

Modeling continuous-time processes via input-to-state filters

Kaushik Mahata and Minyue Fu

Centre for Complex Dynamic Systems and Control,
School of Electrical Engineering and Computer Science,
University of Newcastle, Callaghan, NSW 2308, Australia

Email: Kaushik.Mahata@newcastle.edu.au, Minyue.Fu@newcastle.edu.au

version August 9, 2005

Abstract

A direct algorithm to estimate continuous-time ARMA (CARMA) models is proposed in this paper. In this approach we first pass the observed data through an input-to-state filter and compute the state covariance matrix. The properties of the state covariance matrix are then exploited to estimate the half-spectrum of the observed data at a set of user-defined points on the right-half plane. Finally, the continuous-time parameters are obtained from the half-spectrum estimates by solving an analytic interpolation problem with a positive real constraint. As shown by simulations, the proposed algorithm delivers much more reliable estimates than indirect modeling approaches, which rely on estimating an intermediate discrete-time model.

Keywords: Continuous-time processes, identification, input-to-state filtering, ARMA modeling.

1 Introduction

The identification of continuous-time stochastic processes is a fundamental research issue which has received considerable interest recently. Since many natural processes are continuous-time, it is of interest in many practical applications to identify a continuous-time model instead of a discrete-time model [15]. Although the signal is in continuous-time, in practice one works with sampled data. One popular approach is to identify a discrete-time system from uniformly sampled data, as shown in [15] and references therein. Subsequently, the estimated discrete-time model is converted back to a continuous-time model via a nonlinear transformation [24]. This approach will be referred to as the *indirect* approach. Apart from the obvious difficulty of solving nonlinear equations, this approach also suffers from several other setbacks: (i) At a fast sampling rate, the poles and the zeros of the associated discrete-time system cluster close to the point $1 + i0$ in the complex plane, leading to a numerically ill-conditioned identification problem; (ii) The continuous-time parameters can be very sensitive to the sampled data. These issues have been discussed in detail in [4, 3, 26, 17, 16, 24]. A second approach is to identify the continuous-time parameter *directly*. The basic idea here is to replace the differentiation operator with the *delta operator* [5, 11]. Several methods have been developed using this approach for autoregressive models, see [4, 3, 26, 17]. This approach is advantageous in many cases as it is computationally efficient and one avoids nonlinear transformations if the underlying model is autoregressive. This technique also benefits from non-uniform sampling [17, 15]. However, it is not well understood how we can extend this technique for an ARMA model since the mapping from the continuous-time zeros to the equivalent discrete-time zeros is complicated. It is also not known how to guarantee the positivity of the estimated spectrum. Therefore, for ARMA models, the only available time domain approach seems to be the indirect method described above [16, 24]. A frequency domain method is proposed in [23].

If we consider a discrete-time process, we can ensure the stability of the estimated autoregressive model in a fairly simple way [20, 27, 18]. But unlike the discrete-time counterpart, the mapping from the lagged covariance estimates to the system parameters for a continuous-time system is more complicated. Hence, the standard discrete-time algorithms cannot be extended directly. In this paper, we propose a direct approach for modeling continuous-time stochastic processes. Specifically, we provide an estimation algorithm with the following properties: (i) It is computationally efficient; (ii) The stability of the estimated model is guaranteed; (iii) It can handle irregularly sampled data; (iv) It is possible to circumvent the problems associated with the *sampling zeros* for an ARMA model.

Our approach uses the framework of input-to-state filtering [8, 10, 9] where we first estimate the half-spectrum and its derivatives evaluated at some pre-specified points in the right-half plane. This is achieved using a linear operation on the covariance matrix of the output of an input-to-state filter. Subsequently, we present an approach for estimating a stable rational model from the estimates of the half-spectrum. In this step, we apply linear interpolation combined with a regularization step similar to that proposed in [28, 21]. If the resulting model is unstable then a recent spectral zero assignment algorithm [1], (see also [7]), is used to compute a stable model. The approach can also be used when the data are irregularly sampled. We provide additional insights in the estimation of the half-spectrum, and discuss the numerical and statistical issues involved in the estimation of such statistics. We also carry out an asymptotic statistical analysis of second-order. The proposed algorithm is tested using numerical simulations.

2 Input-to-state filters

The discussion in this section is applicable to any continuous-time wide-sense stationary stochastic process $u(t)$. As a special application, we apply the results derived in this section to continuous-time ARMA processes in later sections. Analogous results for discrete-time processes have been derived in [8, Theorem 1], see also [10, Corollary 1] and [9]. The results mentioned above are somewhat more general than what we require for this paper. The proofs in the aforementioned references have been extended to the continuous-time case [9, Section V].

Consider a *scalar* and real-valued¹ continuous-time stochastic process $u(t)$ having an autocorrelation function

$$r_\tau := \mathcal{E} \{u(t + \tau)u(t)\}.$$

Then, the spectrum of the process is defined as

$$\phi(s) := \int_{-\infty}^{\infty} d\tau r_\tau e^{-s\tau}, \quad \text{Re}(s) = 0.$$

In this work we use the so-called *half-spectrum* $f(s)$ of $u(t)$, which is defined as

$$f(s) := \int_0^{\infty} d\tau r_\tau e^{-s\tau}, \quad \text{Re}(s) \geq 0.$$

Consequently, it is readily verified that

$$\phi(s) = f(s) + f(-s), \quad \text{Re}(s) = 0. \quad (2.1)$$

In the following we estimate $f(s)$ and its derivatives at a predefined set of points $\{s_k\}_{k=1}^m$ from the observed continuous-time signal $u(t)$. The points $\{s_k\}_{k=1}^m$ are referred to as the *interpolation points* and satisfy $\text{Re}(s_k) > 0, \forall k$. The main idea here is to use an input-to-state filter

$$\dot{z}(t) = Fz(t) + gu(t), \quad (2.2)$$

where F has eigenvalues at $\{-s_k\}_{k=1}^m$ and the pair (F, g) is controllable. We assume that the filter in (2.2) has a pole of order n_k at $-s_k$, while the order of the filter is n , *i.e.*, F is a $n \times n$ matrix, and $\sum_{k=1}^m n_k = n$. We show that the covariance matrix of the output $z(t)$ can be used to extract the estimates of $f(s_k)$ and its derivatives. In particular, a pole of order n_k at $-s_k$ enables us to extract the derivatives of $f(s)$ up to order $n_k - 1$ evaluated at s_k . The following proposition is the first step in that direction.

Proposition 1. Assume that $f(\infty)$ is bounded. Let E be the unique positive definite solution to the Lyapunov equation

$$FE + EF' + gg' = 0. \quad (2.3)$$

Then there exist scalar-valued functions $\{w_k\}_{k=0}^{n-1}$ of F and $f(s)$ such that

$$P := \mathcal{E} \{z(t)z'(t)\} = WE + EW', \quad (2.4)$$

where

$$W := \frac{1}{2\pi i} \oint_{C_R} ds f(-s) [sI - F]^{-1} = \sum_{k=0}^{n-1} w_k F^k. \quad (2.5)$$

¹However, the results in this section can be generalized for a complex-valued stochastic process in a fairly straightforward manner.

with C_R being the infinite semicircular contour encircling the entire right-half plane traversed in the *clockwise* direction. Moreover, the coefficients $\{w_k\}_{k=0}^{n-1}$ are invariant of the choice of the coordinates of $z(t)$.

Proof: The proof of (2.4) and (2.5) is similar to Theorem 1 in [8]; see also [19] for a more direct proof. To show the invariance of $\{w_k\}_{k=0}^{n-1}$, consider the state sequence $z_1(t)$ of the input-to-state filter (F_1, g_1) , where

$$F_1 = LFL^{-1}, \quad g_1 = Lg$$

for some nonsingular matrix L . Let the covariance matrix of $z_1(t)$ be P_1 . It follows that $P_1 = LPL'$ and the solution E_1 to the Lyapunov equation $F_1E_1 + E_1F_1' + g_1g_1' = 0$ satisfy $E_1 = LEL'$. Then (2.4) gives

$$P_1 = W_1E_1 + E_1W_1',$$

where $W_1 = LWL^{-1}$. Now from (2.5) we can verify our assertion that $W_1 = \sum_{k=0}^{n-1} w_k F_1^k$. ■

Next, we state a method for computing $\{f(s_k)\}_{k=1}^m$ and its derivatives.

Theorem 1. Let F be chosen such that its Jordan form has the block-diagonal structure

$$\text{diag} \left\{ J(-s_1, n_1), \quad \dots, \quad J(-s_m, n_m) \right\}, \quad (2.6)$$

where for each $k \in \{1, 2, \dots, m\}$, the matrix $J(-s_k, n_k)$ is an elementary $n_k \times n_k$ Jordan block having $-s_k$ along the main diagonal, *ones* along the first upper subdiagonal, and *zeros* elsewhere. Define the polynomial

$$\varpi(s) := \sum_{k=0}^{n-1} w_k s^k.$$

Denote the r th derivative of $f(s)$ evaluated at $s = s_k$ by $f^{(r)}(s_k)$. Then

$$f^{(r)}(s_k) = (-1)^r \varpi^{(r)}(-s_k), \quad 0 \leq r < n_k, \quad 1 \leq k \leq m. \quad (2.7)$$

Proof: The proof is similar to Theorem 2 in [8]; see also [19] for a direct proof. ■

Remarks:

1. If we allow multiple Jordan blocks of F with the same eigenvalue $-s_k$, then the information present in the associated Jordan blocks is redundant, and we are unable to extract n pieces of statistics about $f(s)$. See the proof in [19] for details.
2. Extracting statistics from the data for modeling discrete-time processes is often accomplished by estimating the covariances of the observed data. The theory of input-to-state filtering is not necessary in this development [18]. However, Theorem 1 is vital for continuous-time processes. ■

Next we express the vector of $f^{(j)}(s_k)$ as a linear function of the coefficients $\{w_k\}_{k=0}^{n-1}$. Define

$$\mathbf{f}_k := \begin{bmatrix} f^{(0)}(s_k) & \dots & f^{(n_k-1)}(s_k) \end{bmatrix}', \quad \mathbf{w} := \begin{bmatrix} w_0 & \dots & w_{n-1} \end{bmatrix}'. \quad (2.8)$$

From (2.7) we get $\mathbf{f}_k = \mathfrak{D}_k w$ for $k = 1, \dots, m$, where \mathfrak{D}_k is a $n_k \times n$ matrix defined as

$$[\mathfrak{D}_k]_{ij} = (-1)^{i-1} \left[\frac{d^{i-1}}{ds^{i-1}} \{s^{j-1}\} \right]_{s=-s_k}.$$

Real-valued implementation of the input-to-state filter requires a self-conjugate set of interpolation points. However, if s_j is the complex conjugate of s_k , then \mathbf{f}_j is the complex conjugate of \mathbf{f}_k . Hence it is sufficient to consider the interpolation points having nonnegative imaginary parts. Let the number of real-valued interpolation points be n_r , and the number of interpolation points with strictly positive imaginary parts be n_c . Clearly, $m = n_r + 2n_c$. Without loss of generality, we assume that $\{s_k\}_{k=1}^{n_c}$ are real-valued, and $\{s_k\}_{k=n_r+1}^{n_r+n_c}$ have strictly positive imaginary parts. Define the vectors

$$\mathbf{f}_R := [\mathbf{f}'_1 \ \cdots \ \mathbf{f}'_{n_r}]', \quad \mathbf{f}_C := [\mathbf{f}'_{n_r+1} \ \cdots \ \mathbf{f}'_{n_r+n_c}]', \quad \mathbf{f} := [\mathbf{f}'_R \ \text{Re}(\mathbf{f}'_C) \ \text{Im}(\mathbf{f}'_C)]. \quad (2.9)$$

From the above discussion it follows that

$$\begin{aligned} \mathbf{f} &= \mathbf{D}w, \\ \mathbf{D} &= [\mathfrak{D}'_1 \ \cdots \ \mathfrak{D}'_{n_r} \ \text{Re}(\mathfrak{D}'_{n_r+1}) \ \cdots \ \text{Re}(\mathfrak{D}'_{n_r+n_c}) \ \text{Im}(\mathfrak{D}'_{n_r+1}) \ \cdots \ \text{Im}(\mathfrak{D}'_{n_r+n_c})]'. \end{aligned} \quad (2.10)$$

Using the results derived so far, we propose an algorithm below for estimating of \mathbf{f} . The justification behind the algorithm will be given in Section 3. In the algorithm we implement the input-to-state filter in the controllable canonical form. Let the characteristic polynomial of F be

$$\Delta(s) := \prod_{k=1}^n (s + s_k) = s^n + \sum_{k=1}^n \delta_k s^{n-k}. \quad (2.11)$$

The state space matrices for the input-to-state filter in the controllable canonical form are then given by

$$F_\star = \begin{bmatrix} -\delta_1 & \cdots & -\delta_{n-1} & -\delta_n \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \quad g_\star = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The associated state vector and the state covariance are denoted by $z_\star(t)$ and P_\star , respectively:

$$\dot{z}_\star(t) = F_\star z_\star(t) + g_\star u(t), \quad P_\star := \mathcal{E} \{z_\star(t) z_\star'(t)\}. \quad (2.12)$$

From (2.4) and (2.5) we have

$$P_\star = \sum_{k=0}^{n-1} w_k \{F_\star^k E_\star + E_\star (F_\star')^k\}, \quad (2.13)$$

where E_\star is the solution to the Lyapunov equation

$$F_\star E_\star + E_\star F_\star' + g_\star g_\star' = 0. \quad (2.14)$$

Therefore, computing w from P_\star amounts to solving a least-squares problem. In Section 3 we show that it is enough to consider only the diagonal elements in the matrix-valued equation (2.13). In the following for a $n \times n$ real-valued matrix H we denote

$$\mathfrak{d}[H] = [\ [H]_{1,1} \ \cdots \ [H]_{n,n} \]'.$$

Considering only the diagonal elements of (2.13) we get

$$\mathfrak{d}[P_\star] = X_\star w, \quad X_\star := \begin{bmatrix} \mathfrak{d}[2E_\star] & \mathfrak{d}[F_\star E_\star + E_\star F_\star'] & \cdots & \mathfrak{d}[F_\star^{n-1} E_\star + E_\star (F_\star')^{n-1}] \end{bmatrix}. \quad (2.15)$$

Since F_\star and g_\star being user defined, X_\star is assumed to be known in the following algorithm. We assume further that the observed process $u(t)$ is known for $t \in [0, T]$.

Algorithm.

1. Compute the state output $z_\star(t)$ in (2.12), and obtain an estimate of P_\star by computing

$$\hat{P}_\star := \frac{1}{T} \int_0^T dt z_\star(t) z_\star'(t). \quad (2.16)$$

2. Compute the estimates of w and \mathbf{f} as

$$\hat{w} = X_\star^{-1} \mathfrak{d}[\hat{P}_\star], \quad \hat{\mathbf{f}} = \mathbf{D} \hat{w}. \quad (2.17)$$

From a practical point of view one might prefer to implement (2.16) in discrete-time, where the integral is approximated by a Riemannian summation

$$\check{P}_\star = \frac{1}{N} \sum_{k=0}^{N-1} z_\star(k t_s) z_\star'(k t_s),$$

where t_s is the *sampling interval* and $T = N t_s$. In fact, \check{P}_\star is a consistent estimate of P_\star . Furthermore, if t_s is sufficiently small then \check{P}_\star and \hat{P}_\star have similar statistical properties. Taking $t_s \rightarrow 0$ does increase N , but the accuracy of \check{P}_\star remains the same. This is because $z_\star(k t_s)$ is highly correlated with its neighboring samples when $t_s \rightarrow 0$, and an increase in N does not influence the estimation accuracy [29].

It is often assumed that the observed data are the sampled version of the continuous-time process. Then it is required to implement the input-to-state filter (2.12) using a discretization technique [22, p.33],[6]. The systematic errors introduced due to such approximation are not significant [6] for practical values of t_s . Analysis of such systematic errors is beyond the scope of this paper. However, if the continuous-time signal is available to the user, it is possible to avoid systematic errors by implementing (2.12) using analog devices for which several efficient architectures are readily available [13, p.35]. The output $z_\star(t)$ from the input-to-state filter can then be sampled for the computation of \hat{P}_\star .

3 Some computational and statistical aspects

In this section we focus on the computational and the statistical aspects involved in the estimation of the half-spectrum. The primary aim is to justify the algorithm for computing \hat{w} and $\hat{\mathbf{f}}$ proposed in Section 2. First we examine the rank of the system of equations to be solved in order to determine \hat{w} . We also comment on the choice of the coordinates of the state of the input-to-state filter. Finally, we determine the second-order statistics of \hat{w} and $\hat{\mathbf{f}}$. To this end we need some additional notation. Define

$$\begin{aligned} \mathbb{H} &:= \{H \in \mathbb{R}^{n \times n} : [H]_{i,j} = (-1)^j h_{i+j-1} \text{ for some real } \{h_k\}_{k=1}^{2n-1}\} \\ \mathbb{H}_0 &:= \{H \in \mathbb{H} : [H]_{i,j} = 0 \text{ when } i+j \text{ is odd}\}. \end{aligned}$$

A key property of the set \mathbb{H}_0 is that if $H \in \mathbb{H}_0$ and $i + j = 2\ell$ for some positive integer $\ell \leq n$, then

$$[H]_{i,j} = (-1)^{j-\ell} [H]_{\ell,\ell}.$$

It is also straightforward to verify that $H \in \mathbb{H}$ implies $H + H' \in \mathbb{H}_0$.

Proposition 2. The matrix P_\star defined in (2.12) is a member of \mathbb{H}_0 .

Proof: Denote the differentiation operator by p . Since (F_\star, g_\star) is in the controllable canonical form we have

$$z_\star(t) = [z_1(t) \ \cdots \ z_n(t)]', \quad z_j(t) := \frac{p^{n-j}}{\Delta(p)} u(t),$$

where the polynomial $\Delta(\cdot)$ is defined in (2.11). By the definition of P_\star we get

$$[P_\star]_{jk} = \mathcal{E} \{z_j(t) z_k(t)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \phi(i\omega) \frac{(i\omega)^{n-j} (-i\omega)^{n-k}}{|\Delta(i\omega)|^2}. \quad (3.1)$$

Note that the spectrum $\phi(\omega)$ is an even function of ω . Therefore, if $j + k$ is an odd integer the integral vanishes because the integrand is an odd function of ω . If $j + k = 2\ell$ for some integer ℓ , then we have

$$(i\omega)^{n-j} (-i\omega)^{n-k} = \omega^{2n-2\ell} (i)^{2n-2\ell} (-1)^{n-k} = \omega^{2n-2\ell} (-1)^{k-\ell}.$$

Consequently, it follows from (3.1) that

$$\begin{aligned} [P_\star]_{jk} &= \frac{(-1)^{k-\ell}}{2\pi} \int_{-\infty}^{\infty} d\omega \frac{\omega^{2n-2\ell} \phi(i\omega)}{|\Delta(i\omega)|^2} = (-1)^{k-\ell} \mathcal{E} \left\{ \frac{p^{n-\ell}}{\Delta(p)} u(t) \cdot \frac{p^{n-\ell}}{\Delta(p)} u(t) \right\} \\ &= (-1)^{k-\ell} [P_\star]_{\ell\ell}. \end{aligned}$$

From the last equality it is straightforward to verify that $P_\star \in \mathbb{H}_0$. ■

Remark: If $u(t)$ is a continuous-time white noise (*i.e.* derivative of the Weiner process) of unit variance, then $P_\star = E_\star$; see (2.14). Consequently, $E_\star \in \mathbb{H}_0$. ■

In the following proposition we show that the matrices involved in (2.13) belong to \mathbb{H}_0 .

Proposition 3. The matrix $X_k := F_\star^k E_\star \in \mathbb{H}$ for any positive integer k .

Proof: In this proof we use the Matlab notation $H(j : k, :)$, $j \leq k$ to denote the sub-matrix of H composed of the j th to k th rows of H . Similarly, $H(:, j : k)$, $j \leq k$ denotes the sub-matrix of H composed of the j th to k th columns of H . From the structure of F_\star we have for $k \geq 0$ that

$$X_{k+1}(2 : n, :) = F_\star(2 : n, :) X_k = [I_{n-1} \ 0_{(n-1) \times 1}] X_k = X_k(1 : n-1, :). \quad (3.2)$$

Next, we show using mathematical induction that

$$X_{k+1}(:, 2 : n) = -X_k(:, 1 : n-1). \quad (3.3)$$

Consider the case $k = 0$. Note that $X_0 = E_\star$ is a symmetric matrix. Also,

$$\begin{aligned} X_1 + X_1' &= -g_\star g_\star' \Rightarrow \\ [X_1(:, 2 : n)]' &= -X_1(2 : n, :) = -X_0(1 : n-1, :) = -[X_0(:, 1 : n-1)]', \end{aligned}$$

where in the third equality we have used (3.2). Therefore, (3.3) holds for $k = 0$. Now assume that (3.3) holds for $k = 0, 1, \dots, \ell - 1$. Then

$$X_{\ell+1}(:, 2 : n) = F_\star X_\ell(:, 2 : n) = -F_\star X_{\ell-1}(:, 1 : n - 1) = -X_\ell(:, 1 : n - 1).$$

Hence, (3.3) follows by induction. Now $X_0 = E_\star \in \mathbb{H}$. Applying (3.2) and (3.3) recursively it is readily verified that $X_k \in \mathbb{H}$ for all $k \geq 0$. Hence the proposition follows. \blacksquare

It follows from Proposition 3 that $X_k + X'_k \in \mathbb{H}_0$ for all k . Since $P_\star \in \mathbb{H}_0$, the number of independent equations in the matrix equation (2.13) is only n . If $z(t) = Lz_\star(t)$ for some nonsingular L , then using (2.13) and Proposition 1 we get

$$LP_\star L' = \sum_{k=0}^{n-1} w_k L(X_k + X'_k)L' \quad \Rightarrow \quad [L \otimes L] \text{vec}[P_\star] = \sum_{k=0}^{n-1} w_k [L \otimes L] \text{vec}[X_k + X'_k]. \quad (3.4)$$

This is a weighted version of (2.13). Note that we denote the matrix Kronecker product by \otimes . Since the underlying system of equations is not overdetermined, the weighting has no effect on the final solution. However, in a practical scenario, the true covariance matrix P_\star is unknown, and we work with an estimate \hat{P}_\star not necessarily belonging to \mathbb{H}_0 . Then we might expect that solving an overdetermined system with proper weighting (3.4) may give more accurate estimates of $\{w_k\}_{k=0}^{n-1}$. We explore the second-order statistical properties of \hat{P}_\star to examine this aspect. We first state the following basic result.

Proposition 4. Let $u_A(t)$, $u_B(t)$, $u_C(t)$ and $u_D(t)$ be jointly Gaussian real and scalar-valued continuous-time stationary stochastic processes. Let us define

$$\hat{P}_{AB}(\tau) := \frac{1}{T} \int_0^T dt u_A(t + \tau) u_B(t), \quad P_{AB}(\tau) := \mathcal{E} \{u_A(t + \tau) u_B(t)\}.$$

The cross-spectrum $\Phi_{AB}(i\omega)$ of $u_A(t)$ and $u_B(t)$ is given by

$$\Phi_{AB}(i\omega) = \int_{-\infty}^{\infty} d\tau P_{AB}(\tau) e^{-i\omega\tau}.$$

Assume that each of the functions $P_{AC}(\tau)P_{BD}(\tau)$ and $P_{AD}(\tau)P_{BC}(\tau)$ tends to zero as $\tau \rightarrow \infty$ at a rate faster than τ^{-2} . Then as $T \rightarrow \infty$ the asymptotic covariance between $\hat{P}_{AB}(0)$ and $\hat{P}_{CD}(0)$ is given by

$$\mathcal{E} \{ \hat{P}_{AB}(0) - P_{AB}(0) \} \{ \hat{P}_{CD}(0) - P_{CD}(0) \} = \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega [\Phi_{AC}(i\omega)\Phi_{DB}(i\omega) + \Phi_{AD}(i\omega)\Phi_{CB}(i\omega)].$$

Proof: By straightforward algebra we have

$$\begin{aligned} & \mathcal{E} \{ \hat{P}_{AB}(0) - P_{AB}(0) \} \{ \hat{P}_{CD}(0) - P_{CD}(0) \} = \mathcal{E} \{ \hat{P}_{AB}(0) \hat{P}_{CD}(0) \} - P_{AB}(0) P_{CD}(0) \\ &= \frac{1}{T^2} \int_0^T \int_0^T dt_1 dt_2 [\mathcal{E} \{ u_A(t_1) u_B(t_1) u_C(t_2) u_D(t_2) \} - \mathcal{E} \{ u_A(t_1) u_B(t_1) \} \mathcal{E} \{ u_C(t_2) u_D(t_2) \}] \\ &= \frac{1}{T^2} \int_0^T \int_0^T dt_1 dt_2 [\mathcal{E} \{ u_A(t_1) u_C(t_2) \} \mathcal{E} \{ u_B(t_1) u_D(t_2) \} + \mathcal{E} \{ u_A(t_1) u_D(t_2) \} \mathcal{E} \{ u_B(t_1) u_C(t_2) \}] \\ &= \frac{1}{T^2} \int_0^T \int_0^T dt_1 dt_2 [P_{AC}(t_1 - t_2) P_{BD}(t_1 - t_2) + P_{AD}(t_1 - t_2) P_{BC}(t_1 - t_2)] \\ &= \frac{1}{T^2} \int_{-T}^T dt (T - t) [P_{AC}(t) P_{BD}(t) + P_{AD}(t) P_{BC}(t)], \end{aligned}$$

where in the second equality we have used a well-known formula for the fourth order moment of a jointly Gaussian random variables [12]. Since the functions $P_{AC}(\tau)P_{BD}(\tau)$ and $P_{AD}(\tau)P_{BC}(\tau)$ tends to zero as $\tau \rightarrow \infty$ at a rate faster than τ^{-2} , we can write as $T \rightarrow \infty$

$$\begin{aligned} & \mathcal{E} \{ \hat{P}_{AB}(0) - P_{AB}(0) \} \{ \hat{P}_{CD}(0) - P_{CD}(0) \} \\ &= \frac{1/T}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt d\omega_1 d\omega_2 e^{i(\omega_1 + \omega_2)t} [\Phi_{AC}(i\omega_1)\Phi_{BD}(i\omega_2) + \Phi_{AD}(i\omega_1)\Phi_{BC}(i\omega_2)] \\ &= \frac{1}{2\pi T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\omega_1 d\omega_2 \delta(\omega_1 + \omega_2) [\Phi_{AC}(i\omega_1)\Phi_{BD}(i\omega_2) + \Phi_{AD}(i\omega_1)\Phi_{BC}(i\omega_2)] \\ &= \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega [\Phi_{AC}(i\omega)\Phi_{BD}(-i\omega) + \Phi_{AD}(i\omega)\Phi_{BC}(-i\omega)]. \end{aligned}$$

Hence, the proposition follows. ■

Since $P_{\star} \in \mathbb{H}_0$, each element along an anti-subdiagonal of \hat{P}_{\star} gives the estimates of the same quantity (up to a sign factor). Our next proposition explores the correlation structure of the elements in $\{[\hat{P}_{\star}]_{ij} : i + j = 2\ell\}$ for a fixed ℓ .

Proposition 5. Let \hat{P}_{\star} be given by (2.16). Then for $i + j = k + l = 2\ell$ it holds that

$$\text{cov} \left\{ [\hat{P}_{\star}]_{ij}, [\hat{P}_{\star}]_{kl} \right\} = (-1)^{j-l} \text{cov} \left\{ [\hat{P}_{\star}]_{\ell\ell}, [\hat{P}_{\star}]_{\ell\ell} \right\}, \quad \text{as } T \rightarrow \infty, \quad (3.5)$$

where $\text{cov}(x_1, x_2)$ denotes the covariance between two random variables x_1 and x_2 . Furthermore, for the diagonal elements of \hat{P}_{\star} it holds for $T \rightarrow \infty$ that

$$\text{cov} \left\{ [\hat{P}_{\star}]_{ii}, [\hat{P}_{\star}]_{jj} \right\} = \frac{2}{T} \text{var} \left\{ \frac{p^{2n-i-j}}{\Delta^2(p)} u_s(t) \right\}, \quad (3.6)$$

where $u_s(t)$ is a stationary stochastic process having a spectral density $\Phi_u^2(i\omega)$.

Proof: Recall that $[z_{\star}(t)]_i = \{p^{n-i}/\Delta(p)\}u(t)$, where p denotes the differentiation operator. Consider the case $i + j = k + l = 2\ell$. Then $(-1)^k = (-1)^{2\ell-l} = (-1)^{-l} = (-1)^l$. Proposition 4 gives

$$\begin{aligned} \text{cov} \left\{ [\hat{P}_{\star}]_{ij}, [\hat{P}_{\star}]_{kl} \right\} &= \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega \frac{|\Phi_u(i\omega)|^2}{|\Delta(i\omega)|^4} [(i\omega)^{2n-i-l}(-i\omega)^{2n-j-k} + (i\omega)^{2n-i-k}(-i\omega)^{2n-j-l}] \\ &= \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega \frac{|\Phi_u(i\omega)|^2}{|\Delta(i\omega)|^4} (i\omega)^{4n-i-j-k-l} [(-1)^{j+k} + (-1)^{j+l}] \\ &= \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega \frac{|\Phi_u(i\omega)|^2}{|\Delta(i\omega)|^4} (i\omega)^{4(n-\ell)} [(-1)^{j-l} + (-1)^{j-l}] \\ &= \frac{2(-1)^{j-l}}{2\pi T} \int_{-\infty}^{\infty} d\omega \frac{\omega^{4(n-\ell)} |\Phi_u(i\omega)|^2}{|\Delta(i\omega)|^4} \end{aligned} \quad (3.7)$$

Now putting $j = l = \ell$ in (3.7) we get (3.5). The calculation leading to (3.6) is similar *i.e.*,

$$\text{cov} \left\{ [\hat{P}_{\star}]_{ii}, [\hat{P}_{\star}]_{jj} \right\} = \frac{1}{2\pi T} \int_{-\infty}^{\infty} d\omega \frac{|\Phi_u(i\omega)|^2}{|\Delta(i\omega)|^4} [(i\omega)^{2n-i-j}(-i\omega)^{2n-i-j} + (i\omega)^{2n-i-j}(-i\omega)^{2n-i-j}],$$

which gives (3.6). ■

Since $P_\star \in \mathbb{H}_0$, every element of the set $\mathbb{S}_\ell := \{(-1)^{k-\ell}[\hat{P}_\star]_{j,k} : j+k=2\ell\}$ gives an estimate of $[P_\star]_{\ell,\ell}$. However, from (3.5) it follows that the asymptotic covariance matrix of the elements in \mathbb{S}_ℓ is of rank one. Hence, it is not possible to improve the statistical accuracy of estimated $[P_\star]_{\ell,\ell}$ by taking a linear combination of the elements in \mathbb{S}_ℓ . In practice T must be large enough to ensure the reliability of the extracted statistics. For such practical values of T the inference made from the asymptotic analysis is still valid. Numerical simulation results also confirm this observation. Thus, we conclude that there is no improvement in statistical accuracy if we solve the overdetermined system of equations (3.4), and the coordinates of $z(t)$ have no effect on the statistical accuracy of \hat{w} . The observations so far justify the algorithm described in Section 2.

Proposition 6. As the observation time $T \rightarrow \infty$ the asymptotic covariance matrix $C_{\hat{w}}$ of $\sqrt{T}\hat{w}$ in (2.17) is given by

$$C_{\hat{w}} = X^{-1}C_P(X')^{-1}, \quad (3.8)$$

where \hat{P}_\star is given in (2.16) and C_P is the asymptotic covariance matrix of $\sqrt{T}\mathfrak{d}[\hat{P}_\star]$, which is a Hankel matrix given by

$$[C_P]_{i,j} = 2 \operatorname{var} \left\{ \frac{p^{2n-i-j}}{\Delta^2(p)} u_s(t) \right\}. \quad (3.9)$$

where $u_s(t)$ is a stationary stochastic process having spectral density $\Phi_u^2(i\omega)$. Consequently, the asymptotic covariance matrix of $\sqrt{T}\hat{f}$ is given by

$$C_f = \mathbf{D}C_{\hat{w}}\mathbf{D}'. \quad (3.10)$$

Proof: The expression for C_P in (3.9) follows directly from (3.7), while (3.8) and (3.10) are consequences of (2.17). ■

4 Fitting a rational model

Assume that $u(t)$ has a strictly proper rational spectrum of order ν , *i.e.*,

$$\phi(s) = \frac{c(s)c(-s)}{a(s)a(-s)}, \quad (4.1)$$

where

$$a(s) = s^\nu + \sum_{k=1}^{\nu} a_k s^{\nu-k}, \quad c(s) = \sum_{k=1}^{\nu} c_k s^{\nu-k}.$$

Then the half-spectrum $f(s)$ also admits a strictly proper rational representation as

$$f(s) = \frac{b(s)}{a(s)}, \quad b(s) = \sum_{k=1}^{\nu} b_k s^{\nu-k}.$$

such that

$$\varphi(s) := c(s)c(-s) = a(s)b(-s) + b(s)a(-s). \quad (4.2)$$

Our approach in this work is to identify the *real-valued* parameters $\{a_k\}_{k=1}^{\nu}$ and $\{b_k\}_{k=1}^{\nu}$ from the data. Subsequently we can evaluate the right-hand side of the equation (4.2). Then a spectral factorization of (4.2) will lead to the parameters $\{c_k\}_{k=1}^{\nu}$. Note that the right-hand side of (4.2) needs

to be positive real in order to ensure the existence of a stable spectral factor. Another important issue is to ensure the stability of the estimated polynomial $a(s)$. The problem of computing $a(s)$ and $b(s)$ from the *interpolation conditions* originating from $f(s)$ and its derivatives evaluated at $\{s_k\}_{k=1}^m$ is in fact a linear problem. However, when we impose the stability constraint on $a(s)$ and positivity constraint on the right-hand side of (4.2), we have to solve a Nevanlinna-Pick interpolation problem with a degree constraint [2, 14], which is more difficult.

In order to keep the description simple, we do not consider interpolation conditions involving first and higher order derivatives of $f(s)$. However, the following discussion can be generalized² to account for the interpolation conditions involving first and higher order derivatives of $f(s)$. Thus from now on we consider the case $n_k = 1$ for $k \in \{1, \dots, m\}$.

Let \hat{f}_k be the estimate of $f(s_k)$. Once $n \geq 2\nu$, it is straightforward to solve

$$\hat{f}_k a(s_k) = b(s_k), \quad k \in \{1, \dots, n\}, \quad (4.3)$$

in the least-squares sense. It is also straightforward to incorporate weightings and solve a weighted least-squares problem. In order to derive the statistical properties of the resulting estimates, we need to express the problem in terms of the real-valued vector $\hat{\mathbf{f}}$. For that we need some notation. Define

$$\begin{aligned} \psi_k &:= [s_k^{\nu-1} \ \cdots \ s_k \ 1], \quad \Psi_R := [\psi'_1 \ \cdots \ \psi'_{n_r}]', \quad \Psi_C := [\psi'_{n_r+1} \ \cdots \ \psi'_{n_r+n_c}]'. \\ \gamma_R &:= [s_1^\nu \ \cdots \ s_{n_r}^\nu]', \quad \gamma_C := [s_{n_r+1}^\nu \ \cdots \ s_{n_r+n_c}^\nu]', \quad \gamma := [\gamma'_R \ \text{Re}(\gamma'_C) \ \text{Im}(\gamma'_C)]' \\ \mathbf{F} &:= \begin{bmatrix} \text{diag}(f_R) & 0 & 0 \\ 0 & \text{diag}\{\text{Re}(f_C)\} & -\text{diag}\{\text{Im}(f_C)\} \\ 0 & \text{diag}\{\text{Im}(f_C)\} & \text{diag}\{\text{Re}(f_C)\} \end{bmatrix}, \quad \Psi := \begin{bmatrix} \Psi_R \\ \text{Re}(\Psi_C) \\ \text{Im}(\Psi_C) \end{bmatrix}. \end{aligned}$$

See (2.9) for definitions of f_R and f_C . In the following we denote the estimate of \mathbf{F} derived from $\hat{\mathbf{f}}$ by $\hat{\mathbf{F}}$. Let us introduce

$$\theta_1 := [a_1 \ \cdots \ a_\nu \ b_1 \ \cdots \ b_\nu]', \quad G = [-\mathbf{F}\Psi \ \Psi], \quad \hat{G} = [-\hat{\mathbf{F}}\Psi \ \Psi].$$

It can be verified that $G\theta_1 = \mathbf{F}\gamma$. Consequently, the least-squares estimate of θ_1 is given by

$$\hat{\theta}_1 = [\hat{G}'\Delta\hat{G}]^{-1}\hat{G}'\Delta\hat{\mathbf{F}}\gamma,$$

where Δ is a positive-definite weighting matrix chosen by the user. The following proposition quantifies the asymptotic covariance matrix of $\hat{\theta}_1$.

Proposition 7. Define the matrix U as

$$U_R = \begin{bmatrix} a(s_1) \\ \vdots \\ a(s_{n_r}) \end{bmatrix}, \quad U_C = \begin{bmatrix} a(s_{n_r+1}) \\ \vdots \\ a(s_{n_r+n_c}) \end{bmatrix}, \quad U = \begin{bmatrix} \text{diag}(U_R) & 0 & 0 \\ 0 & \text{diag}\{\text{Re}(U_C)\} & -\text{diag}\{\text{Im}(U_C)\} \\ 0 & \text{diag}\{\text{Im}(U_C)\} & \text{diag}\{\text{Re}(U_C)\} \end{bmatrix}.$$

²To illustrate how to set up a linear problem in terms of interpolation constraints involving derivatives of $f(s)$ note that

$$\frac{df(s)}{ds}a(s) + \frac{da(s)}{ds}f(s) = \frac{db(s)}{ds}.$$

Here we estimate $f(s)$ and $df(s)/ds$ for known s values, giving the interpolation conditions. Thus, the above equation is still linear in $\{a_k\}_{k=1}^\nu$ and $\{b_k\}_{k=1}^\nu$. Now the procedure can be repeated successively for higher order derivatives.

Then the asymptotic covariance matrix of $\sqrt{T}\hat{\theta}_1$ is given by

$$C_{\hat{\theta}_1} = T\mathcal{E}\{(\hat{\theta}_1 - \theta_1)(\hat{\theta}_1 - \theta_1)'\} = [G'\Delta G]^{-1}G'\Delta UC_f U'\Delta G[G'\Delta G]^{-1}. \quad (4.4)$$

The optimum choice of Δ is given by

$$\Delta = (UC_f U')^{-1}.$$

and the associated covariance matrix of the optimum estimate of $\sqrt{T}\theta_1$ is $[G'(UC_f U')^{-1}G]^{-1}$.

Proof: In this proof we denote $\theta_a = [a_1 \cdots a_\nu]'$. Using a standard technique of deriving the asymptotic accuracy of the least-squares based estimates [27, p.285] one can show that the asymptotic estimation error is given by

$$\begin{aligned} \hat{\theta}_1 - \theta_1 &= [\hat{G}'\Delta\hat{G}]^{-1}\hat{G}'\Delta(\hat{\mathbf{F}}\gamma - \hat{G}\theta_1) \\ &\approx [G'\Delta G]^{-1}G'\Delta(\hat{\mathbf{F}}\gamma - \hat{G}\theta_1) \\ &= [G'\Delta G]^{-1}G'\Delta(\hat{\mathbf{F}} - \mathbf{F})(\gamma + \Psi\theta_a) = [G'\Delta G]^{-1}G'\Delta U(\hat{\mathbf{f}} - \mathbf{f}). \end{aligned}$$

Now it is straightforward to derive (4.4). The remaining part of the proposition follows from the theory of the best linear unbiased estimates [18, p.555]. \blacksquare

The implementation of the optimally weighted least-squares estimator requires a bootstrapping procedure. First we need an initial estimate of the system parameters obtained without any weighting. This knowledge is used to compute C_f , which is then used to implement the optimally weighted estimator.

Let $\hat{a}(s)$, $\hat{b}(s)$ and $\hat{\varphi}(s)$ denote the estimates of $a(s)$, $b(s)$ and $\varphi(s)$, respectively, derived from $\hat{\theta}_1$. However the requirement for spectral factorization

$$\hat{\varphi}(s) > 0, \quad \text{Re}(s) = 0. \quad (4.5)$$

may not hold in general. This can be fixed by using a regularization procedure described below. The idea here is to perturb the coefficients of $\hat{\varphi}(s)$, so that (4.5) is enforced. This actually amounts to solving a linear matrix inequality (LMI). In fact we can use the results in [28], (see also [21]), for estimation of moving average processes. Using a bilinear transformation $s = (z - 1)/(z + 1)$ let us define

$$\hat{\varphi}_d(z) := \frac{(z + 1)^{2\nu}}{z^\nu} \hat{\varphi}\left\{\frac{z - 1}{z + 1}\right\} = [(1 + z)(1 + z^{-1})]^\nu \hat{\varphi}\left\{\frac{z - 1}{z + 1}\right\} = \sum_{j=-\nu}^{\nu} \beta_j z^j.$$

It is straightforward to verify the symmetric structure of $\hat{\varphi}_d(z)$, also that $\hat{\varphi}_d(z)$ is real-valued for $|z| = 1$. Using the property of the bilinear transformation we can express (4.5) equivalently as

$$\hat{\varphi}_d(z) > 0, \quad |z| = 1. \quad (4.6)$$

However, (4.6) is equivalent to the constraint that $\hat{\varphi}_d(e^{-i\omega})$ is the spectral density function of a moving average process. This problem occurs naturally in identifying an order ν moving average process where $\{\beta_j\}_{j=0}^\nu$ represent the covariances estimated from the data. Also it is not guaranteed that (4.6) holds. Therefore one looks for the modified coefficients $\{\hat{\beta}_j\}_{j=0}^\nu$ such that

$$\sum_{j=0}^{\nu} (\beta_j - \hat{\beta}_j)^2$$

is minimized subject to the constraint

$$\sum_{j=-\nu}^{\nu} \hat{\beta}_j z^j > 0, \quad |z| = 1.$$

The above optimization problem can be cast as a semidefinite programming problem [28], thus can be solved numerically in a computationally efficient way. We denote the spectral density computed using the optimal solution by $\check{\varphi}_d(z)$. An inverse bilinear transformation gives a refined version $\check{\varphi}(s)$ of $\hat{\varphi}(s)$ such that $\check{\varphi}(i\omega) > 0$ for all real-valued ω . We compute $\check{\varphi}(s)$ as

$$\check{\varphi}(s) = \frac{z^\nu}{(z+1)^{2\nu}} \check{\varphi}_d(z) \Big|_{z=\frac{1+s}{1-s}}$$

From $\check{\varphi}(s)$ we can compute a consistent estimate $\check{c}(s)$ of $c(s)$ by solving the spectral factorization problem

$$\check{\varphi}(s) = \check{c}(s)\check{c}(-s).$$

For a large enough observation length the condition (4.5) is satisfied since $\hat{\theta}_1$ comes sufficiently close to θ_1 due to the consistency properties. Then the regularization procedure is not needed. Therefore, the asymptotic covariance expression in Proposition 7 remains valid. Also by the same reasoning we can extend the result to derive the asymptotic covariance of the estimated c -parameters. This is done in the following proposition. Here for the simplicity of notation we extend the definitions of a_k , b_k and c_k so that $a_k = b_k = c_k = 0$ if $k > \nu$ and $k < 0$. Also $b_0 = c_0 = 0$ and $a_0 = 1$.

Proposition 8. Denote $\theta_a = [a_1 \ \cdots \ a_\nu]'$, $\theta_b = [b_1 \ \cdots \ b_\nu]'$ and $\theta_c = [c_1 \ \cdots \ c_\nu]'$. Also define

$$\theta = [\theta'_a \ \theta'_b \ \theta'_c]'$$

and let $\hat{\theta}$ denote the estimate of θ derived in above. Let us define the matrix valued function $\mathcal{V}(\cdot)$ such that the $\nu \times \nu$ matrix $\mathcal{V}(\theta_a)$ is defined element-wise as

$$[\mathcal{V}(\theta_a)]_{jk} = (-1)^{\nu-k} a_{2j-k}, \quad (4.7)$$

and similarly define $\mathcal{V}(\theta_b)$ and $\mathcal{V}(\theta_c)$. Then the asymptotic covariance matrix of $\sqrt{T}\hat{\theta}$ is given by

$$C_{\hat{\theta}} := T\mathcal{E}\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)'\} = VC_{\hat{\theta}_1}V',$$

where $C_{\hat{\theta}_1}$ is given in Proposition 7, and

$$V = \begin{bmatrix} I & 0 \\ 0 & I \\ \mathcal{V}^{-1}(\theta_c)\mathcal{V}(\theta_b) & \mathcal{V}^{-1}(\theta_c)\mathcal{V}(\theta_a) \end{bmatrix}. \quad (4.8)$$

Proof: Denote the estimation error in θ_a , θ_b and θ_c by $\tilde{\theta}_a$, $\tilde{\theta}_b$ and $\tilde{\theta}_c$, respectively. Also $\tilde{a}(s)$, $\tilde{b}(s)$ and $\tilde{c}(s)$ will denote the associated perturbations in the polynomials $a(s)$, $b(s)$ and $c(s)$, respectively. Then we have the following asymptotic perturbation expansion of the first order:

$$a(s)\tilde{b}(-s) + a(-s)\tilde{b}(s) + b(s)\tilde{a}(-s) + b(-s)\tilde{a}(s) = c(s)\tilde{c}(-s) + c(-s)\tilde{c}(s). \quad (4.9)$$

Now we equate the coefficients of $\{s^k\}_{k=0}^{2\nu}$ on both sides of (4.9). In matrix notation we have

$$\{\mathcal{S}_1(\theta_a) + \mathcal{S}_2(\theta_a)\} \begin{bmatrix} 0 \\ \tilde{\theta}_b \end{bmatrix} + \{\mathcal{S}_1(\theta_b) + \mathcal{S}_2(\theta_b)\} \begin{bmatrix} 0 \\ \tilde{\theta}_a \end{bmatrix} = \{\mathcal{S}_1(\theta_c) + \mathcal{S}_2(\theta_c)\} \begin{bmatrix} 0 \\ \tilde{\theta}_c \end{bmatrix}, \quad (4.10)$$

where $\mathcal{S}_1(\theta_a)$, $\mathcal{S}_2(\theta_a)$ are $(2\nu + 1) \times (\nu + 1)$ matrices defined element-wise as

$$[\mathcal{S}_1(\theta_a)]_{jk} = (-1)^{\nu-j+k} a_{j-k}, \quad [\mathcal{S}_2(\theta_a)]_{jk} = (-1)^{\nu+1-k} a_{j-k}.$$

Then for an even j we see that

$$[\mathcal{S}_1(\theta_a)]_{jk} + [\mathcal{S}_2(\theta_a)]_{jk} = \{(-1)^{\nu-2\nu+k} + (-1)^{\nu+1-k}\} a_{j-k} = \{(-1)^{\nu-k} + (-1)^{\nu+1-k}\} a_{j-k} = 0.$$

Similarly, for an odd j we have

$$[\mathcal{S}_1(\theta_a)]_{jk} = (-1)^{\nu-2\nu-1+k} a_{j-k} = (-1)^{\nu+1-k} a_{j-k} = [\mathcal{S}_2(\theta_a)]_{jk}.$$

Therefore, every even numbered row of (4.10) vanishes. Also the first row of (4.10) vanishes, because the coefficient of $s^{2\nu+1}$ vanishes on both sides of (4.9). Now retaining only the relevant terms in (4.10) corresponding to the 3-rd, 5-th, \dots , $(2\nu + 1)$ -th rows we get

$$\mathcal{V}(\theta_a)\tilde{\theta}_b + \mathcal{V}(\theta_a)\tilde{\theta}_b = \mathcal{V}(\theta_c)\tilde{\theta}_c \quad \Rightarrow \quad \hat{\theta} - \theta = V(\hat{\theta}_1 - \theta_1),$$

where $[\mathcal{V}(\theta_a)]_{jk} = [\mathcal{S}_2(\theta_a)]_{2j+1,k+1} = (-1)^{\nu-k} a_{2j-k}$. Now (4.8) is straightforward from the last equality, since $\mathcal{V}(\theta_c)$ is nonsingular by construction (4.7). \blacksquare

If the roots of $\hat{a}(s)$ obtained from the linear interpolation approach are in the left-half plane, then we get a consistent estimate of the CARMA process transfer function as $\check{c}(s)/\hat{a}(s)$. However, $\hat{a}(s)$ may have roots in the right-half plane in some rare occasions. In that case a second regularization step is required. One popular but heuristic approach to handle this problem is to reflect the unstable roots of $\hat{a}(s)$ to the left-half-plane about the imaginary axis. Another possibility is to use the spectral zero assignability approach in [7]. Note that we have a reliable estimate $\check{c}(s)$ of $c(s)$. Therefore, the convex optimization algorithm in [1] can be used to re-estimate $a(s)$ using $\nu + 1$ interpolation data. A natural way to choose this subset of size $\nu + 1$ is to pick up the interpolation data with lower statistical variation. The expression (3.10) can be used for that purpose. Numerical simulations show that the heuristic strategy of reflecting the unstable poles of $\hat{a}(s)$ to the left-half plane works as good as the Byrnes-Georgiou-Lindquist algorithm in [1].

If we want to fit an AR model, the spectral zeros are *a priori* known. Then there is no need for estimating $c(s)$. It is also possible to extract an accurate estimate of the AR model using the Byrnes-Georgiou-Lindquist algorithm [1].

5 An illustrative example

In this section we illustrate the proposed direct modeling approach using numerical simulations. To conduct the simulation we first need to simulate the sampled version of a continuous-time stochastic process $u(t)$. We do so by using the method in [25, 15] as follows. First we express the continuous-time process $u(t)$ in state space:

$$\begin{aligned} \dot{\zeta}(t) &= A_c \zeta(t) + b_c e(t), \\ u(t) &= c_c \zeta(t), \end{aligned} \quad (5.1)$$

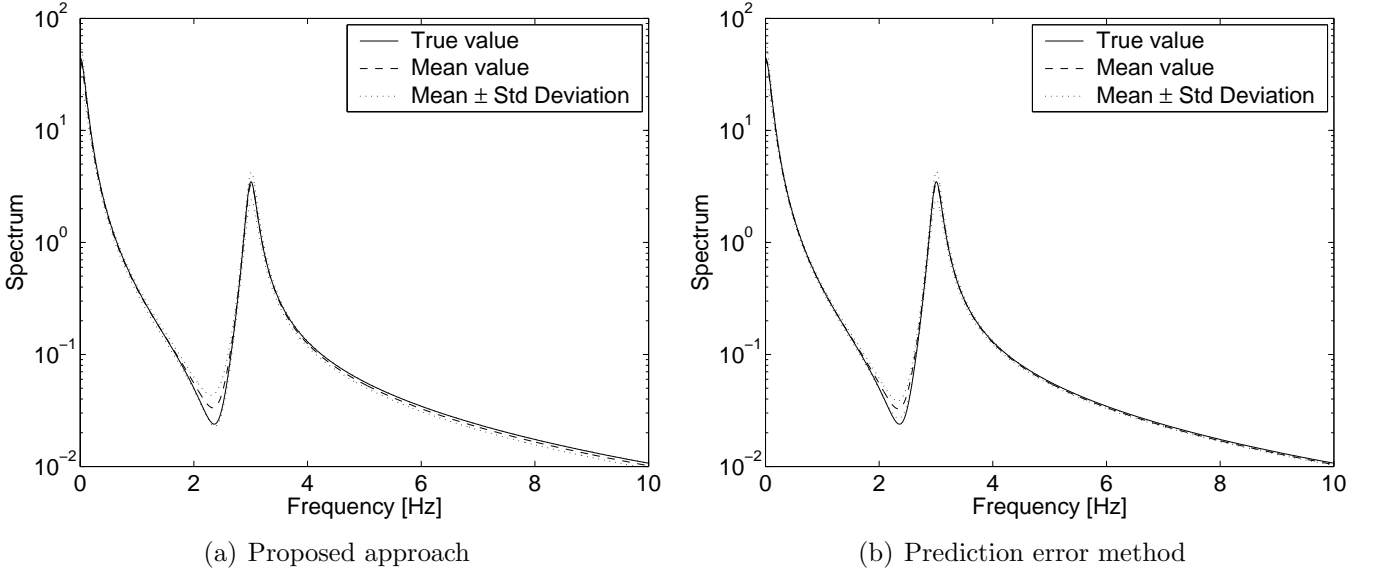


Figure 1: Comparison of the mean of the estimated spectrum (dashed line) and the true spectrum (solid line). The mean \pm standard deviation of the estimated spectrum is shown in dotted lines.

where $\zeta(t)$ is the state vector, and $e(t)$ is a unity variance continuous-time white noise, *i.e.* $\mathcal{E}\{e(t+\tau)e(t)\} = \delta(\tau)$. Note that $\delta(\cdot)$ is the Dirac's delta function. Let the process be sampled at a sampling interval t_s . We want to seek for an *equivalent* discrete-time state space model such that the second-order statistics of the discrete-time model output are the same as that of the underlying continuous-time process at the sampling instants. This discrete-time model is given by [25, 15]

$$\begin{aligned} \zeta\{(k+1)t_s\} &= e^{A_c t_s} \zeta(kt_s) + e_d(kt_s), \\ u(kt_s) &= c_c \zeta(kt_s), \end{aligned} \quad (5.2)$$

where $e_d(kt_s)$ is a vector-valued discrete-time white noise sequence with $\mathcal{E}\{e_d(t)e_d'(t)\} = R_d$. The covariance matrix R_d is related to the underlying continuous-time model via two Lyapunov equations as follows:

$$A_c Q + Q A_c' + b_c b_c' = 0, \quad (5.3)$$

$$Q = e^{A_c t_s} Q e^{A_c' t_s} + R_d \quad (5.4)$$

More precisely, we need to solve for Q in (5.3), and subsequently compute R_d from (5.4). We point out that $Q = \mathcal{E}\{\zeta(t)\zeta'(t)\}$, which is easy to verify from (5.3).

In order to compute the estimates of the half-spectrum at the selected interpolation points we need to compute the output of the input-to-state filter using the samples $u(kt_s)$ at the sampling instants. This is done by a popular discretization technique known as state variable filtering (SVF); see [6] and references therein. In the discretization of the input-to-state filter, the input signal is assumed to vary linearly in between the sampling instants (commonly referred to as the first-order hold).

In the simulations we consider a CARMA model with [see also (4.1)]

$$a(s) = s^3 + 0.3s^2 + 9s + 0.9, \quad c(s) = s^2 + 0.5s + 6.$$

The observation time is 500 sec, and the sampling interval is 0.05 sec. This means that the total number of samples is 10^4 . We estimate the half-spectrum at the interpolation points $\{1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 1 \pm 2i, 1 \pm 3i, 1 \pm 4i\}$. No derivative constraint is used in the estimation process. The interpolation data are then used to fit a CARMA model using the procedure outlined in Section 4. The performance of the proposed algorithm is compared with that of the so-called Prediction Error Method (PEM) based approach [16, 24]. Although PEM is known as the most accurate estimator, it does not always have a solution [16, 24]. Approximately 25% of our simulations fail to give a PEM solution. The results presented here are based on only those realizations for which we get a PEM solution. The estimation results obtained from 100 Monte-Carlo simulations are shown in Figure 1, where the true spectrum is compared with the estimated mean value \pm standard deviation. As can be seen in Figure 1, the estimation accuracy of the proposed approach is comparable with the prediction error method. For about 2% of the cases PEM deviates significantly from the true parameter vector (the optimization routine fails to reach the global minimum). These special cases excluded in the evaluation of the statistical performance. In Table 1 we show the analytical and empirical standard deviations of the proposed estimates. The performance of the proposed approach is similar to PEM as far as estimation of

	a_1	a_2	a_3	c_1	c_2	c_3
True value	0.3	9	0.9	1	0.5	6
Mean	0.3186	9.0331	0.9961	0.9740	0.5887	5.9368
Std-Dev (empirical)	0.0396	0.1240	0.2393	0.0290	0.1201	0.2211
Std-Dev (analytical)	0.0408	0.1126	0.1876	0.0224	0.1152	0.1792

Table 1: Parameter estimation performance of the proposed algorithm.

$\{a_k\}_{k=1}^3$ are concerned. However, PEM is slightly better than the proposed method in estimation of $\{c_k\}_{k=1}^3$ when it gives a solution. The analytical standard deviations of the proposed estimates match well with the empirical standard deviations in Table 1. The algorithms are implemented using Matlab 6.1 on a 2.8 GHz Pentium IV processor with 1 GB RAM. The discrete-time PEM estimate is computed using the `pem` routine in the System Identification Toolbox. The average time required to compute the proposed estimate is 0.2 sec, while it takes 1.0 sec on average to compute the PEM estimate.

6 Conclusions

In this paper we have proposed a novel direct approach for modeling continuous-time stochastic processes. The main idea is to use an input-to-state filter to compute the half-spectrum in some prescribed points in the right-half plane. The estimated samples of the half-spectrum are then used to obtain a rational model of the half-spectrum using linear interpolation with a positivity constraint. This is done by solving a semidefinite programming problem. The unique feature of the approach is two fold. Firstly, it is not required to estimate an intermediate discrete-time model. Thus, we can avoid many numerical difficulties associated with an indirect method. Secondly, it offers estimates which are comparable to PEM in terms of accuracy. However the PEM technique often fails to give a solution, which is not the case with the proposed method. One important open research question is to understand how the interpolation points affect the estimation results. More interpolation points tend to give better estimates (but at an expense of more computation). However, it is not clear how to best choose their locations.

References

- [1] C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 46:6:822–839, June 2001.
- [2] Ph. Delsarte, Y. Genin, Y. Kamp, and P. Van Dooren. Speech modeling and the trigonometric moment problem. *Philips Journal of Research*, 37:277–292, 1982.
- [3] H. Fan, T. Söderström, M. Mossberg, B. Carlsson, and Y. Zhou. Estimation of continuous-time AR process parameters from discrete-time data. *IEEE Transactions on Signal Processing*, 47:1232–1244, 1999.
- [4] H. Fan, T. Söderström, and Y. Zhou. Continuous-time AR process parameter estimation in presence of additive white noise. *IEEE Transactions on Signal Processing*, pages 3392–3398, December 1999.
- [5] A. Feuer and G. C. Goodwin. *Sampling in digital signal processing and control*. Birkhäuser, Boston, 1996.
- [6] H. Garnier and P.C. Young. Time-domain approaches for continuous-time model identification from sampled data. In *Invited tutorial paper for the American Control Conference (ACC'2004)*, Boston, MA (USA), June 2004.
- [7] T. T. Georgiou. The interpolation problem with a degree constraint. *IEEE Transactions on Automatic Control*, 44:3:631–635, March 1999.
- [8] T. T. Georgiou. Spectral estimation via selective harmonic amplification. *IEEE Transactions on Automatic Control*, 46:1:29–42, January 2001.
- [9] T. T. Georgiou. Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parametrization. *IEEE Transactions on Automatic Control*, 47:11:1811–1823, November 2002.
- [10] T. T. Georgiou. The structure of state covariances and its relation to the power spectrum of the input. *IEEE Transactions on Automatic Control*, 47:7:1056–1066, July 2002.
- [11] G. C. Goodwin, R. H. Middleton, and H. V. Poor. High-speed digital signal processing and control. *Proceedings of IEEE*, 80:2:240–259, 1992.
- [12] P. Janssen and P. Stoica. On the expectation of the product of four matrix-valued Gaussian random variables. *IEEE Transactions on Automatic Control*, AC-33:867–870, 1988.
- [13] T. Kailath. *Linear Systems*. Prentice Hall, 1980.
- [14] H. Kimura. Positive partial realization of covariance sequences. In C. I. Byrnes and A. Lindquist, editors, *Modeling, Identification and Robust Control*, pages 499–513. North-Holland, Amsterdam, The Netherlands, 1987.
- [15] E. Larsson. *Identification of stochastic continuous-time systems*. PhD thesis, Uppsala dissertations from the Faculty of Science and Technology: 52, Department of Information Technology, Uppsala University, 2004.

- [16] E. Larsson and M. Mossberg. On possibilities for estimating continuous-time arma parameters. In *IFAC Symposium on System Identification*, Rotterdam, The Netherlands, August 27-29 2003.
- [17] E. K. Larsson and T. Söderström. Identification of continuous-time ar processes from unevenly sampled data. *Automatica*, 38:4:709–718, April 2002.
- [18] L. Ljung. *System Identification - Theory for the User, 2nd edition*. Prentice Hall, Upper Saddle River, NJ, USA, 1999.
- [19] K. Mahata and M. Fu. Modelling continuous-time stochastic processes using input-to-state filters. In *16th IFAC world congress*, Prague, 2005.
- [20] J. Makhoul. Linear prediction: a tutorial overview. *Proceedings of IEEE*, 63:4:561–580, April 1975.
- [21] J. Mari, P. Stoica, and T. McKelvey. Vector arma estimation: a reliable subspace approach. *IEEE Transactions on Signal Processing*, 48:7:2092–2104, July 2000.
- [22] R. H. Middleton and G. C. Goodwin. *Digital Control and Estimation: A unified approach*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.
- [23] R. Pintelon and J. Schoukens. Time series analysis in the frequency domain. *IEEE Transactions on Signal Processing*, 47:1:206–210, 1999.
- [24] T. Söderström. Computing stochastic continuous-time models from ARMA models. *International Journal of Control*, 53:1311–1326, June 1991.
- [25] T. Söderström. *Discrete-time Stochastic Systems, 2 ed*. Springer-Verlag, London, 2002.
- [26] T. Söderström and M. Mossberg. Performance evaluation of methods for identifying continuous-time autoregressive processes. *Automatica*, 36:53–59, January 2000.
- [27] T. Söderström and P. Stoica. *System Identification*. Prentice Hall International, Hemel Hempstead, UK, 1989.
- [28] P. Stoica, T. McKelvey, and J. Mari. MA estimation in polynomial time. *IEEE Transactions on Signal Processing*, 48:7:1999–2012, July 2000.
- [29] B. Wahlberg. The effect of rapid sampling in system identification. *Automatica*, 26:167–170, 1990.