted by $S(z)$ is defined as

$$S(z) \triangleq G(z)G^{\mathrm{T}}(z^{-1}). \tag{3}$$

System (1) is called the *innovation form*, unity variance, minimum phase spectral factor associated with power spectrum $S(z)$. From (3) and Assumption 2.1, note that

$$S(\mathrm{e}^{\mathrm{j}\theta}) > 0 \quad \text{for all } \theta. \tag{4}$$

This is the positive realness condition, and it imposes a constraint on the given spectrum samples S_k, i.e. $S_k > 0$ for each k, as well as on the identified power spectrum denoted by $\widehat{S}_N(z)$.

We will assume that the noise η corrupting the spectrum samples is a zero mean complex white noise process with a covariance function satisfying

$$\mathbf{E}\begin{bmatrix} \operatorname{Re}\eta_k \\ \operatorname{Im}\eta_k \end{bmatrix} [\operatorname{Re}\eta_s^{\mathrm{T}} \ \ \operatorname{Im}\eta_s^{\mathrm{T}}] = \begin{bmatrix} \frac{1}{2}\mathscr{R}_k & 0 \\ 0 & \frac{1}{2}\mathscr{R}_k \end{bmatrix} \delta_{ks}. \tag{5}$$

Here $\mathbf{E}(x)$ denotes the expected value of random variable x; A^{T}, A^*, and A^{H} are, respectively, the transpose, the complex conjugate, and the complex conjugate transpose of A; $\operatorname{Re} x$ and $\operatorname{Im} x$ are the real and the imaginary parts of x; and δ_{ks} is the Kronecker delta. Furthermore, we assume that the fourth-order moments are bounded above by some $M_\eta < \infty$ as

$$\mathbf{E}\,\|\eta_k\|_F^4 \leqslant M_\eta \quad \text{for all } k, \tag{6}$$

where $\|\eta_k\|_F \triangleq [\operatorname{Tr}(\eta_k^{\mathrm{H}}\eta_k)]^{1/2}$ denotes the Frobenius norm of η_k.

Given a set X, we denote the number of elements in X by $\mathscr{C}(X)$. Thus, $\mathscr{C}(\{\theta_k\}_{k=1}^N \cap [a,b])$ is the number of frequencies contained in $[a,b] \subseteq [0, 2\pi]$. We will assume that the frequencies satisfy

$$\lim_{N\to\infty} \inf \frac{1}{N} \mathscr{C}(\{\theta_k\}_{k=1}^N \cap [a,b]) \geqslant \delta(b-a) \tag{7}$$

for every $[a,b] \subseteq [0, 2\pi]$ and some fixed $\delta > 0$. This means that every point on the unit circle has a nonzero asymptotic density of frequencies relative to N.

For a complex measurable function $G(z)$, we define the supremum norm by

$$\|G\|_\infty \triangleq \sup_{0 \leqslant \theta < 2\pi} \sigma_{\max}(G(\mathrm{e}^{\mathrm{j}\theta})),$$

where $\sigma_{\max}$ denotes the largest singular value.

The problem studied in this paper can be stated as follows:

Given: N noisy samples $S_k \in \mathbf{C}^{m\times m}$ of the power spectrum $S(z)$ evaluated at N points on the unit circle:

$$S_k = S(e^{j\theta_k}) + \eta_k, \quad k = 1, 2, \ldots, N. \tag{8}$$

Find: A quadruplet $(\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D})$ such that the estimated power spectrum

$$\widehat{S}_N(z) = \widehat{G}(z)\widehat{G}^{\mathrm{T}}(z^{-1}) \tag{9}$$

is strongly consistent, i.e.

$$\lim_{N\to\infty} \|\widehat{S}_N - S\|_\infty = 0, \quad \text{w.p.1.}, \tag{10}$$

where

$$\widehat{G}(z) \triangleq \widehat{C}(zI_n - \widehat{A})^{-1}\widehat{B} + \widehat{D}. \tag{11}$$

We also require the algorithm to produce the true model if the noise is zero given a finite amount of data N, i.e. there exists an $N_0 < \infty$ such that

$$\|\widehat{S}_N - S\|_\infty = 0 \quad \text{for all } N \geqslant N_0. \tag{12}$$

An identification algorithm which satisfies (12) is called *correct algorithm*. In this paper, we present an algorithm which has these properties. Strong consistency is a most natural requirement for any useful algorithm. As the amount of data increases, asymptotically the correct model should be obtained. In practice, any algorithm has to use a finite amount of data. Then, correctness of an algorithm becomes important. This is particularly important for spectra with sharp peaks.

The above identification problem can be thought as the design of a linear shaping filter (A, B, C, D) from (corrupted) power spectrum measurements. In this procedure, the zeros of $G(z)$ can be restricted, without loss of generality, to be minimum phase.

3. Identification algorithm

Let us first consider the noise-free case to motivate the derivation of the identification algorithm. We begin by splitting $S(z)$ into the *so-called* spectral summands as follows.

Theorem 1. *Consider the power spectrum $S(z)$ in (3). Suppose that Assumption 2.1 holds. Let P be the solution of the discrete-time Lyapunov equation*:

$$P = APA^{\mathrm{T}} + BB^{\mathrm{T}}. \tag{13}$$

Let

$$E \triangleq CPC^{\mathrm{T}} + DD^{\mathrm{T}}, \tag{14}$$

$$F \triangleq APC^{\mathrm{T}} + BD^{\mathrm{T}}. \tag{15}$$

Then $S(z)$ can be split into the sum of two system transfer matrices as follows:

$$S(z) = H(z) + H^{\mathrm{T}}(z^{-1}) \tag{16}$$

with

$$H(z) \triangleq \frac{1}{2}E + C(zI_n - A)^{-1}F. \tag{17}$$

Proof. See, for example, Caines (1988).

This splitting of $S(z)$ into the sum of a causal transfer function $H(z)$ and an anti-causal transfer function $H^{\mathrm{T}}(z^{-1})$ is the first step of our subspace-based identification algorithm. It is also the starting point of the subspace algorithm in Van Overschee et al. (1997). As in Van Overschee et al. (1997), from the samples S_k we identify a quadruplet $(A, F, C, \frac{1}{2}E)$ which describes the spectral summand $H(z)$. The algorithm proposed in Van Overschee et al. (1997) uses biased Markov parameters of $S(z)$ as in McKelvey et al. (1996a); and requires the discrete frequencies θ_k, $k = 1, 2, \ldots, N$ be uniformly spaced in the interval $[0, \pi]$. The contribution of this paper is to remove this restriction on the frequencies.

Next, from (16) and (17) we write a state-representation of $S(z)$ as follows:

$$x^{\mathrm{c}}(t+1) = Ax^{\mathrm{c}}(t) + Fu(t), \tag{18}$$

$$x^{\mathrm{ac}}(t-1) = A^{\mathrm{T}}x^{\mathrm{ac}}(t) + C^{\mathrm{T}}u(t), \tag{19}$$

$$y^{\mathrm{s}}(t) = Cx^{\mathrm{c}}(t) + F^{\mathrm{T}}x^{\mathrm{ac}}(t) + Eu(t). \tag{20}$$

These equations are the special cases of the equations considered in Verhaegen (1996) for the time-domain subspace identification of mixed causal and anti-causal linear-time invariant systems.

Following Fraanje et al. (2003), we take the discrete Fourier transforms of Eqs. (18)–(20), where we shift Eq. (19) by $p-1$ samples forward in time:

$$e^{j\theta}X^{\mathrm{c}}(\theta) = AX^{\mathrm{c}}(\theta) + FU(\theta),$$

$$e^{-j\theta}X^{\mathrm{ac,p}}(\theta) = A^{\mathrm{T}}X^{\mathrm{ac,p}}(\theta) + C^{\mathrm{T}}e^{j(p-1)\theta}U(\theta),$$

$$Y^{\mathrm{s}}(\theta) = CX^{\mathrm{c}}(\theta) + F^{\mathrm{T}}e^{-j(p-1)\theta}X^{\mathrm{ac,p}}(\theta) + EU(\theta),$$

where $X^{\mathrm{c}}(\theta)$, $X^{\mathrm{ac},p}(\theta)$, $U(\theta)$, and $Y^{\mathrm{s}}(\theta)$ denote the discrete Fourier transforms of $x^{\mathrm{c}}(t)$, $x^{\mathrm{ac}}(t+p-1)$, $u(t)$, and $y(t)$, respectively, and $p > 2n$. Let $X_i^{\mathrm{c}}(\theta)$ be the resulting state transform when $U(\theta) = e_i$, the unit vector with 1 on the ith position; and $X_i^{\mathrm{ac},p}(\theta)$ is defined similarly. By defining the compound state matrices:

$$X_{\mathrm{C}}^{\mathrm{c}}(\theta) \triangleq [X_1^{\mathrm{c}}(\theta)\ X_2^{\mathrm{c}}(\theta)\ \cdots\ X_m^{\mathrm{c}}(\theta)],$$

$$X_{\mathrm{C}}^{\mathrm{ac},p}(\theta) \triangleq [X_1^{\mathrm{ac},p}(\theta)\ X_2^{\mathrm{ac},p}(\theta)\ \cdots\ X_m^{\mathrm{ac},p}(\theta)], \tag{21}$$

$S(e^{j\theta})$ can be implicitly described as

$$S(e^{j\theta}) = CX_{\mathrm{C}}^{\mathrm{c}}(\theta) + F^{\mathrm{T}}e^{-j(p-1)\theta}X_{\mathrm{C}}^{\mathrm{ac},p}(\theta) + E$$

with

$$e^{j\theta}X_{\mathrm{C}}^{\mathrm{c}}(\theta) = AX_{\mathrm{C}}^{\mathrm{c}}(\theta) + F,$$

$$e^{-j\theta}X_{\mathrm{C}}^{\mathrm{ac},p}(\theta) = A^{\mathrm{T}}X_{\mathrm{C}}^{\mathrm{ac},p}(\theta) + C^{\mathrm{T}}e^{j(p-1)\theta}. \tag{22}$$

By iteratively substituting the state-equations, we obtain the relation

$$\begin{bmatrix} S(e^{j\theta}) \\ e^{j\theta}S(e^{j\theta}) \\ \vdots \\ e^{j(p-2)\theta}S(e^{j\theta}) \\ e^{j(p-1)\theta}S(e^{j\theta}) \end{bmatrix} = \Gamma_p \begin{bmatrix} I_m \\ e^{j\theta}I_m \\ \vdots \\ e^{j(p-2)\theta}I_m \\ e^{j(p-1)\theta}I_m \end{bmatrix} + \mathcal{O}_p \begin{bmatrix} X_C^c(\theta) \\ X_C^{ac,p}(\theta) \end{bmatrix}, \tag{23}$$

where

$$\mathcal{O}_p \triangleq \begin{bmatrix} C & F^T(A^T)^{p-1} \\ CA & F^T(A^T)^{p-2} \\ \vdots & \vdots \\ CA^{p-2} & F^TA^T \\ CA^{p-1} & F^T \end{bmatrix} \tag{24}$$

and

$$\Gamma_p \triangleq \begin{bmatrix} E & F^TC^T & \cdots & F^T(A^T)^{p-2}C^T \\ CF & E & \cdots & \vdots \\ \vdots & \vdots & \ddots & F^TC^T \\ CA^{p-2}F & \cdots & CF & E \end{bmatrix}. \tag{25}$$

By repeating (23) for θ_k, $k = 1, 2, \ldots, N$, we get

$$\mathcal{S}_C = \mathcal{O}_p\mathcal{X}_C + \Gamma_p\mathcal{W}_C \tag{26}$$

where $z_k = e^{j\theta_k}$, $k = 1, 2, \ldots, N$ and

$$\mathcal{S}_C \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} S(z_1) & \cdots & S(z_N) \\ e^{j\theta_1}S(z_1) & \cdots & e^{j\theta_N}S(z_N) \\ \vdots & \ddots & \vdots \\ e^{j(p-1)\theta_1}S(z_1) & \cdots & e^{j(p-1)\theta_N}S(z_N) \end{bmatrix}, \tag{27}$$

$$\mathcal{W}_C \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} I_m & \cdots & I_m \\ e^{j\theta_1}I_m & \cdots & e^{j\theta_N}I_m \\ \vdots & \ddots & \vdots \\ e^{j(p-1)\theta_1}I_m & \cdots & e^{j(p-1)\theta_N}I_m \end{bmatrix}, \tag{28}$$

$$\mathcal{X}_C \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} X_C^c(\theta_1) & \cdots & X_C^c(\theta_N) \\ X_C^{ac,p}(\theta_1) & \cdots & X_C^{ac,p}(\theta_N) \end{bmatrix}. \tag{29}$$

Now, we consider the noisy data case. From (8), (26), and (27), we get

$$\widehat{\mathcal{S}}_C = \mathcal{O}_p\mathcal{X}_C + \Gamma_p\mathcal{W}_C + \mathcal{N}_C, \tag{30}$$

where

$$\widehat{\mathcal{S}}_C \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} S_1 & \cdots & S_N \\ e^{j\theta_1}S_1 & \cdots & e^{j\theta_N}S_N \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(p-1)\theta_1}S_1 & \cdots & e^{j(p-1)\theta_N}S_N \end{bmatrix}, \tag{31}$$

$$\mathcal{N}_C \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} \eta_1 & \cdots & \eta_N \\ e^{j\theta_1}\eta_1 & \cdots & e^{j\theta_N}\eta_N \\ \vdots & \ddots & \vdots \\ e^{j(p-1)\theta_1}\eta_1 & \cdots & e^{j(p-1)\theta_N}\eta_N \end{bmatrix}. \tag{32}$$

Since $\mathcal{O}_p$ is a real matrix and we are interested in the real range space, we convert (26) into a relation involving only real valued matrices:

$$\begin{aligned} \widehat{\mathcal{S}} &= \mathcal{O}_p\mathcal{X} + \Gamma_p\mathcal{W} + \mathcal{N} \\ &= \mathcal{S} + \mathcal{N}, \end{aligned} \tag{33}$$

where

$$\widehat{\mathcal{S}} \triangleq [\mathrm{Re}\,\widehat{\mathcal{S}}_C \;\; \mathrm{Im}\,\widehat{\mathcal{S}}_C], \tag{34}$$

$$\mathcal{S} \triangleq [\mathrm{Re}\,\mathcal{S}_C \;\; \mathrm{Im}\,\mathcal{S}_C], \tag{35}$$

$$\mathcal{X} \triangleq [\mathrm{Re}\,\mathcal{X}_C \;\; \mathrm{Im}\,\mathcal{X}_C], \tag{36}$$

$$\mathcal{W} \triangleq [\mathrm{Re}\,\mathcal{W}_C \;\; \mathrm{Im}\,\mathcal{W}_C], \tag{37}$$

$$\mathcal{N} \triangleq [\mathrm{Re}\,\mathcal{N}_C \;\; \mathrm{Im}\,\mathcal{N}_C]. \tag{38}$$

Let $\mathcal{W}^\perp$ be the projection matrix onto the null space of $\mathcal{W}$ given by

$$\mathcal{W}^\perp \triangleq I_{2mN} - \mathcal{W}^H(\mathcal{W}\mathcal{W}^H)^{-1}\mathcal{W}. \tag{39}$$

The term $\Gamma_p\mathcal{W}$ in (33) is canceled when multiplied from right by $\mathcal{W}^\perp$. Thus,

$$\begin{aligned} \widehat{\mathcal{S}}\mathcal{W}^\perp &= \mathcal{O}_p\mathcal{X}\mathcal{W}^\perp + \mathcal{N}\mathcal{W}^\perp \\ &= \mathcal{S}\mathcal{W}^\perp + \mathcal{N}\mathcal{W}^\perp. \end{aligned} \tag{40}$$

The range space of $\mathcal{S}\mathcal{W}^\perp$ equals the range space of $\mathcal{O}_p$ unless rank cancelations occur. A sufficient condition for the range spaces to be equal is that the intersection between the row spaces of $\mathcal{W}$ and $\mathcal{X}$ is empty. In the following, we present sufficient conditions in terms of the data and the system.

Lemma 2. *Let $N \geqslant (p/2) + n + 1$, $\mathcal{W}_C$, and $\mathcal{X}_C$ be given by (28) and (29) with distinct frequencies θ_k such that z_k*

is not an eigenvalue of A. Then

$$\operatorname{rank}\begin{bmatrix}\mathscr{W}\\ \mathscr{X}\end{bmatrix} = pm+2n \;\Leftrightarrow\; (A,B,C,D)\ \textit{minimal}. \tag{41}$$

Proof. See Appendix A. □

If the frequencies are distinct, the number of data satisfies $N \geqslant (p/2)+n+1$, and (A,B,C,D) is minimal, then the two row spaces of $\mathscr{W}$ and $\mathscr{X}$ do not intersect and the range space of $\mathscr{S}\mathscr{W}^{\perp}$ coincides with the range space of $\mathscr{O}_p$. Now, a study of the relation between the column range spaces of $\mathscr{S}\mathscr{W}^{\perp}$ and $\widehat{\mathscr{S}}\mathscr{W}^{\perp}$ for large N is in order.

In Moor (1993), it was shown that by using the singular value decomposition of $\widehat{\mathscr{S}}\mathscr{W}^{\perp}$, the $2n$ left singular vectors corresponding to the $2n$ largest singular values form a strongly consistent estimate of the range space of $\mathscr{S}\mathscr{W}^{\perp}$ if the following conditions hold w.p.1

$$\text{(i)}\quad \lim_{N\to\infty} \mathscr{S}\mathscr{W}^{\perp}(\mathscr{N}\mathscr{W}^{\perp})^{\mathrm{T}} = 0, \tag{42}$$

$$\text{(ii)}\quad \lim_{N\to\infty} \mathscr{N}\mathscr{W}^{\perp}(\mathscr{N}\mathscr{W}^{\perp})^{\mathrm{T}} = \alpha I_{pm} \tag{43}$$

for some scalar $\alpha \geqslant 0$. In McKelvey et al. (1996a), it was shown under assumption (6) that (42) holds and

$$\lim_{N\to\infty} \mathscr{N}\mathscr{W}^{\perp}(\mathscr{N}\mathscr{W}^{\perp})^{\mathrm{T}} = \mathscr{K}\mathscr{K}^{\mathrm{T}}, \quad \text{w.p.1.}, \tag{44}$$

where $\mathscr{K} \in \mathbf{R}^{pm\times pm}$ is a matrix defined by

$$\mathscr{K}\mathscr{K}^{\mathrm{T}} \triangleq \operatorname{Re}(\mathscr{W}_{\mathrm{C}}\mathscr{R}\mathscr{W}_{\mathrm{C}}^{\mathrm{H}}),$$

$$\mathscr{R} \triangleq \begin{bmatrix}\mathscr{R}_1 & 0 & \cdots & 0\\ 0 & \mathscr{R}_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \mathscr{R}_N\end{bmatrix}. \tag{45}$$

The matrix $\mathscr{K}$ can be found by a Cholesky decomposition. Thus, from (40) we have the weighted version

$$\mathscr{K}^{-1}\widehat{\mathscr{S}}\mathscr{W}^{\perp} = \mathscr{K}^{-1}\mathscr{S}\mathscr{W}^{\perp} + \mathscr{K}^{-1}\mathscr{N}\mathscr{W}^{\perp} \tag{46}$$

satisfying (42) and (43) with $\alpha = 1$. Hence, the $2n$ left singular vectors corresponding to the $2n$ largest singular values of $\mathscr{K}^{-1}\widehat{\mathscr{S}}\mathscr{W}^{\perp}$ will form a strongly consistent estimate of the range space of $\mathscr{K}^{-1}\mathscr{S}\mathscr{W}^{\perp}$ which equals to the range space of $\mathscr{K}^{-1}\mathscr{O}_p$.

A numerically efficient way of forming $\widehat{\mathscr{S}}\mathscr{W}^{\perp}$ is to use the QR-factorization:

$$\begin{bmatrix}\mathscr{W}\\ \widehat{\mathscr{S}}\end{bmatrix} = \begin{bmatrix}R_{11} & 0\\ R_{21} & R_{22}\end{bmatrix}\begin{bmatrix}Q_1^{\mathrm{T}}\\ Q_2^{\mathrm{T}}\end{bmatrix}. \tag{47}$$

A simple derivation yields

$$\widehat{\mathscr{S}}\mathscr{W}^{\perp} = R_{22}Q_2^{\mathrm{T}} \tag{48}$$

and it suffices to use R_{22} since Q_2^{T} is a matrix of full rank. Thus, the $2n$ left singular vectors corresponding to the $2n$ largest singular values of $\mathscr{K}^{-1}\widehat{\mathscr{S}}\mathscr{W}^{\perp}$ are obtained from the singular value decomposition:

$$\mathscr{K}^{-1}R_{22} = [\widehat{U}_{2n}\ \ \tilde{U}]\begin{bmatrix}\widehat{\Sigma}_{2n} & 0\\ 0 & \tilde{\Sigma}\end{bmatrix}\begin{bmatrix}\widehat{V}_{2n}\\ \tilde{V}\end{bmatrix}, \tag{49}$$

where this decomposition is partitioned such that $\widehat{\Sigma}_{2n}$ contains the $2n$ largest singular values.

Our consistency analysis has shown that

$$\lim_{N\to\infty} \mathscr{K}\widehat{U}_{2n} = \mathscr{O}_p T, \quad \text{w.p.1} \tag{50}$$

for some nonsingular matrix T. In the calculation of $\widehat{U}_{2n}$, $2n$ elements with fixed indices can be chosen freely, subject to the constraint that magnitudes are not greater than unity. Thus, by fixing values of those elements for all N, we see from (50) that $\widehat{U}_{2n}$ converges to a matrix denoted by U_{2n} w.p.1 as $N \to \infty$. Hence,

$$\mathscr{K}U_{2n} = \mathscr{O}_p T. \tag{51}$$

This asymptotic formula (in the number of data) will be the key in the development of our algorithm. Before undertaking this study, let us record the following result which will be used later.

Lemma 3. *Let S_k, $k=1,\dots,N$ be noise-free samples of the power spectrum of a discrete-time system of order n satisfying Assumption 2.1 at N distinct frequencies θ_k. Furthermore, let $N \geqslant (p/2)+n+1$ and $\mathscr{K} \in \mathbf{R}^{pm\times pm}$ be any nonsingular matrix. Then, for some nonsingular T*

$$\mathscr{K}\widehat{U}_{2n} = \mathscr{O}_p T. \tag{52}$$

Thus, the equations derived from the asymptotic formula are also valid for a finite number of data under the conditions stated in Lemma 3.

Let J_{u} and J_{d} be the upward and downward shift matrices defined by

$$J_{\mathrm{u}}\mathscr{O}_p \triangleq \begin{bmatrix} CA & F^{\mathrm{T}}(A^{\mathrm{T}})^{p-2}\\ \vdots & \vdots\\ CA^{p-2} & F^{\mathrm{T}}A^{\mathrm{T}}\\ CA^{p-1} & F^{\mathrm{T}}\end{bmatrix}, \tag{53}$$

$$J_{\mathrm{d}}\mathscr{O}_p \triangleq \begin{bmatrix} C & F^{\mathrm{T}}(A^{\mathrm{T}})^{p-1}\\ \vdots & \vdots\\ CA^{p-2} & F^{\mathrm{T}}A^{\mathrm{T}}\end{bmatrix}. \tag{54}$$

Then,

$$J_{\mathrm{u}}\mathscr{O}_p = J_{\mathrm{d}}\mathscr{O}_p A', \tag{55}$$

where

$$A' \triangleq \begin{bmatrix} A & 0 \\ 0 & (A^{\mathrm{T}})^{-1} \end{bmatrix}. \tag{56}$$

Hence,

$$A' = (J_{\mathrm{d}}\mathscr{O}_p)^{\dagger} J_{\mathrm{u}}\mathscr{O}_p = TA''T^{-1}, \tag{57}$$

where $M^{\dagger} \triangleq (M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}$ is the Moore–Penrose inverse of full column-rank matrix M and

$$A'' \triangleq (J_{\mathrm{d}}\mathscr{K}U_{2n})^{\dagger} J_{\mathrm{u}}\mathscr{K}U_{2n}. \tag{58}$$

From (57), we see that A' and A'' are similar matrices. This means that they have the same Jordan blocks in their Jordan canonical representations. Likewise, we have from (24) and (51),

$$C' \triangleq [C \;\; F^{\mathrm{T}}(A^{\mathrm{T}})^{p-1}] = J_{\mathrm{f}}\mathscr{O}_p = C''T^{-1}, \tag{59}$$

where

$$J_{\mathrm{f}} \triangleq [I_m \;\; 0_{m\times(p-1)}], \tag{60}$$

$$C'' \triangleq J_{\mathrm{f}}\mathscr{K}U_{2n}. \tag{61}$$

Let us put A'' into the following Jordan canonical form:

$$A'' = [\Pi_{\mathrm{c}} \;\; \Pi_{\mathrm{ac}}] \begin{bmatrix} \Sigma_{\mathrm{c}} & 0 \\ 0 & \Sigma_{\mathrm{ac}} \end{bmatrix} [\Pi_{\mathrm{c}} \;\; \Pi_{\mathrm{ac}}]^{-1}, \tag{62}$$

where the eigenvalues of Σ_{c} lie inside the unit circle. Since $H(z)$ is invariant to similarity transformations, we may let

$$A \triangleq \Sigma_{\mathrm{c}} \tag{63}$$

in (1). The canonical form (62) is invariant to the ordering of eigenvalues as long as the eigenvalues and the corresponding eigenvectors of Σ_{c} are permuted accordingly, in complex pairs. Moreover, from the similarity of A'' to A', in (1) we may let

$$\Sigma_{\mathrm{ac}} = (\Sigma_{\mathrm{c}}^{\mathrm{T}})^{-1}. \tag{64}$$

This, of course, imposes a certain structure on T. Let

$$\Pi = [\Pi_{\mathrm{c}} \;\; \Pi_{\mathrm{ac}}]. \tag{65}$$

Then, (62) can be written as

$$A' = \Pi^{-1}A''\Pi. \tag{66}$$

Hence from (57) and (66),

$$A' = TA''T^{-1} = \Pi^{-1}A''\Pi. \tag{67}$$

The relations among Σ, Π, and T are captured in the following lemma. Recall that A has distinct eigenvalues.

Lemma 4. *Let A'' be as in (58). Consider the Jordan canonical form of A'' given by (62) where A and Σ_{ac} satisfy (63) and (64). Then, Σ_{c} is a block diagonal matrix*

$$\Sigma_{\mathrm{c}} = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_k \end{bmatrix}, \quad \Sigma_i \in \mathbf{R}^{n_i \times n_i}, \tag{68}$$

where $n_i \in \{1,2\}$, $v_i \neq 0$, and

$$\Sigma_i \triangleq \begin{cases} \mu_i & \text{if } n_i = 1, \\ \begin{bmatrix} \mu_i & v_i \\ -v_i & \mu_i \end{bmatrix} & \text{if } n_i = 2. \end{cases} \tag{69}$$

Also, Σ_{ac} is a block diagonal matrix with block types and sizes compatible with Σ_{c}. For some Λ_{c} and Λ_{ac} compatible with Σ_{c}, the following holds:

$$\Pi = T^{-1}\Lambda, \tag{70}$$

where

$$\Lambda \triangleq \begin{bmatrix} \Lambda_{\mathrm{c}} & 0 \\ 0 & \Lambda_{\mathrm{ac}} \end{bmatrix}. \tag{71}$$

Let X and Y be two block diagonal matrices with block sizes and types compatible with Σ_{c}, then X^{T}, XY and X^{-1} are also compatible with Σ_{c} and $XY = YX$.

Proof. See Appendix B. □

Now, multiplying (51) from right by Π, we get

$$\mathscr{K}U_{2n}\Pi = \mathscr{O}_p\Lambda. \tag{72}$$

Hence, from (24), (65), and (71)

$$\mathscr{K}U_{2n}\Pi_{\mathrm{c}} = \begin{bmatrix} C\Lambda_{\mathrm{c}} \\ C\Sigma_{\mathrm{c}}\Lambda_{\mathrm{c}} \\ \vdots \\ C\Sigma_{\mathrm{c}}^{p-2}\Lambda_{\mathrm{c}} \\ C\Sigma_{\mathrm{c}}^{p-1}\Lambda_{\mathrm{c}} \end{bmatrix},$$

$$\mathscr{K}U_{2n}\Pi_{\mathrm{ac}} = \begin{bmatrix} F^{\mathrm{T}}(\Sigma_{\mathrm{c}}^{\mathrm{T}})^{p-1}\Lambda_{\mathrm{ac}} \\ F^{\mathrm{T}}(\Sigma_{\mathrm{c}}^{\mathrm{T}})^{p-2}\Lambda_{\mathrm{ac}} \\ \vdots \\ F^{\mathrm{T}}\Sigma_{\mathrm{c}}^{\mathrm{T}}\Lambda_{\mathrm{ac}} \\ F^{\mathrm{T}}\Lambda_{\mathrm{ac}} \end{bmatrix}.$$

Thus,

$$C\Lambda_{\mathrm{c}} = J_{\mathrm{f}}\mathscr{K}U_{2n}\Pi_{\mathrm{c}}, \quad F^{\mathrm{T}}\Lambda_{\mathrm{ac}} = J_{\mathrm{l}}\mathscr{K}U_{2n}\Pi_{\mathrm{ac}}, \tag{73}$$

where

$$J_{\mathrm{l}} \triangleq [0_{m\times(p-1)} \;\; I_m]. \tag{74}$$

The problem of finding the state-space matrices C, F, and E is now reduced to estimating E, Λ_{c}, and Λ_{ac} from the spectral data (8).

From Lemma 4, $S(z)$ in (16) can be written as

$$S(z) = E + C\Lambda_{\mathrm{c}}(z\Lambda_{\mathrm{ac}}^{\mathrm{T}}\Lambda_{\mathrm{c}} - \Lambda_{\mathrm{ac}}^{\mathrm{T}}\Sigma_{\mathrm{c}}\Lambda_{\mathrm{c}})^{-1}\Lambda_{\mathrm{ac}}^{\mathrm{T}}F$$
$$+ F^{\mathrm{T}}\Lambda_{\mathrm{ac}}(z^{-1}\Lambda_{\mathrm{c}}^{\mathrm{T}}\Lambda_{\mathrm{ac}} - \Lambda_{\mathrm{c}}^{\mathrm{T}}\Sigma_{\mathrm{c}}^{\mathrm{T}}\Lambda_{\mathrm{ac}})^{-1}\Lambda_{\mathrm{c}}^{\mathrm{T}}C^{\mathrm{T}}$$

$$= E + C\Lambda_c(zI_n - \Sigma_c)^{-1}\Lambda_c^{-1}F$$
$$+ F^T(\Lambda_c^T)^{-1}(z^{-1}I_n - \Sigma_c^T)^{-1}\Lambda_c^T C^T$$
$$= E + \chi(z)Z + Z^T\chi^T(z^{-1}), \tag{75}$$

where

$$\chi(z) \triangleq J_f \mathscr{K} U_{2n}\Pi_c(zI_n - \Sigma_c)^{-1}, \tag{76}$$

$$Z \triangleq \Lambda_c^{-1}F. \tag{77}$$

Thus, E and Z can be estimated from data (8) by solving the following linear least-squares problem:

$$E^\sharp, Z^\sharp \triangleq \arg\min_{\check{E},\check{Z}} \sum_{k=1}^{N} \|\mathscr{R}_k^{-1/2}(\chi(z_k)\check{Z} + \check{Z}^T\chi^T(z_k^{-1}) + \check{E} - S(z_k))\|_F^2. \tag{78}$$

Formula (78) is nonasymptotic in N though asymptotic quantities are used in it. However, it suggests a scheme to consistently estimate the state-space parameters A, C, E, and F.

Recall that when the spectrum samples are noise-free, we can replace U_{2n} with $\widehat{U}_{2n}$ in the above formulae. Thus, we have the following result.

Lemma 5. *Let $S(z)$ be the power spectrum of a discrete-time system of order n satisfying Assumption 2.1. Let $\chi(z)$ and Z be as in (76) and (77), respectively. Consider the linear least-squares problem (78). If $N \geqslant (p/2) + n + 1$, then*

$$E^\sharp = E, \quad Z^\sharp = Z. \tag{79}$$

Proof. The proof of this lemma is contained in the proof of Theorem 7. □

Once we find Z, we calculate C and F from the first equation in (73) and (77) as

$$C = J_f\mathscr{K}U_{2n}\Pi_c, \quad F = Z \tag{80}$$

which is due to the fact that $H(z)$ defined by (17) is invariant to post-multiplication of C by Λ_c and pre-multiplication of F by Λ_c^{-1} since from Lemma 4, we have $\Lambda_c^{-1}(zI_n - \Sigma_c)^{-1}\Lambda_c = (zI_n - \Sigma_c)^{-1}$.

We are left with the determination of the system matrices B and D. To this end, we first solve the following *Riccati* equation for P:

$$P = APA^T + (F - APC^T)(E - CPC^T)^{-1} \cdot (F - APC^T)^T. \tag{81}$$

Then, we compute B and D as follows:

$$B = (F - APC^T)(E - CPC^T)^{-1/2}, \tag{82}$$

$$D = (E - CPC^T)^{1/2}. \tag{83}$$

Now, we return to the normal case to outline the proposed algorithm. Let

$$\tilde{A} \triangleq (J_d\mathscr{K}\widehat{U}_{2n})^\dagger J_u\mathscr{K}\widehat{U}_{2n} \tag{84}$$

and put $\tilde{A}$ into the Jordan canonical form:

$$\tilde{A} \triangleq [\widehat{\Pi}_c \;\; \widehat{\Pi}_{ac}]\begin{bmatrix}\widehat{\Sigma}_c & 0\\ 0 & \widehat{\Sigma}_{ac}\end{bmatrix}[\widehat{\Pi}_c \;\; \widehat{\Pi}_{ac}]^{-1}, \tag{85}$$

where the eigenvalues of $\widehat{\Sigma}_c$ lie inside the unit circle. Let

$$\widehat{A} \triangleq \widehat{\Sigma}_c, \tag{86}$$

$$\widehat{C} \triangleq J_f\mathscr{K}\widehat{U}_{2n}\widehat{\Pi}_c. \tag{87}$$

From (50) and (58), we have

$$\lim_{N\to\infty} \tilde{A} = A'', \quad \text{w.p.1.} \tag{88}$$

As in the calculation of $\widehat{U}_{2n}$, we can freely choose $2n$ elements of $\widehat{\Pi}_c$ and $\widehat{\Pi}_{ac}$ with fixed indices, subject to the constraint that magnitudes are not greater than unity. Then, by fixing values of those elements equal to the values of the corresponding elements in Π_c and Π_{ac} for all N, we see from (88) and (85) that

$$\lim_{N\to\infty} \widehat{\Sigma}_c = \Sigma_c \quad \text{and} \quad \lim_{N\to\infty} \widehat{\Sigma}_{ac} = \Sigma_{ac}, \quad \text{w.p.1,}$$

$$\lim_{N\to\infty} \widehat{\Pi}_c = \Pi_c \quad \text{and} \quad \lim_{N\to\infty} \widehat{\Pi}_{ac} = \Pi_{ac}, \quad \text{w.p.1.} \tag{89}$$

Let

$$\widehat{\chi}(z) \triangleq \widehat{C}(zI_n - \widehat{\Sigma}_c)^{-1}. \tag{90}$$

Then, from (89) and the fact that $\widehat{U}_{2n} \to U_{2n}$ w.p.1 as $N \to \infty$ we have

$$\lim_{N\to\infty} \|\widehat{\chi} - \chi\|_\infty, \quad \text{w.p.1.} \tag{91}$$

The uniform convergence is due to the fact that the spectral radius of the limit matrix Σ_c is less than one.

The estimates of E and F are obtained by solving the following linear least-squares problem:

$$\widehat{E}, \widehat{F} \triangleq \arg\min_{\check{E},\check{F}} \sum_{k=1}^{N} \|\mathscr{R}_k^{-1/2}(\widehat{\chi}(z_k)\check{F} + \check{F}^T\widehat{\chi}^T(z_k^{-1}) + \check{E} - S_k)\|_F^2. \tag{92}$$

Before concluding the consistency analysis, let us summarize the final algorithm in the following.

Algorithm 1. Subspace algorithm with nonuniformly spaced spectrum samples:

(1) Given the data S_k, θ_k, and the covariance data $\mathscr{R}_k$, form matrices $\mathscr{S}$, $\mathscr{W}_C$, $\mathscr{W}$, and $\mathscr{K}$ defined by (34), (28), (37), and (45).
(2) Calculate the QR-factorization in (47).
(3) Calculate the SVD in (49).

(4) Determine the system order n by inspecting the singular values and partition the SVD such that $\widehat{\Sigma}_{2n}$ contains the $2n$ largest singular values.
(5) With J_u, J_d, and $\widehat{U}_{2n}$ defined by (53), (54), and (49), calculate $\tilde{A}$ from (84).
(6) Block-diagonalize $\tilde{A}$ as in (85) and let $\widehat{\Pi}_c$ and $\widehat{A}$ be as in (85) and (86).
(7) With J_f defined by (60), let $\widehat{C}$ be as in (87).
(8) Solve the least-squares problem (92) for $\widehat{E}$ and $\widehat{F}$ where $\widehat{\chi}$ is defined by (90).
(9) Solve the Riccati equation for $\widehat{P}$:

$$\widehat{P} = \widehat{A}\widehat{P}\widehat{A}^{\mathrm{T}} + (\widehat{F} - \widehat{A}\widehat{P}\widehat{C}^{\mathrm{T}})(\widehat{E} - \widehat{C}\widehat{P}\widehat{C}^{\mathrm{T}})^{-1} \cdot (\widehat{F} - \widehat{A}\widehat{P}\widehat{C}^{\mathrm{T}})^{\mathrm{T}} \tag{93}$$

and calculate $\widehat{B}$ and $\widehat{D}$ from

$$\widehat{B} = (\widehat{F} - \widehat{A}\widehat{P}\widehat{C}^{\mathrm{T}})(\widehat{E} - \widehat{C}\widehat{P}\widehat{C}^{\mathrm{T}})^{-1/2}, \tag{94}$$

$$\widehat{D} = (\widehat{E} - \widehat{C}\widehat{P}\widehat{C}^{\mathrm{T}})^{1/2}. \tag{95}$$

(10) Calculate $\widehat{G}(z)$ and $\widehat{S}_N(z)$ from (11) and (9). □

Combination of Lemmas 3 and 5 yields our first result captured in the following.

Theorem 6. *Consider Algorithm* 1 *with N noise-free samples of the power spectrum of a discrete-time system of order n satisfying Assumption* 2.1 *at N distinct frequencies θ_k. Let $\mathscr{K} \in \mathbf{R}^{pm\times pm}$ be any nonsingular matrix. If $N \geqslant (p/2) + n + 1$, then Algorithm* 1 *is correct.*

Now, we finish the consistency analysis of Algorithm 1.

Theorem 7. *Consider Algorithm* 1 *with corrupted measurements of the power spectrum of a discrete-time system of order n satisfying Assumption* 2.1 *where the corruptions and the frequencies satisfy assumptions* (5)–(7). *Then, Algorithm* 1 *is strongly consistent.*

Proof. See Appendix C. □

The algorithm described in Van Overschee et al. (1997) is a special case of Algorithm 1. The only difference between the algorithms is the choice of the annihilator $\mathscr{W}^{\perp}$. In Algorithm 1, a maximal rank annihilator is used whereas in Van Overschee et al. (1997) an annihilator of much smaller rank is used. In the nonuniform case, we cannot a priori derive a smaller matrix to cancel $\Gamma_p\mathscr{W}$ in (33) since there is a risk of canceling some of the row space of $\mathscr{X}$. The details can be found in McKelvey et al. (1996a).

Another issue to be addressed is the positivity of the power spectrum. Any physically meaningful power spectrum must be positive real. The power spectrum estimated by the above algorithm may not satisfy this requirement due to noise and undermodeling. This requirement manifests itself as the existence of a positive definite solution of (81). If a positive definite solution fails to exist, then the spectral factor cannot be computed. Thus, the positivity of the spectrum should be enforced after the identification. There are many possibilities. Two methods enforcing the positivity condition (4) are outlined in Van Overschee et al. (1997). These methods can be integrated into Algorithm 1 without modification.

4. Examples

In this section, we use two identification examples to illustrate the properties of the developed algorithm. The first example is based on simulated data. This example will show us the role played by the noise covariance information. The second example deals with the design of a linear shaping filter from measured road data.

4.1. Simulation example

Let the true system $G(z) = C(zI_4 - A)^{-1}B + D$ be a fourth-order system described by the state-space model:

$$A = \begin{bmatrix} 0.8876 & 0.4494 & 0 & 0 \\ -0.4494 & 0.7978 & 0 & 0 \\ 0 & 0 & -0.6129 & 0.0645 \\ 0 & 0 & -6.4516 & -0.7419 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.2247 \\ 0.8989 \\ 0.0323 \\ 0.1290 \end{bmatrix},$$

$$C = [0.4719 \;\; 0.1124 \;\; 9.6774 \;\; 1.6129],$$

$$D = 0.9626.$$

We assume N noisy samples S_k of the power spectrum $S(z)$ evaluated at N points on the unit circle are given as

$$S_k = S(\mathrm{e}^{\mathrm{j}\theta_k}) + \tilde{S}(\mathrm{e}^{\mathrm{j}\theta_k})v_k, \quad k = 1,\ldots,N,$$

where the noise term $\tilde{S}(\mathrm{e}^{\mathrm{j}\theta_k})v_k$ is composed of a noise transfer function $\tilde{S}(z)$, given by a second-order state-space model:

$$\tilde{S}(z) = \tilde{C}(zI_2 - \tilde{A})^{-1}\tilde{B} + \tilde{D}$$

with

$$\tilde{A} = \begin{bmatrix} 0.6296 & 0.0741 \\ -7.4074 & 0.4815 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 0.04 \\ 0.9 \end{bmatrix},$$

$$\tilde{C} = [1.6300 \;\; 0.0740], \quad \tilde{D} = 0.2$$

and v_k being independent complex identically distributed normal random variables with zero mean and unit variance. The variance of the noise process at each frequency equals $\mathscr{R}_k = |\tilde{S}(z_k)|^2$. We picked the frequencies randomly and

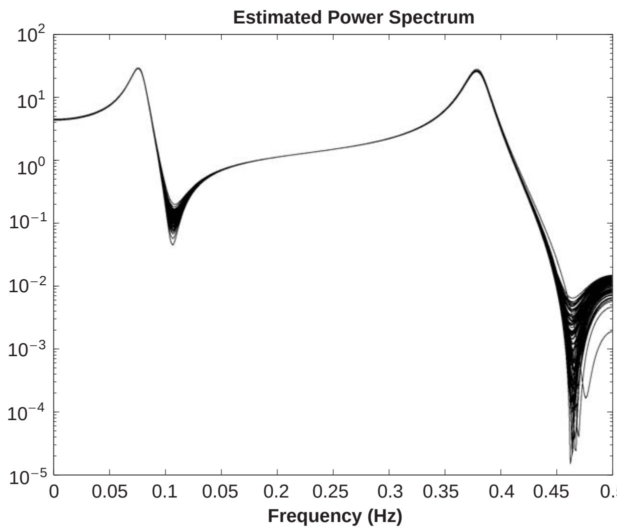

Fig. 1. The results from Monte Carlo simulations for the 100 estimated models using the covariance information.

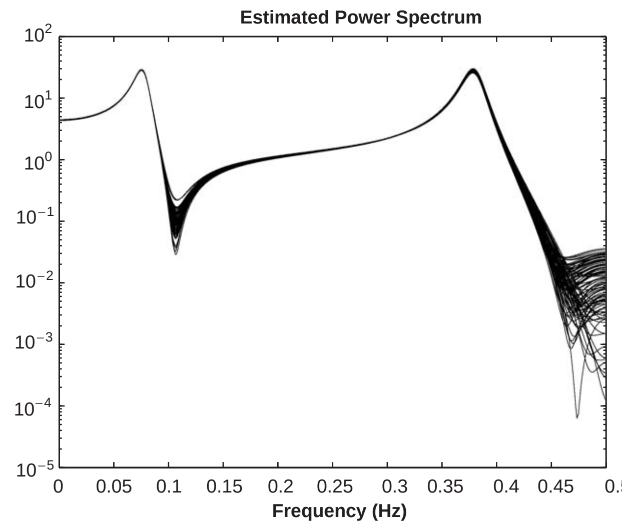

Fig. 2. The results from Monte Carlo simulations for the 100 estimated models without using the covariance information.

independently from the intervals

$$\left[\frac{\pi}{N}\left(k-\frac{1}{2}\right), \frac{\pi}{N}\left(k+\frac{1}{2}\right)\right], \quad k=1,\ldots,N.$$

Thus, each θ_k has a uniform distribution.

To examine the consistency properties of Algorithm 1, we performed Monte Carlo simulations estimating the power spectrum, given the samples S_k, using different noise realizations of v_k. For $N=400$ and fixed frequencies, 100 different noise realizations were generated, and Algorithm 1 with $p=50$ estimated 100 models. To assess the quality of the resulting model, both the (measured) supremum norm

$$\|\widehat{S}_n - S\|_{m,\infty} \triangleq \max_{1\leqslant k\leqslant N} |\widehat{S}_N(z_k) - S(z_k)|$$

and the (measured) $\mathscr{H}_2$ norm

$$\|\widehat{S}_n - S\|_{m,2} \triangleq \left(\frac{1}{N}\sum_{k=1}^{N} |\widehat{S}_N(z_k) - S(z_k)|^2\right)^{1/2}$$

of the estimation error were determined for each estimated model and averaged over the 100 estimated models. In Fig. 1, the results for the 100 estimated models using the covariance information are shown. We computed $\|\widehat{S}_n - S\|_{m,2} = 0.3307$ and $\|\widehat{S}_n - S\|_{m,\infty} = 2.4208$.

In Fig. 2, the results for the 100 estimated models without using the covariance information, i.e. $\mathscr{R}_k = 1$ for all k, are shown. We computed $\|\widehat{S}_n - S\|_{m,2} = 0.4500$ and $\|\widehat{S}_n - S\|_{m,\infty} = 2.6240$.

Comparing $\mathscr{H}_2$ errors, as predicted by the analysis, using the noise covariance information in Algorithm 1 reduces the estimation error by about 30%.

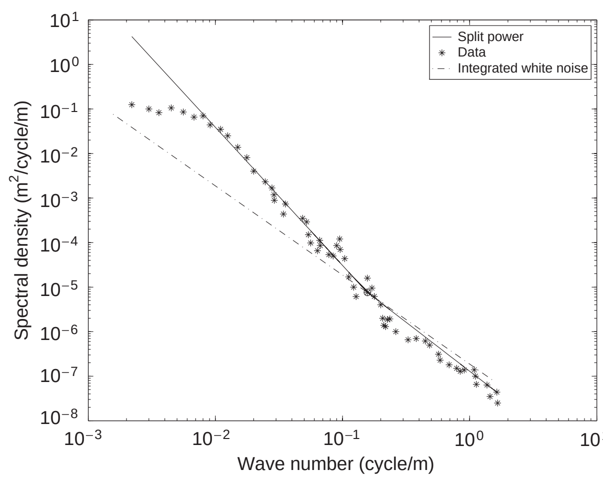

Fig. 3. The road power spectrum (Dodds & Robson, 1973) and its approximate modeling by the split power law and the integrated white noise.

4.2. Stochastic road modeling example

In this subsection, we consider one practical application of Algorithm 1. Provided that occasional large irregularities such as potholes are removed from the analysis, the road surface may be described as a realization of a stationary random process. This assumption enables one to determine the response of a vehicle traversing a road by accepted techniques of the theory of random vibration. If the road surface is further assumed to be homogenous and isotrophic, then a road profile can be completely described by a single power spectral density evaluated from any single track.

In Fig. 3 (Dodds & Robson, 1973), the spectral density of a typical road and its split power law

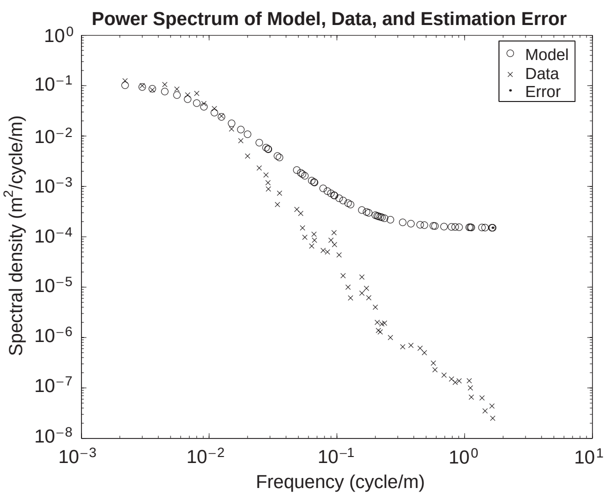

Fig. 4. The spectral data and its modeling by a rational model of order one produced by Algorithm 1 with $\mathscr{R} = I_N$.

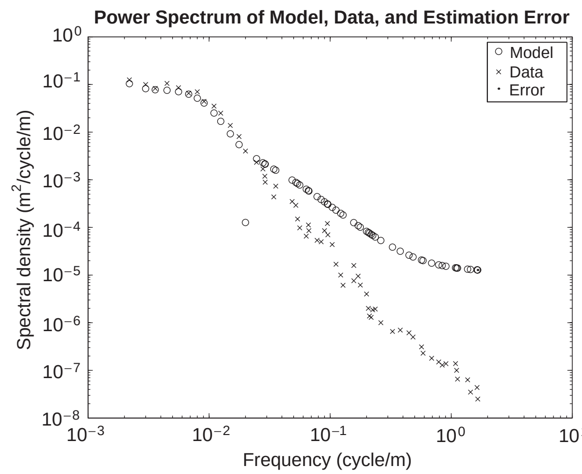

Fig. 5. The spectral data and its modeling by a rational model of order 7 produced by Algorithm 1 with $\mathscr{R} = I_N$.

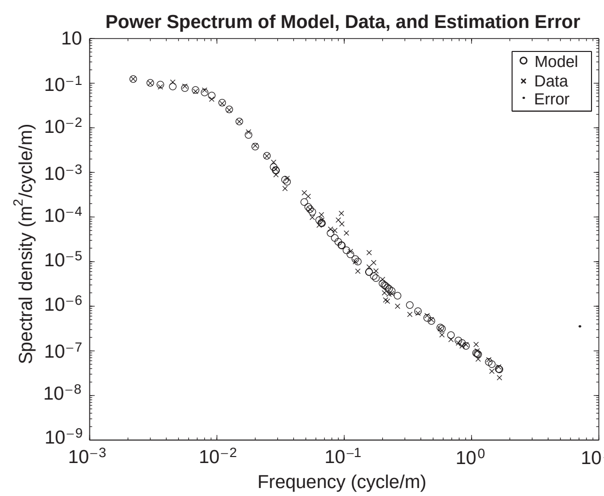

Fig. 6. The spectral data and its modeling by a rational model of order 7 produced by Algorithm 1 with $\mathscr{R}_k = S_k$.

approximation:

$$\widehat{S}(j2\pi\tilde{n}) = \begin{cases} C|\tilde{n}/\tilde{n}_0|^{-2\delta_1}, & 0 < |\tilde{n}| < \tilde{n}_0, \\ C|\tilde{n}/\tilde{n}_0|^{-2\delta_2}, & \tilde{n}_0 \leqslant |\tilde{n}| < \infty \end{cases}$$

obtained by trial and error for $\tilde{n}_0 = 0.15708$ cycles/m, $\delta_1 = 1.6$, $\delta_2 = 1.1$, and $C = 0.76 \times 10^{-5}$ are plotted. In the figure, we also show the integrated white noise approximation to the data: $C|\tilde{n}/\tilde{n}_0|^{-2}$ which is commonly used in stochastic road modeling. It is clear that the fit by the integrated white-noise modeling is rather poor, in particular at the frequencies below $\tilde{n}_0$. The problem with the split power approximation is that it cannot be generated by shape filters. Hence, it is not suitable for simulating the response of vehicle. Besides, it is unbounded at $\tilde{n} = 0$.

In this application, we seek a low order shape filter whose output spectrum matches the spectral data in Fig. 3 as closely as possible. The continuous-time estimation problem is converted to a discrete one by using the bilinear map:

$$s = \psi(z) = \lambda \frac{z-1}{z+1} \quad (\lambda > 0).$$

The number of data is $N = 63$. We picked $\lambda = 0.2$ and $p = 32$ in Algorithm 1. In the first trial, we choose $n = 1$ and $\mathscr{R}_N = I_N$. The continuous-time spectral factor was obtained by substituting $z = \psi^{-1}(s)$ in the discrete-time spectral factor. Thus,

$$\widehat{G}_N(s) = 0.0122\,\frac{s + 1.1154}{s + 0.0404}.$$

In Fig. 4, the output spectrum and the estimation error of this transfer function are compared with the road data. This figure tells us that the first order rational filter produced by Algorithm 1 is accurate up to 0.02 cycles/m.

To interpret this result, assume that the bandwidth of the vehicle suspension is 10 Hz and the forward velocity of the vehicle is 30 m/s. (The vehicle frequency response rolls off at least 20 decibels per decade.) This corresponds to a spectral bandwidth of $\frac{1}{3}$ cycles/m in Fig. 3. Since the road power spectrum rapidly rolls off, we conclude that the first order model is a good first-degree approximation. It should be noted that the power spectrum of this model is not integrable. A convergence factor rolling the frequency response off at high frequencies may be introduced.

Next, we tried higher model orders with $\mathscr{R}$ as a design variable. In Figs. 5 and 6, the output spectra and the estimation errors are compared with the road data for $n = 7$, $p = 32$, and the two cases $\mathscr{R} = I_N$ and $\mathscr{R}_k = S_k$. Clearly, Fig. 6 indicates improvement in the high frequency rolling caused by weighting.

The purpose of modeling a power spectrum by a rational function of reasonably low order is to use this approximation for the design of a linear shaping filter with a white noise input. Then, the identified road spectrum is used, for example in a quarter-car model, to study the response of the vehicle to random road inputs (Türkay & Akçay, 2004).

5. Conclusions

In this paper, we presented a strongly consistent subspace algorithm for the identification of square multi-input/multi-output, discrete-time, linear-time invariant systems from nonuniformly spaced power spectrum measurements. The algorithm was illustrated with one practical example that solves a stochastic road modeling problem.

Acknowledgements

The authors would like to thank Professor M. Verhaegen of the Delft University of Technology for sending them a preprint of Fraanje et al. (2003).

Appendix A. Proof of Lemma 2

Eq. (22) implies that

$$X_{\mathrm{C}}^{\mathrm{c}}(\theta)=(\mathrm{e}^{\mathrm{j}\theta}I_n-A)^{-1}F,$$

$$X_{\mathrm{C}}^{\mathrm{ac},p}(\theta)=\mathrm{e}^{\mathrm{j}(p-1)\theta}(\mathrm{e}^{-\mathrm{j}\theta}I_n-A^{\mathrm{T}})^{-1}C^{\mathrm{T}}.$$

The matrix $\begin{bmatrix}\mathscr{W}_{\mathrm{C}}\\ \mathscr{X}_{\mathrm{C}}\end{bmatrix}$ is rank deficient if and only if there exists a row vector

$$[\alpha_1 \;\cdots\; \alpha_p \;\; \beta \;\; \gamma]\neq 0 \tag{A.1}$$

with $\alpha_k^{\mathrm{T}}\in\mathbf{R}^m$, $k=1,\dots,p$ and $\beta^{\mathrm{T}},\gamma^{\mathrm{T}}\in\mathbf{R}^n$ such that

$$[\alpha_1 \;\cdots\; \alpha_p \;\; \beta \;\; \gamma]\begin{bmatrix}\mathscr{W}_{\mathrm{C}}\\ \mathscr{X}_{\mathrm{C}}\end{bmatrix}=0$$

$$\Leftrightarrow$$

$$\mathscr{J}(z_k)=0,\quad k=1,\dots,N, \tag{A.2}$$

where

$$\mathscr{J}(z)=\sum_{k=1}^{p}\alpha_k z^{k-1}+\gamma z^p(I_n-zA^{\mathrm{T}})^{-1}C^{\mathrm{T}}$$
$$+\beta(zI_n-A)^{-1}F. \tag{A.3}$$

Since $\mathscr{J}(z)$ is a real-rational matrix, z_k is a zero of it if and only if z_k^* is a zero of it. Thus, each element of $\mathscr{J}(z)$ has at least $2N-2$ zeros whenever (A.2) holds. (If $z_k\notin\mathbf{R}$ for all k, then the number of zeros is precisely $2N$.)

Let $\rho(A)$ be the spectral radius of A, i.e. the largest of the magnitudes of the eigenvalues of A. The Laurent series of $\mathscr{J}(z)$ converges in the annulus

$$\mathbf{D}_A \triangleq \{z\in\mathbf{C}:\rho(A)<|z|<[\rho(A)]^{-1}\}. \tag{A.4}$$

Each element in the rational vector $\mathscr{J}(z)$ is either identically zero or has at least $2N-2$ zeros at $\mathrm{e}^{\pm\mathrm{j}\theta_k}$; but each element of $\mathscr{J}(z)$ can have at most $p+2n-1$ zeros. Since $2N\geqslant p+2n+2$, we then have $\mathscr{J}(z)\equiv 0$. This implies that all the coefficients in the Laurent expansion of $\mathscr{J}(z)$ are zero. Since the Laurent series of $\gamma z^p(I_n-zA^{\mathrm{T}})^{-1}C^{\mathrm{T}}$ starts with $\gamma C^{\mathrm{T}}z^p$ and converges in the disk

$$\mathbf{D}_\rho \triangleq \{z\in\mathbf{C}:\;|z|<[\rho(A)]^{-1}\} \tag{A.5}$$

the three terms on the right-hand side of (A.3) are independent; and therefore they are identically zero. Hence,

$$\alpha_k=0,\quad k=1,\dots,p,$$

$$\gamma(z^{-1}I_n-A^{\mathrm{T}})^{-1}C^{\mathrm{T}}\equiv 0,$$

$$\beta(zI_n-A)^{-1}F\equiv 0.$$

The minimality of (A,B,C,D) implies the minimality of $(A,F,C,\frac{1}{2}E)$. Thus, from the controllability of the pairs (A,F) and $(A^{\mathrm{T}},C^{\mathrm{T}})$ we have $\beta=\gamma=0$. Hence, (A.1) is violated. Finally, note that $\begin{bmatrix}\mathscr{W}_{\mathrm{C}}\\ \mathscr{X}_{\mathrm{C}}\end{bmatrix}$ is rank deficient if and only $\begin{bmatrix}\mathscr{W}\\ \mathscr{X}\end{bmatrix}$ is rank deficient. The last assertion is due to the fact that for any complex matrix Z,

$$x^{\mathrm{T}}Z=0 \;\;\Leftrightarrow\;\; x\,[\mathrm{Re}\,Z \;\; \mathrm{Im}\,Z]=0. \quad \square$$

Appendix B. Proof of Lemma 4

The first two claims are obvious. From (67),

$$A''T^{-1}=T^{-1}A',\qquad A''\Pi=\Pi A'. \tag{B.1}$$

Now, partition T^{-1} and Π as

$$T^{-1}=[t_1 \;\cdots\; t_{2n}],\qquad \Pi=[\pi_1 \;\cdots\; \pi_{2n}].$$

If $n_i=1$, put $l=n_1+\cdots+n_i$. Then, from (67)

$$A''t_l=\mu_i t_l,\qquad A''\pi_l=\mu_i\pi_l$$

which shows that t_l is an eigenvector of Σ_{c} associated with the eigenvalue μ_i. Thus, for some $\Lambda_{n_i}\in\mathbf{R}$

$$\pi_l=\Lambda_{n_i}t_l.$$

This equality is due to the fact that eigenvectors corresponding to a simple real eigenvalue span a one dimensional subspace of $\mathbf{R}^n$. If $n_i=2$, again putting $l=n_1+\cdots+n_i$, from (67) we get

$$A''[t_l \;\; t_{l+1}]=[t_l \;\; t_{l+1}]\Sigma_i,$$

$$A''[\pi_l \;\; \pi_{l+1}]=[\pi_l \;\; \pi_{l+1}]\Sigma_i$$

which shows that t_l and t_{l+1} are eigenvectors of Σ_{c} associated with the eigenvalues $\mu_i\pm jv_i$. It is known that eigenvectors corresponding to a pair of simple complex eigenvalues form a two dimensional subspace of $\mathbf{R}^n$. Hence, for some $\beta\in\mathbf{R}^{2\times 2}$

$$\pi_l=\beta_{11}t_l+\beta_{12}t_{l+1},$$

$$\pi_{l+1}=\beta_{21}t_l+\beta_{22}t_{l+1}. \tag{B.2}$$

Multiplying both sides of the first equation in (B.2) with A'' and using the equations in (B.2), we get

$$(\beta_{12}+\beta_{21})v_i t_l + (\beta_{22}-\beta_{11})v_i t_{l+1} = 0.$$

Since $v_i \neq 0$ and t_l and t_{l+1} are linearly independent vectors, we then must have

$$\beta_{11}=\beta_{22}, \quad \beta_{21}=-\beta_{12}.$$

It follows that

$$[\pi_l \ \ \pi_{l+1}] = \Lambda_i[t_l \ \ t_{l+1}],$$

where

$$\Lambda_i \triangleq \begin{bmatrix} \beta_{11} & \beta_{12} \\ -\beta_{12} & \beta_{11} \end{bmatrix}.$$

Let

$$\Lambda_c \triangleq \begin{bmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_k \end{bmatrix}.$$

Then, Λ_c is compatible with Σ_c and

$$[\pi_1 \ \cdots \ \pi_n] = \Lambda_c\,[t_1 \ \cdots \ t_n]. \tag{B.3}$$

Likewise, for some Λ_{ac} compatible with Σ_{ac} we get

$$[\pi_{n+1} \ \cdots \ \pi_{2n}] = \Lambda_{ac}[t_{n+1} \ \cdots \ t_{2n}]. \tag{B.4}$$

Since Σ_{ac} is compatible with Σ_c, Λ_{ac} is compatible with Σ_c. Thus, combining (B.3) and (B.4), we get (70). The last claims are easy to verify. □

Appendix C. Proof of Theorem 7

Let

$$\tilde{\mathscr{S}}_N(z) \triangleq \hat{\chi}(z)\check{F} + \check{F}^{\mathrm{T}}\hat{\chi}^{\mathrm{T}}(z^{-1}) + \check{E}.$$

The least-squares problem (92) can be written as

$$\hat{E},\hat{F} = \arg\min_{\check{E},\check{F}}(\hat{Q}_N - \hat{T}_N + D_N), \tag{C.1}$$

where

$$\hat{Q}_N \triangleq \frac{1}{N}\sum_{k=1}^{N}\|\mathscr{R}_k^{-1/2}\tilde{\mathscr{S}}_N(z_k)\|_F^2,$$

$$\hat{T}_N \triangleq \frac{1}{N}\sum_{k=1}^{N}\mathrm{Tr}\{S_k^{\mathrm{H}}\mathscr{R}_k^{-1}\tilde{\mathscr{S}}_N(z_k) + \tilde{\mathscr{S}}_N^{\mathrm{H}}(z_k)\mathscr{R}_k^{-1}S_k\},$$

$$D_N \triangleq \frac{1}{N}\sum_{k=1}^{N}\|\mathscr{R}_k^{-1/2}S_k\|_F^2.$$

In the derivation of (C.1), we have used the facts that $\mathrm{Tr}(AB)=\mathrm{Tr}(BA)$ and $\mathrm{Tr}(A)=\mathrm{Tr}(A^{\mathrm{T}})$ for any matrices A and B of compatible sizes. The boundedness of fourth-order moments means the boundedness of second order moments. More precisely, $\mathbf{E}\|\eta_k\|_F^2 \leqslant [\mathbf{E}\|\eta_k\|_F^4]^{1/2}$. Hence from (6) and the chain of (in)equalities

$$[\sigma_{\max}^{1/2}(\mathscr{R}_k)]^2 = \sigma_{\max}^2(\mathscr{R}_k^{1/2}) \leqslant \|\mathscr{R}_k^{1/2}\|_F^2 = \mathrm{Tr}(\mathscr{R}_k) = \mathbf{E}\|\eta_k\|_F^2,$$

we get $\sigma_{\max}(\mathscr{R}_k) \leqslant M_\eta^{1/2}$. Let $\sigma_{\min}$ denote the smallest singular value. The inequality $\|XY\|_F \geqslant \sigma_{\min}(X)\|Y\|_F$ valid for any matrices of X and Y of compatible sizes then yields

$$\hat{Q}_N \geqslant \frac{1}{M_\eta^{1/2}N}\sum_{k=1}^{N}\|\tilde{\mathscr{S}}_N(z_k)\|_F^2. \tag{C.2}$$

From (91), we have for each $\check{E}$ and $\check{F}$

$$\lim_{N\to\infty}\|\tilde{\mathscr{S}}_N - \check{S}\|_\infty = 0, \quad \text{w.p.1}, \tag{C.3}$$

where

$$\check{S}(z) \triangleq \chi(z)\check{F} + \check{F}^{\mathrm{T}}\chi^{\mathrm{T}}(z^{-1}) + \check{E}. \tag{C.4}$$

Hence,

$$\lim_{N\to\infty}\inf \hat{Q}_N \geqslant M_\eta^{-1/2}\lim_{N\to\infty}\inf\frac{1}{N}\sum_{k=1}^{N}\|\check{\mathscr{S}}(z_k)\|_F^2, \quad \text{w.p.1.}$$

We claim that if $\check{E}$ and $\check{F}$ is a nontrivial pair, then $\check{S}(z)$ can vanish only at a finite number of points z_k. To establish this claim, suppose that

$$\check{S}(z_k)=0, \quad k=1,\ldots,M.$$

Then, from $\check{S}(z_k^{-1})=\check{S}^{\mathrm{T}}(z_k)$ we see that these inequalities are also satisfied with z_k^{-1}, $k=1,\ldots,M$. Thus, using the same argument in the proof of Lemma 2, if $2M-2>2n$ and the frequencies are distinct, we conclude that $\check{S}(z)$ is identically zero (since its each entry can have at most $2n$ zeros). Let $\mathbf{D}_A$ and $\mathbf{D}_\rho$ be as in (A.4) and (A.5), respectively. Then, all the coefficients in the Laurent expansion of $\check{\mathscr{S}}(z)$, which converges in $\mathbf{D}_A$, are zero. The Laurent series of $\check{F}^{\mathrm{T}}\chi^{\mathrm{T}}(z^{-1})$ starts with $z\check{F}^{\mathrm{T}}[J_f\mathscr{K}U_{2n}\Pi_c]^{\mathrm{T}}$ and converges in the disk $\mathbf{D}_\rho$. Therefore, the three terms on the right-hand side of (C.4) are independent; and thus they are identically zero. Hence,

$$\check{E}=0, \quad \check{F}^{\mathrm{T}}(z^{-1}I_n - \Sigma_c^{\mathrm{T}})^{-1}\Lambda_c^{\mathrm{T}}C^{\mathrm{T}} \equiv 0. \tag{C.5}$$

Let $x\in\mathbf{R}^m$ be such that $\check{F}x \triangleq \beta \neq 0$. This is possible since $\check{E}=0$ implies $\check{F}\neq 0$. Then, from the second equation in (C.5) we have

$$\beta^{\mathrm{T}}(z^{-1}I_n - \Sigma_c^{\mathrm{T}})^{-1}\Lambda_c^{\mathrm{T}}C^{\mathrm{T}} \equiv 0$$

which means that $(\Sigma_c^{\mathrm{T}}, \Lambda_c^{\mathrm{T}}C^{\mathrm{T}})$ is not a controllable pair. Since Λ_c is nonsingular, this means $(A^{\mathrm{T}}, C^{\mathrm{T}})$ is not controllable, i.e. (A,C) is not observable. Thus, we reach a contradiction and $\check{S}(z)$ is nonzero in the complement of at most $2n$ points. Since $\|\check{S}(z)\|_F^2$ is uniformly continuous on the unit circle, a standard compactness argument then yields

$$\|\check{S}(\mathrm{e}^{\mathrm{j}\theta})\|_F^2 \geqslant \gamma, \quad \theta\in\bigcup_{i=1}^{r}[a_i,b_i]$$

for some $\gamma > 0$ and disjoint intervals $[a_i, b_i] \subseteq [0, 2\pi]$ satisfying $\sum_{i=1}^{r}(b_i - a_i) > \pi$. Thus, from (7) we obtain for all sufficiently large N

$$\frac{1}{N}\sum_{k=1}^{N}\|\check{\mathscr{S}}(z_k)\|_F^2 \geqslant \delta\gamma\pi.$$

We have shown that

$$(\check{E},\check{F}) \neq 0 \quad \Leftrightarrow \quad \liminf_{N\to\infty} \widehat{Q}_N > 0, \qquad \text{w.p.1.} \tag{C.6}$$

Let $\mathrm{vec}(S_k)$ denote the vector formed by stacking the columns of S_k into one long vector:

$$\mathrm{vec}(S_k) \triangleq \begin{bmatrix} S_{k,11} \\ \vdots \\ S_{k,m1} \\ S_{k,12} \\ \vdots \\ S_{k,m2} \\ S_{k,1m} \\ \vdots \\ S_{k,mm} \end{bmatrix}.$$

Let

$$\check{\Gamma} \triangleq \begin{bmatrix} \mathrm{vec}(\check{E}) \\ \mathrm{vec}(\check{F}) \end{bmatrix}.$$

The *Kronecker product* of two matrices $X \in \mathbf{C}^{m\times n}$ and $Y \in \mathbf{C}^{p\times q}$ is defined as

$$X \otimes Y \triangleq \begin{bmatrix} X_{11}Y & X_{12}Y & \cdots & X_{1n}Y \\ X_{21}Y & X_{22}Y & \cdots & X_{2n}Y \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1}Y & X_{m2}Y & \cdots & X_{mn}Y \end{bmatrix} \in \mathbf{C}^{mp\times nq}.$$

For each k, we can write $\mathrm{vec}(\tilde{\mathscr{S}}_N(z_k))$ and $\mathrm{vec}(\check{\mathscr{S}}_N(z_k))$ as linear functions in $\check{\Gamma}$:

$$\tilde{\mathscr{A}}_{N,k}\check{\Gamma} = \mathrm{vec}(\tilde{\mathscr{S}}_N(z_k)), \quad \check{\mathscr{A}}_{N,k}\check{\Gamma} = \mathrm{vec}(\check{\mathscr{S}}_N(z_k)) \tag{C.7}$$

for some matrices $\tilde{\mathscr{A}}_{N,k}$ and $\check{\mathscr{A}}_{N,k}$. To be specific on this, let $\widehat{\chi}_i(z_k)$ and $\chi_i(z_k)$ denote the ith rows of $\widehat{\chi}$ and $\chi(z_k)$, respectively. Let for each k,

$$\tilde{\mathscr{B}}_{N,k} \triangleq \begin{bmatrix} \mathrm{e}_1^{\mathrm{T}} \otimes \widehat{\chi}(z_k) + I_m \otimes \widehat{\chi}_1^*(z_k) \\ \vdots \\ \mathrm{e}_m^{\mathrm{T}} \otimes \widehat{\chi}(z_k) + I_m \otimes \widehat{\chi}_m^*(z_k) \end{bmatrix},$$

$$\check{\mathscr{B}}_{N,k} \triangleq \begin{bmatrix} \mathrm{e}_1^{\mathrm{T}} \otimes \chi(z_k) + I_m \otimes \chi_1^*(z_k) \\ \vdots \\ \mathrm{e}_m^{\mathrm{T}} \otimes \chi(z_k) + I_m \otimes \chi_m^*(z_k) \end{bmatrix}.$$

Then,

$$\tilde{\mathscr{A}}_{N,k} \triangleq \begin{bmatrix} I_{m^2} & 0 \\ 0 & \tilde{\mathscr{B}}_{N,k} \end{bmatrix}, \quad \check{\mathscr{A}}_{N,k} \triangleq \begin{bmatrix} I_{m^2} & 0 \\ 0 & \check{\mathscr{B}}_{N,k} \end{bmatrix}.$$

Hence,

$$\widehat{Q}_N = \frac{1}{N}\sum_{k=1}^{N}\|(I_m \otimes \mathscr{R}_k^{-1/2})\mathrm{vec}(\tilde{\mathscr{S}}_N(z_k))\|_F^2 = \check{\Gamma}^{\mathrm{T}}\widehat{\Xi}_N\check{\Gamma},$$

where

$$\widehat{\Xi}_N \triangleq \frac{1}{N}\sum_{k=1}^{N}\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m \otimes \mathscr{R}_k^{-1})\tilde{\mathscr{A}}_{N,k}.$$

From (C.6), note that $\widehat{\Xi}_N$ is positive definite for all large N w.p.1. Likewise, we can write $\widehat{T}_N$ as

$$\widehat{T}_N = \Upsilon_N\check{\Gamma},$$

where

$$\Upsilon_N \triangleq \frac{1}{N}\sum_{k=1}^{N}[\mathrm{vec}(S_k^*)]^{\mathrm{T}}(I_m \otimes \mathscr{R}_k^{-1})\tilde{\mathscr{A}}_{N,k} + \frac{1}{N}\sum_{k=1}^{N}[\mathrm{vec}(S_k)]^{\mathrm{T}}(I_m \otimes \mathscr{R}_k^{-1})\tilde{\mathscr{A}}_{N,k}^*.$$

Let $\widehat{\Gamma}_N$ denote the least-squares solution of (92) in stacked form

$$\widehat{\Gamma}_N \triangleq \begin{bmatrix} \mathrm{vec}(\widehat{E}_N) \\ \mathrm{vec}(\widehat{F}_N) \end{bmatrix}.$$

Then,

$$\widehat{\Gamma}_N = \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\frac{1}{N}\sum_{k=1}^{N}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m \otimes \mathscr{R}_k^{-1})\mathrm{vec}(S_k)\}.$$

Split $\widehat{\Gamma}_N$ as

$$\widehat{\Gamma}_N = \bar{\Gamma}_N + \tilde{\Gamma}_N,$$

where

$$\bar{\Gamma}_N \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\frac{1}{N}\sum_{k=1}^{N}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m \otimes \mathscr{R}_k^{-1})\mathrm{vec}(S(z_k))\},$$

$$\tilde{\Gamma}_N \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\frac{1}{N}\sum_{k=1}^{N}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m \otimes \mathscr{R}_k^{-1})\mathrm{vec}(\eta_k)\}.$$

Let Γ denote the unknowns in stacked form:

$$\Gamma \triangleq \begin{bmatrix} \mathrm{vec}(E) \\ \mathrm{vec}(F) \end{bmatrix}.$$

Observe that if $N \geqslant (p/2) + n + 1$ and $\eta_k = 0$ for all k, then $\widehat{\chi}(z) = \chi(z)$ for all z and from (C.7)we have

$$\mathrm{vec}(S(z_k)) = \check{\mathscr{A}}_{N,k}\Gamma = \tilde{\mathscr{A}}_{N,k}\Gamma \quad \text{for all } k.$$

Hence,

$$\widehat{\Gamma}_N = \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\frac{1}{N}\sum_{k=1}^{N}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m\otimes\mathscr{R}_k^{-1})\tilde{\mathscr{A}}_{N,k}\,\Gamma\}$$

$$= \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\mathrm{Re}\{\widehat{\Xi}_N\}\,\Gamma = \Gamma.$$

This proves Lemma 5. □

Now, from (91) we have uniformly in k

$$\lim_{N\to\infty}\|\tilde{\mathscr{A}}_{N,k}-\check{\mathscr{A}}_{N,k}\|_F, \quad \text{w.p.1.} \tag{C.8}$$

Recall that $\widehat{\Xi}_N^{-1}$ is bounded away from zero w.p.1 for all large N; and $\tilde{\mathscr{A}}_{N,k}$ is also uniformly bounded in k w.p.1 for all large N. Thus, the following series

$$\bar{\Gamma}_N = \Gamma + \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\frac{1}{N}\sum_{k=1}^{N}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m\otimes\mathscr{R}_k^{-1})$$

$$\cdot(\tilde{\mathscr{A}}_{N,k}-\check{\mathscr{A}}_{N,k})\}\Gamma$$

converges w.p.1 to Γ as N tends to infinity.

Finally, we study the noise term $\tilde{\Gamma}_N$. Let

$$c_{N,k} \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\mathrm{Re}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m\otimes\mathscr{R}_k^{-1})\},$$

$$d_{N,k} \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\mathrm{Im}\{\tilde{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m\otimes\mathscr{R}_k^{-1})\},$$

$$\xi_k \triangleq \mathrm{vec}(\mathrm{Re}(\eta_k)),$$

$$\zeta_k \triangleq \mathrm{vec}(\mathrm{Im}(\eta_k)).$$

Then, we can write $\tilde{\Gamma}_N$ as

$$\tilde{\Gamma}_N = \frac{1}{N}\sum_{k=1}^{N}c_{N,k}\xi_k + \frac{1}{N}\sum_{k=1}^{N}d_{N,k}\zeta_k.$$

Let us assume for a moment that $c_{N,k}$ and $d_{N,k}$ are bounded sequences of deterministic matrices denoted by $\bar{c}_{N,k}$ and $\bar{d}_{N,k}$. Then, $\bar{c}_{N,k}\xi_k$ and $\bar{d}_{N,k}\zeta_k$ are sequences of independent zero mean random variables with uniformly bounded fourth-order moments. Thus, from the strong law of large numbers (Chung, 1968) each series above tends to zero w.p.1 as N tends to infinity. Now, let

$$\check{\Xi}_N \triangleq \frac{1}{N}\sum_{k=1}^{N}\check{\mathscr{A}}_{N,k}^{\mathrm{H}}(I_m\otimes\mathscr{R}_k^{-1})\check{\mathscr{A}}_{N,k}$$

and

$$\bar{c}_{N,k} \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\mathrm{Re}\{\check{\mathscr{A}}_{N,k}^{\mathrm{T}}(I_m\otimes\mathscr{R}_k^{-1})\},$$

$$\bar{d}_{N,k} \triangleq \mathrm{Re}\{\widehat{\Xi}_N^{-1}\}\mathrm{Im}\{\check{\mathscr{A}}_{N,k}^{\mathrm{T}}(I_m\otimes\mathscr{R}_k^{-1})\}.$$

From (C.8),

$$\lim_{N\to\infty}\|\widehat{\Xi}_N-\check{\Xi}_N\|_F, \quad \text{w.p.1,}$$

and thus

$$\lim_{N\to\infty}\|c_N-\bar{c}_N\|_\infty = 0, \quad \text{w.p.1,}$$

$$\lim_{N\to\infty}\|d_N-\bar{d}_N\|_\infty = 0, \quad \text{w.p.1,} \tag{C.9}$$

where $\|c_N\|_\infty \triangleq \sup_{1\leqslant k\leqslant N}\sigma_{\max}(c_{N,k})$. The series

$$\tilde{\vartheta}_N = \frac{1}{N}\sum_{k=1}^{N}(c_{N,k}-\bar{c}_{N,k})\xi_k - \frac{1}{N}\sum_{k=1}^{N}(d_{N,k}-\bar{d}_{N,k})\zeta_k$$

is dominated (absolutely) by the series

$$\|c_N-\bar{c}_N\|_\infty\frac{1}{N}\sum_{k=1}^{N}\|\xi_k\|_2 + \|d_N-\bar{d}_N\|_\infty\frac{1}{N}\sum_{k=1}^{N}\|\zeta_k\|_2,$$

where $\|x\|_2$ is the Euclidean norm of $x\in\mathbf{R}^n$ defined by $\|x\|_2 \triangleq (\sum_{j=1}^{n}|x_j|^2)^{1/2}$. From the strong law of large numbers, we have

$$\lim_{N\to\infty}\left[\frac{1}{N}\sum_{k=1}^{N}\|\xi_k\|_2 - \frac{1}{N}\sum_{k=1}^{N}\mathbf{E}\|\xi_k\|_2\right] = 0, \quad \text{w.p.1,}$$

$$\lim_{N\to\infty}\left[\frac{1}{N}\sum_{k=1}^{N}\|\zeta_k\|_2 - \frac{1}{N}\sum_{k=1}^{N}\mathbf{E}\,\|\zeta_k\|_2\right] = 0, \quad \text{w.p.1.}$$

From (5), we have

$$\frac{1}{N}\sum_{k=1}^{N}\mathbf{E}\|\xi_k\|_2 = \frac{1}{N}\sum_{k=1}^{N}\mathbf{E}\,\|\mathrm{Re}\,\eta_k\|_F = \frac{1}{2N}\sum_{k=1}^{N}\mathrm{Tr}(\mathscr{R}_k)$$

$$\leqslant \frac{m}{2N}\sum_{k=1}^{N}\sigma_{\max}(\mathscr{R}_k) \leqslant \frac{m}{2}M_\eta^{1/2}.$$

Thus,

$$\lim_{N\to\infty}\sup\frac{1}{N}\sum_{k=1}^{N}\|\xi_k\|_2 \leqslant \frac{m}{2}M_\eta^{1/2}, \quad \text{w.p.1.}$$

Likewise,

$$\lim_{N\to\infty}\sup\frac{1}{N}\sum_{k=1}^{N}\|\zeta_k\|_2 \leqslant \frac{m}{2}M_\eta^{1/2}, \quad \text{w.p.1.}$$

Hence, from (C.9) $\tilde{\vartheta}_N$ converges to zero w.p.1 as N tends to infinity; and therefore $\tilde{\Gamma}_N$ converges to zero w.p.1 as N tends to infinity. It follows that $\widehat{\Gamma}_N$ converges to Γ w.p.1 as N tends to infinity.

References

Caines, P. (1988). *Linear stochastic systems*, *Series in probability and mathematical statistics*. New York: Wiley.

Chung, K. L. (1968). *A course in probability theory*. New York: Harcourt, Brace & World.

Dodds, C. J., & Robson, J. D. (1973). The description of road surface roughness. *Journal of Sound and Vibration*, *31*, 175–183.

Fraanje, R., Verhaegen, M., Verdult, V., & Pintelon, R. (2003). A frequency domain subspace algorithm for mixed causal, anti-causal LTI systems. In *Proceedings CD-ROM of the 13th IFAC symposium on system identification*, Rotterdam, The Netherlands, August 2003 (pp. 893–898).

Glover, K., & Willems, J. (1974). Parametrization of linear dynamical systems: Canonical forms and identifiability. *IEEE Transactions on Automatic Control*, *AC-19*, 640–645.

Guidorzi, R. (1974). Canonical structures in the identification of multivariable systems. *Automatica*, *11*, 361–374.

Guidorzi, R. (1981). Invariants and canonical forms for systems structural and parametric identification. *Automatica*, *17*, 117–133.

Kay, S. (1988). *Modern spectral estimation, theory & application*. Englewood Cliffs, NJ: Prentice-Hall.

Larimore, W. E. (1990). Canonical variate analysis in identification, filtering and adaptive control. In *Proceedings of the 29th conference on decision and control*, Hawai, USA (pp. 596–604).

Liu, K., Jacques, R. N., & Miller, D. W. (1994). Frequency domain structural system identification by observability range space extraction. In *Proceedings of the American control conference*, Vol. 1, Baltimore, MD (pp. 107–111).

Ljung, L. (2000). *System identification: Theory for the user* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

McKelvey, T., Akçay, H., & Ljung, L. (1996a). Subspace-based multivariable system identification from frequency response data. *IEEE Transactions on Automatic Control*, *41*, 960–979.

McKelvey, T., Akçay, H., & Ljung, L. (1996b). Subspace-based identification of infinite-dimensional multivariable systems from frequency response data. *Automatica*, *32*, 885–902.

Moor, B. D. (1993). The singular value decomposition and long and short spaces of noisy matrices. *IEEE Transactions on Signal Processing*, *41*, 2826–2838.

Van Overbeek, A. J. M., & Ljung, L. (1982). On-line structure selection for multivariable state space models. *Automatica*, *18*, 529–543.

Van Overschee, P., & De Moor, B. (1996a). *Subspace identification for linear systems: Theory—implementation—applications*. Dordrecht: Kluwer Academic Publishers.

Van Overschee, P., & De Moor, B. (1996b). Continuous-time frequency domain subspace identification (special issue on subspace methods for detection and estimation). *Signal Processing*, *52*, 179–194.

Van Overschee, P., De Moor, B., Dehandschutter, W., & Swevers, J. (1997). A subspace algorithm for the identification of discrete time frequency domain power spectra. *Automatica*, *33*, 2147–2157.

Priestley, M. B. (1989). *Spectral analysis and time series*. London, England: Academic Press.

Schoukens, J., & Pintelon, R. (1991). *Identification of linear systems: A practical guideline to accurate modeling*. London: Pergamon.

Stoica, P., & Moses, R. (1997). *Introduction to spectral analysis*. Upper Saddle River, NJ: Prentice-Hall.

Türkay, S., & Akçay, H. (2004). A study of random vibration characteristics of the quarter-car model. *First IFAC symposium on advances in automotive control*, Salerno, Italy, to appear.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data (special issue on statistical signal processing and control). *Automatica*, *30*, 61–74.

Verhaegen, M. (1996). A subspace model identification solution to the identification of mixed causal, anti-causal LTI systems. *SIAM Journal on Matrix Analysis and Applications*, *17*, 332–347.

Viberg, M. (1995). Subspace methods in system identification (special issue on trends in system identification). *Automatica*, *31*, 1835–1852.

Hüseyin Akçay was born in Antalya, Turkey in 1958. He received the Engineer degree from the Istanbul Technical University in 1981, the M.Sc. degree from the Massachusetts Institute of Technology in 1988, and the Ph.D. degree from the University of Michigan, Ann Arbor in 1992, all in Mechanical Engineering, and the M.A. degree in Mathematics from the University of Michigan, Ann Arbor in 1991. He held visiting positions with Linköping, Newcastle, and Bremen Universities. He received a research fellowship award in 1998 from the Alexander Von Humbold Foundation. He worked at the Tübitak, Marmara Research Center, Gebze, Turkey as Research Scientist. He is currently Professor of Electrical and Electronics Engineering in Anadolu University. His research interests include system identification, signal processing, condition monitoring and fault detection.

Semiha Türkay was born in Kircali, Bulgaria, in 1979. She received the B.S. degree from Osmangazi University and the M.S. degree from the Anadolu University, Turkey, in 2001 and 2003, respectively, both in Electrical and Electronics Engineering.

Since 2001 she has been working as a Research Assistant at Anadolu University and is currently a Ph.D. student. Her research interests include random vibrations, estimation, system identification and control.