Rational interpolation from phase data by subspace methods

Hüseyin Akçay

Abstract—In this paper, two simple subspace-based identification algorithms to identify stable linear-time-invariant systems from corrupted phase samples of frequency response function are developed. The first algorithm uses data sampled at nonuniformly spaced frequencies and is strongly consistent if corruptions are zero-mean additive random variables with a known covariance function. However, this algorithm is biased when corruptions are multiplicative, yet it exactly retrieves finite-dimensional systems from noise-free phase data using a finite amount of data. The second algorithm uses phase data sampled at equidistantly spaced frequencies and also has the same interpolation and strong consistency properties if corruptions are zero-mean additive random variables. The latter property holds also for the multiplicative noise model provided that some noise statistics are known a priori. Promising results are obtained when the algorithms are applied to simulated data.

Index Terms—Rational interpolation, phase data, time delay estimation, subspace-based identification, strong consistency.

I. INTRODUCTION

Let \( f(z) \) denote the transfer function of a given \( n \)-th order, stable, single-input/single-output, discrete-time system. In this paper, we study construction of \( f(z) \) from its corrupted phase measurements given at a set of points \( e^{i\theta_k}, k = 1, \ldots, N \) in the unit circle \( T \).

Some studies utilizing phase information for parameter estimation have been reported in the literature on signal processing. Time delay estimation between signals acquired by different sensors is intrinsic in many signal processing problems [1] and blind equalization in communications [2].

Time delay estimation procedures can be viewed as inverse Fourier transforms of appropriately normalized or weighted cross-spectral density function measurements which yield delay estimates directly in time domain terms. There are instances where it can be advantageous to base time delay estimates on properly interpreted phase data directly in the frequency domain [3].

Recovery of analytic functions from magnitude or real/imaginary part information on finite subsets of \( T \) is related to the problem studied in this paper. In [4], the problem considered in this paper and the related ones were studied assuming that frequencies are uniformly spaced and finite-impulse response models that interpolate the given data were constructed. Complexity of these models is typically in the order of the number of the data. In [5], a subspace-based algorithm to construct transfer functions of stable, discrete-time systems from their real or imaginary parts evaluated on finite subsets of \( T \) were presented. The consistency properties of this algorithm with respect to the corruptions in the data were also discussed. The major advantage of the subspace-based identification algorithms over the prediction error methods is the absence of nonlinear parametric optimization problems. The subspace-based methods are non-iterative and therefore do not suffer from convergence problems.

The organization of the paper is as follows. In Section II, we reformulate the identification problem as a nonlinear least-squares curve fit problem. Then, this curve fit problem is transformed into another curve fit problem, which will play a key role in the derivation of our subspace algorithms. In Section III, we discuss noise models and put forward noise assumptions. In Sections IV and V, two subspace-based identification algorithms are developed and their consistency properties are studied. The properties of the developed algorithms are illustrated in Section VI by a numerical example. Section VII concludes the paper.

II. NONLINEAR LEAST-SQUARES CURVE FIT METHODS

Assuming that \( f(z) \) is minimum-phase and has no zeros at either the origin or infinity, let us write it as a ratio of two coprime polynomials:

\[
f(z) = \frac{b(z)}{a(z)}
\]  

(1)

where the zeros of \( b(z) \) and \( a(z) \) are in complex conjugate pairs so that the underlying impulse-response is real-valued. Since the phase of \( f(z) \) is invariant to scalings of \( f(z) \) by positive numbers, we may assume that \( a(z) \) and \( b(z) \) are monic polynomials.

Given the phase data, a natural approach to estimate the coefficients of the polynomials \( a(z) \) and \( b(z) \) in (1) will be to minimize the quadratic norm of the residuals:

\[
\sum_{k=1}^{N} |\psi(1, \theta_k) - \psi_k|^2
\]  

(2)

where \( \psi_k \) is the phase of \( f(e^{i\theta_k}) \) corrupted by noise. The dependence of \( \psi(1, \theta) \) on \( \theta \) is highly nonlinear.

The optimization problem above can be recast in the following numerically more tractable form:

\[
\min_{a, b} \sum_{k=1}^{N} \left[ \frac{f(e^{i\theta_k})}{f(e^{-i\theta_k})} - e^{i2\psi_k} \right]^2
\]  

(3)

by noting that

\[
e^{i2\psi(1, \theta)} = \frac{f(z)}{f(z^{-1})} = \frac{b(z)a(z^{-1})}{a(z)b(z^{-1})} = g(z), \quad z = e^{i\theta}.
\]  

(4)

Since \( f(z) \) was assumed to be minimum-phase with no zeros either at the origin or at infinity, \( b(z)z^n a(z^{-1}) \) and
$a(z)z^nb(z^{-1})$ are 2n-degree coprime polynomials and (3) fits 2n-degree rational function $g(z)$ to the frequency response data $e^{2\nu_k} \theta_k = 1, \ldots, N$ in the least-squares sense. The formula (4) will also play a key role in the derivation of the subspace-based algorithms of this paper. The solutions of the above optimization problems are obtained by iterations and these iterations may not converge to global minimum.

III. COMPLEX NOISE MODELS

Let us first assume that the phase data are generated by

$$\psi_k = \psi(1, \theta_k) + \psi_k, \quad k = 1, \ldots, N$$ \hspace{1cm} (5)

where the corruptions $\psi_k$ are independent zero-mean random variables. These corruptions are translated into (3) as multiplicative disturbances. They can be put into the additive disturbance form back by defining

$$e^{2\nu_k} = g(z_k) + \eta_k, \quad k = 1, \ldots, N,$$ \hspace{1cm} (6)

where $\eta_k = g(z_k)(e^{2\nu_k} - 1)$. The new random variables $\eta_k$ are independent. However, their means denoted by $E[\eta_k]$ are nonzero for all but few possible exceptions of $k$.

Alternatively, we propose the following more tractable uncertainty model for phase measurements:

$$e^{2\nu_k} = g(z_k) + \nu_k, \quad k = 1, \ldots, N$$ \hspace{1cm} (7)

where $\nu_k$ are independent zero-mean complex random variables with a covariance function satisfying

$$E \left[ \begin{array}{c} \text{Re} \nu_k \\ \text{Im} \nu_k \end{array} \right] \left[ \begin{array}{c} \text{Re} \nu_l \\ \text{Im} \nu_l \end{array} \right] = \frac{1}{2} \left[ \begin{array}{cc} R_k & 0 \\ 0 & R_k \end{array} \right] \delta_{kl}.$$ \hspace{1cm} (8)

IV. IDENTIFICATION ALGORITHM WITH NONUNIFORMLY SPACED DATA

Let us first consider the noise-free case. By partial fraction expansion, $g(z)$ can be decomposed as

$$g(z) = g_c(z) + g_{ac}(z^{-1})$$ \hspace{1cm} (9)

where $g_{ac}(z^{-1})$ is strictly proper and

$$g_c(z) = \frac{b_c(z)}{a(z)}, \quad g_{ac}(z^{-1}) = \frac{b_{ac}(z^{-1})}{b(z^{-1})}.$$ \hspace{1cm} (10)

From (9) and (10), we write a space-state representation of $g(z)$ as follows:

$$x_c[t+1] = A_c x_c[t] + B_c u[t]$$ \hspace{1cm} (11)

$$x_{ac}[t+1] = A_{ac} x_{ac}[t] + B_{ac} u[t]$$ \hspace{1cm} (12)

$$y[t] = C x_c[t] + C_{ac} x_{ac}[t] + u[t].$$ \hspace{1cm} (13)

We take the discrete Fourier transforms of (11)–(13) where we shift (12) by $p - 1$ samples forward and plug in $U(\theta) = 1$:

$$e^{i\theta} X_c(\theta) = A_c X_c(\theta) + B_c$$ \hspace{1cm} (14)

$$e^{-i\theta} X_{ac,p}(\theta) = A_{ac} X_{ac,p}(\theta) + B_{ac} e^{i(p-1)0} \theta$$ \hspace{1cm} (15)

$$Y(\theta) = C_c X_c(\theta) + C_{ac} e^{-i(p-1)0} X_{ac,p}(\theta) + 1$$

where $X_c(\theta)\ ,\ X_{ac,p}(\theta),\ U(\theta),\ \text{and}\ \ Y(\theta)$ denote the discrete Fourier transforms of $x_c[t]\ ,\ x_{ac}[t+p-1]\ ,\ u[t]\ ,\ \text{and}\ y[t]\ ,\ \text{respectively, and}\ p > 2n$. Thus,

$$g(e^{i\theta}) = C_c X_c(\theta) + C_{ac} e^{-i(p-1)0} X_{ac,p}(\theta) + 1$$

with the state equations satisfying (14) and (15). By iteratively substituting the state equations, we obtain the relation

$$\begin{bmatrix} g(z) \\
\vdots \\
z^{p-1}g(z) \end{bmatrix} = \Gamma_p \begin{bmatrix} 1 \\
\vdots \\
z^{p-1} \end{bmatrix} + \Theta_p \begin{bmatrix} X_c(\theta) \\
X_{ac,p}(\theta) \end{bmatrix}$$

where $z = e^{i\theta}$ and

$$\Theta_p = \begin{bmatrix} C_c & C_{ac} A_{ac}^{-1} \\
\vdots & \vdots \\
C_{ac} A_{ac}^{-1} & C_ac \end{bmatrix},$$

$$\Gamma_p = \begin{bmatrix} 1 & C_{ac} B_c & \cdots & C_{ac} A_{ac}^{-2} B_{ac} \\
\vdots & \ddots & \ddots & \vdots \\
C_{ac} A_{ac}^{-2} B_{ac} & \cdots & 1 \end{bmatrix}.$$ \hspace{1cm} (16)

By repeating for $\theta_k, k = 1, 2, \ldots, N$, we get

$$\mathcal{H} = \mathcal{G} + \Gamma_p \mathcal{W}$$ \hspace{1cm} (17)

$$\mathcal{Y} = \frac{1}{\sqrt{N}} \begin{bmatrix} g(z_1) & \cdots & g(z_N) \\
\vdots & \ddots & \vdots \\
g(z_1) & \cdots & g(z_N) \end{bmatrix},$$ \hspace{1cm} (18)

$$\mathcal{Z} = \frac{1}{\sqrt{N}} \begin{bmatrix} X_c(\theta_1) & \cdots & X_c(\theta_N) \\
\vdots & \ddots & \vdots \\
X_{ac,p}(\theta_1) & \cdots & X_{ac,p}(\theta_N) \end{bmatrix}.$$ \hspace{1cm} (19)

Now, we consider the noisy data case assuming that the corruptions are as in (7). From (7), (16), and (17) we get

$$\mathcal{Y} = \mathcal{G} \mathcal{Z} + \Gamma_p \mathcal{W} + \mathcal{N}$$ \hspace{1cm} (20)

where

$$\mathcal{Y} = \frac{1}{\sqrt{N}} \begin{bmatrix} e^{i\theta_{11}} & \cdots & e^{i\theta_{1N}} \\
\vdots & \ddots & \vdots \\
e^{i\theta_{p1}} & \cdots & e^{i\theta_{pN}} \end{bmatrix},$$

$$\Phi_{kl} = (k-1) \theta_l + 2 \psi_l, 1 \leq k \leq p, 1 \leq l \leq N,$$

$$\mathcal{N} = \frac{1}{\sqrt{N}} \begin{bmatrix} v_1 & \cdots & v_N \\
\vdots & \ddots & \vdots \\
v_1 & \cdots & v_N \end{bmatrix}.$$
Let $\mathcal{W}^\perp$ be the projection matrix onto the null space of $\mathcal{W}$ given by $\mathcal{W}^\perp = I_{2N} - \mathcal{W}^H (\mathcal{W} \mathcal{W}^H)^{-1} \mathcal{W}$. The term $\Gamma_p \mathcal{W}$ is canceled when multiplied from right by $\mathcal{W}^\perp$. Thus,

$$\mathcal{G} \mathcal{W}^\perp = \Omega_p \mathcal{G} \mathcal{X}^\perp + \mathcal{N} \mathcal{W}^\perp = \mathcal{K} \mathcal{W}^\perp + \mathcal{N} \mathcal{W}^\perp.$$ 

The range space of $\mathcal{K} \mathcal{W}^\perp$ equals the range space of $\mathcal{O}_p$ unless rank cancellations occur. In the following, we present sufficient conditions in terms of the data and the system. Given a set $X$, we denote its cardinality by $\mathcal{R}(X)$. Let $N^*$ be a consistent estimate of the range space of $X$. Suppose (20) and the following conditions hold w.p.1

$$\begin{align} &\lim_{N \to \infty} \mathcal{R}(\{z_k : \alpha_k = \pm 1\}) + 2 \mathcal{R}(\{z_k : \alpha_k \neq \pm 1\}) \geq N^* \geq p + 2n, \quad \text{w.p.1} \quad \text{(20)} \\ &\lim_{N \to \infty} \mathcal{N} \mathcal{W}^\perp (\mathcal{N} \mathcal{W}^\perp)^T = \alpha I_p \quad \text{w.p.1} \quad \text{(22)} \end{align}$$

for some scalar $\alpha \geq 0$. If the fourth order moments of $v_k$ in (7) are uniformly bounded, then (21) holds and

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

where $\mathcal{K}$ is defined by $\mathcal{K} \mathcal{K}^T = \text{Re}(\mathcal{V} \mathcal{R} \mathcal{V}^H)$ and

$$\mathcal{R} = \begin{bmatrix} R_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & R_N \end{bmatrix}.$$ 

The matrix $\mathcal{K}$ can be found by a Cholesky decomposition.

Thus, we have the weighted version

$$\mathcal{K}^{-1} \mathcal{G} \mathcal{W}^\perp = \mathcal{K}^{-1} \mathcal{K} \mathcal{W}^\perp + \mathcal{K}^{-1} \mathcal{N} \mathcal{W}^\perp$$

satisfying (21) and (22) with $\alpha = 1$. A numerically efficient way of forming $\mathcal{G} \mathcal{W}^\perp$ is to use the QR-factorization:

$$\begin{bmatrix} \mathcal{W} \\ \mathcal{G} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} Q_{11}^T \\ Q_{22}^T \end{bmatrix}. \quad \text{(23)}$$

A simple derivation yields $\mathcal{G} \mathcal{W}^\perp = R_{22} Q_{22}^T$ and it suffices to use $R_{22}$ since $Q_{22}^T$ is a matrix of full rank. Thus, the 2n left singular vectors corresponding to the 2n largest singular values of $\mathcal{K}^{-1} \mathcal{G} \mathcal{W}^\perp$ are obtained from the SVD:

$$\mathcal{K}^{-1} R_{22} = \begin{bmatrix} \tilde{U}_{2n} \\ \mathcal{V}_{2n} \end{bmatrix} \begin{bmatrix} \Sigma_{2n} & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} \tilde{V}_{2n} \\ \mathcal{V} \end{bmatrix}. \quad \text{(24)}$$

Hence, from the consistency analysis in [8] or [7] we obtain

$$\lim_{N \to \infty} \mathcal{K} \tilde{U}_{2n} = \mathcal{O}_p \tau, \quad \text{w.p.1}$$

for some nonsingular matrix $\tau$. We see that $\tilde{U}_{2n}$ converges to a matrix denoted by $U_{2n}$ w.p.1 as $N \to \infty$ and

$$\mathcal{K} U_{2n} = \mathcal{O}_p \tau. \quad \text{(25)}$$

This asymptotic formula (in the number of data) is the key in the development of our algorithm.

Let $J_a$ and $J_d$ be the upward and downward shift matrices defined by

$$J_a \mathcal{O}_p = \begin{bmatrix} C_a \mathcal{A}_c & C_a \mathcal{A}_{ac}^{-2} \\ \vdots & \vdots \\ C_a \mathcal{A}_{ac}^{p-1} & C_a \mathcal{A}_{ac}^{-1} \end{bmatrix} \quad \text{(26)}$$

and

$$J_d \mathcal{O}_p = \begin{bmatrix} C_c & C_a \mathcal{A}_{ac}^{-1} \\ \vdots & \vdots \\ C_a \mathcal{A}_{ac}^{p-2} & C_a \mathcal{A}_{ac} \end{bmatrix} \quad \text{(27)}$$

Then, $J_a \mathcal{O}_p = J_d \mathcal{O}_p A'$ where

$$A' = \begin{bmatrix} A_c & 0 \\ 0 & A_{ac}^{-1} \end{bmatrix}. \quad \text{(28)}$$

Hence, $A' = (J_d \mathcal{O}_p) J_a \mathcal{O}_p = T A'' T^{-1}$ where $(\cdot)^T$ denotes the Moore-Penrose inverse of a full column-rank matrix and $A'' = (J_a \mathcal{K} U_{2n})^T J_a \mathcal{K} U_{2n}$. Thus, $A'$ and $A''$ are similar matrices. This means that they have the same eigenvalues.

Due to the assumptions on the zeros of $b(z)$, there is a one-to-one correspondence between the eigenvalues of $A''$ and the zeros of $a(z)$ and $b(z)$. The eigenvalues of $A''$ inside $\mathcal{T}$ are the poles of $f(z)$ and the eigenvalues of $A''$ outside $\mathcal{T}$ are the inverses of the zeros of $f(z)$. We summarize our algorithm in the following.

**Algorithm 4.1:**

1. Given $\{\psi_k, \theta_k, R_k : k = 1, \cdots, N\}$, form $\mathcal{G}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{K}$.
2. Calculate the QR-factorization in (23).
3. Calculate the SVD in (24).
4. Determine $n$ by inspecting the singular values.
5. Calculate $A = (J_a \mathcal{K} U_{2n})^T J_d \mathcal{K} U_{2n}$.
6. Let the eigenvalues of $A$ inside $\mathcal{T}$ be $\hat{p}_k$ and the inverses of the eigenvalues of $A$ outside $\mathcal{T}$ be $\hat{z}_k, k = 1, \cdots, n$.
7. Calculate an estimator of $f(z)$ as

$$\hat{f}(z) = \frac{1}{\pi} \sum_{k=1}^{n} \frac{z - \hat{z}_k}{z - \hat{p}_k}. \quad \text{(29)}$$

The following is the first interpolation result of this paper.

**Theorem 4.1:** Let $f(z)$ be a scalar, stable, minimum-phase, discrete-time system of order $n$ satisfying $f(0) \neq 0$ and $f(\infty) = 1$. Consider Algorithm 4.1 with $N$ noise-free samples of $\arg f(e^{j\theta})$ at distinct frequencies $\theta_k \in [0, \pi]$ and let $\hat{f}(z)$ be as in (29). Let $N^*$ be as in (20) and assume that $N^* \geq p + 2n$ where $p > 2n$. Let $\mathcal{K} \in \mathcal{R}^{p \times p}$ be any nonsingular matrix. Then, $f = \hat{f}$.

For the noise (7), we have the following consistency result.

**Theorem 4.2:** Let $f(z)$ be as in Theorem 4.1. Consider Algorithm 4.1 with noisy phase measurements of $f(z)$ where the corruptions are as in (7). Let $\hat{f}(z)$ be as in (29). Then,

$$\lim_{N \to \infty} \sup_{z \in \mathcal{T}} |\hat{f}(z) - f(z)| = 0, \quad \text{w.p.1.} \quad \text{(25)}$$
A. Multiplicative noise case

Let us now assume that the corruptions are as in (5). The results derived in this subsection will form the basis of the second algorithm of this paper.

Let 
\[ m_k = E[e^{2\psi_k}] - 1, \quad k = 1, \cdots, N. \]  
(30)

Then, \( \eta_k = g(z_k)m_k + g(z_k)\hat{\xi}_k \) where \( \hat{\xi}_k = e^{2\psi_k} - E[e^{2\psi_k}] \) are zero-mean independent random variables. Hence, assuming \( m_k \) known, we get
\[
(1 + m_k)^{-1}e^{2\psi_k} = g(z_k) + \bar{\nu}_k, \quad k = 1, \cdots, N
\]  
(31)

where \( \bar{\nu}_k = (1 + m_k)^{-1}g(z_k)\hat{\xi}_k \). Thus,
\[
\bar{\gamma} = \theta_p\bar{\mathcal{X}} + \Gamma_p\mathcal{Y} + \bar{\mathcal{N}} = \bar{H} + \bar{\mathcal{N}}
\]

where \( \bar{\gamma} = \tilde{G}\mathcal{M}^{-1} \) and
\[
\mathcal{M} = I_N + \begin{bmatrix} m_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & m_N \end{bmatrix}
\]  
(32)

\[
\bar{\mathcal{N}} = \frac{1}{\sqrt{N}} \begin{bmatrix} \nu_1 & \cdots & \nu_N \\ \vdots & \ddots & \vdots \\ e^{i\theta_1}\nu_1 & \cdots & e^{i\theta_N}\nu_N \end{bmatrix}
\]

\[
\tilde{G} = \tilde{G}\mathcal{M}^{-1}
\]

A modified version of Algorithm 4.1, which we call Algorithm 4.1', is obtained by substituting \( \bar{\gamma} \) and \( I_p \) in places of \( \tilde{G} \) and \( \mathcal{M} \). Algorithm 4.1' is not consistent either since the covariance function of \( \bar{\nu}_k \) depends on the unknown transfer function \( g(z) \). When the corruptions are independent identically distributed random variables, Algorithm 4.1' coincides with Algorithm 4.1. It remains to be settled if there exist a consistent subspace algorithm over the multiplicative noise model (5). The answer is affirmative and the algorithm we propose is based on the key observation that \( \bar{\nu}_k \) are bounded, zero-mean independent random variables.

V. IDENTIFICATION ALGORITHM WITH UNIFORMLY SPACED DATA

Let us assume that \( M + 1 \) phase data \( \psi_k \) on a set of uniformly spaced frequencies \( \omega_k = \pi k / M, k = 0, \cdots, M \) are given. Since \( f(z) \) has a real-valued impulse response, its frequency response on \( [0, \pi] \) can be extended to \( [\pi, 2\pi] \) by complex conjugation which implies that the phase data be extended as \( \psi_{M+k} = -\psi_{M-k}, k = 1, \cdots, M - 1 \) and we set \( N = 2M \). Let
\[
\bar{\mathcal{F}} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ e^{i\pi(N-1)/N} & \cdots & e^{i\pi(N-1)/N} \end{bmatrix}
\]  
(33)

with \( r \) fixed. Then,
\[
\bar{\mathcal{F}} = \bar{H}\mathcal{F} + \bar{\mathcal{N}}\mathcal{F}
\]  
(34)

and (21) and (22) are satisfied with \( \alpha = 0 \). Thus, the \( 2n \) left singular vectors corresponding to the \( 2n \) largest singular values of \( \bar{\mathcal{F}} \) form a strongly consistent estimate of the range space of \( \bar{\mathcal{H}}\mathcal{F} \). This conclusion does not change either for the additive noise model (7).

For the bias, we consider the term \( \bar{H}\mathcal{F} \) in (34). Note that \( \mathcal{Y} = 0 \). From (17), we see that the \( k \)th row and the \( l \)th column entry of \( \bar{H}\mathcal{F} \) is obtained by taking the \( N \)-point inverse discrete Fourier transform of the noise-free data:
\[
[\bar{H}\mathcal{F}]_{k,l} = \frac{1}{N} \sum_{m=0}^{N-1} e^{\frac{2\pi}{N}m(k-l-1)}g(e^{\frac{2\pi}{N}m}) = \xi_{k+l-1}.
\]

Note from (11)–(13) that
\[
g(z) = C_c(zI_n - A_c)^{-1}B_c + 1 + C_{ac}(zI_n - A_{ac})^{-1}B_{ac}.
\]  
(35)

Lemma 5.1: Let \( g(z) \) be as in (35). Then,
\[
\xi_k = \begin{cases} 1 + C_cA_c^{-1}I_cB_c + C_{ac}A_{ac}^{-1}I_cB_{ac}, & k = 0, \\ C_cA_c^{-1}I_cB_c + C_{ac}A_{ac}^{-1}I_cB_{ac}, & k > 0. \end{cases}
\]

where \( I_c = (I_n - A_c)^{-1} \) and \( I_{ac} = (I_n - A_{ac})^{-1} \).

Now, let
\[
\Delta_r = \begin{bmatrix} B_c & \cdots & A_c^{-1}B_c \\ A_{ac}^{-1}B_{ac} & \cdots & B_{ac} \end{bmatrix}
\]

and \( P = N - p - r \). It is straightforward to show that the Hankel matrix \( \bar{\mathcal{H}}\mathcal{F} \) can be factorized as follows
\[
\bar{\mathcal{H}}\mathcal{F} = \theta_p\begin{bmatrix} I_c & 0 \\ 0 & A_p^\alpha \end{bmatrix} \Delta_r
\]  
(37)

yielding the observability range space \( \theta_p \). A dual factorization can be derived as follows.

Let
\[
\bar{\mathcal{H}} = \frac{1}{\sqrt{N}} \begin{bmatrix} g(z_1) & \cdots & g(z_N) \\ \vdots & \ddots & \vdots \\ e^{i\pi(N-1)/N}g(z_1) & \cdots & e^{i\pi(N-1)/N}g(z_N) \end{bmatrix}
\]  
(38)

where \( z_k = e^{i\omega_k}, k = 1, \cdots, N \). Then, a factorization of \( \bar{\mathcal{H}}\mathcal{F} \) analogous to (37) is given by
\[
\bar{\mathcal{H}}\mathcal{F} = \bar{\theta}_p\begin{bmatrix} A_p^\alpha & I_c & 0 \\ 0 & A_p^\alpha & I_{ac} \end{bmatrix} \tilde{\Delta}_r
\]  
(39)

where
\[
\bar{\theta}_p = \begin{bmatrix} C_cA_c^{-1} & C_{ac} \\ \vdots & \vdots \\ C_c & C_{ac}A_c^{-1} \end{bmatrix},
\]

\[
\tilde{\Delta}_r = \begin{bmatrix} A_c^{-1}B_c & \cdots & B_c \\ B_{ac} & \cdots & A_{ac}^{-1}B_{ac} \end{bmatrix}.
\]  
(40)  
(41)

Pre and post multiplication of the latter factorization formula with the permutation matrices \( \mathcal{I}_p \) and \( \mathcal{I}_r \):
\[
\mathcal{I}_p\bar{\mathcal{H}}\mathcal{F}\mathcal{I}_r = \theta_p\begin{bmatrix} A_p^\alpha & I_c & 0 \\ 0 & A_p^\alpha & I_{ac} \end{bmatrix} \Delta_r
\]  
(42)
Let
\[ Y = \hat{H} F + \tilde{S} \hat{H} F \Sigma_r. \] (43)

Then, from (37), (39), and (43)
\[ Y = \mathcal{O} \begin{bmatrix} \mathcal{I}_c + \mathcal{A}_c^p \mathcal{J}_c & 0 \\ 0 & \mathcal{I}_{ac} + \mathcal{A}_{ac}^p \mathcal{J}_{ac} \end{bmatrix} \Delta_r. \] (44)

Since the spectra of \( \mathcal{A}_c \) and \( \mathcal{A}_{ac} \) are both inside \( \mathbf{T} \), the middle term on the right-hand side of (44) is always nonsingular. Hence, the range space of \( \mathcal{O} \) equals to the range space of \( Y \). Moreover, as \( N \to \infty \), \( \mathcal{T} \to \mathcal{O} \Delta_r \). Finally, the entries of \( \mathcal{T} \) are computed via the \( N \)-point inverse Fourier transform of \( \hat{H} F \), that is, for \( 1 \leq k \leq p \) and \( 1 \leq l \leq r \):
\[ \mathcal{G}_{kl} = \hat{g}_{k+l-1} + \tilde{g}_{p+k-l-1}. \] (45)

Replacing \( \hat{H} \) with \( \tilde{H} \) and unifying (44) with the stages of Algorithm 4.1 we obtain the following algorithm.

**Algorithm 5.1:**

1. Given \( \psi_k, \omega_k, m_k, k = 0, \ldots, M \), expand the statistic data according to \( m_{M+k} = E[e^{-2j\omega_k}] - 1 \) to obtain signals of lengths \( N = 2M \).
2. With \( p \) and \( r \) fixed as \( p > 2n, r \geq 2n \), and \( p + r \leq N \), using the expanded data compute the Hankel matrix \( \mathcal{Y} \) in (45) with \( \tilde{H} \) replacing \( \hat{H} \).
3. Calculate the SVD
\[ \mathcal{Y} = \begin{bmatrix} U_{2n}^t & U_n^t \end{bmatrix} \begin{bmatrix} \Sigma_{2n}^2 & 0 \\ 0 & \Sigma_n \end{bmatrix} \begin{bmatrix} V_{2n}^t \\ V_n^t \end{bmatrix}. \] (46)
4. Determine \( n \) by inspecting the singular values.
5. Calculate \( \mathcal{A}_c^p = (J_a U_{2n}^{\prime p})^t J_d U_n^{\prime p} \).
6. Let the eigenvalues of \( \mathcal{A}_c^p \) inside \( \mathbf{T} \) be \( \lambda_k^c \) and the inverses of the eigenvalues of \( \mathcal{A}_c^p \) outside \( \mathbf{T} \) be \( \lambda_k^{-c} \), \( k = 1, \ldots, n \).
7. Calculate an estimator of \( f(z) \) as
\[ f^\sharp(z) = \prod_{k=1}^n \frac{z - \lambda_k^c}{z - \lambda_k^{-c}}. \] (47)

Interpolation property of this algorithm is summarized in the following result.

**Theorem 5.1:** Let \( f(z) \) be a scalar, stable, minimum-phase, discrete-time system of order \( n \) satisfying \( f(0) \neq 0 \) and \( f(\infty) = 1 \). Consider Algorithm 5.1 with \( M + 1 \) noise-free samples of \( \arg f(e^{j\theta}) \) at uniformly spaced frequencies and let \( f^\sharp(z) \) be as in (47). Then, \( f^\sharp = f \).

By picking \( p - 1 = r = 2n \), for Algorithm 5.1 we have the requirement \( 4n + 1 \leq 2M \) to interpolate noise-free phase data. This requirement is satisfied by letting \( M = 2n + 1 \). On the other hand, for Algorithm 4.1 applied to phase data on uniform grids of frequencies, from (20) we read off \( N^* = 2M - 2 \). Since \( p > 2n \), we pick \( p = 2n + 1 \). Then, from Theorem 4.1 the interpolation condition is \( 2M - 2 \geq 4n + 1 \), which is satisfied by letting \( M = 2n + 2 \).

Consistency properties of Algorithm 5.1 are captured in the following result.

**Theorem 5.2:** Let \( f(z) \) be as in Theorem 5.1. Consider Algorithm 5.1 with corrupted phase measurements of \( f(z) \) where the corruptions are as in (5) or (7). In the former case, suppose that the means in (30) are known. Let \( f^\sharp(z) \) be as in (47). Then,
\[ \lim_{M \to \infty} |f^\sharp(z) - f(z)| = 0, \quad \text{w.p.1.} \] (48)

When the corruptions are additive (7), Algorithm 5.1 is consistent without requiring noise covariance information whereas Algorithm 4.1 is consistent only if the first and the second order noise moments are known.

**A biased yet asymptotically unbiased form**

In this rest of this section, we will briefly dwell on a modified form of Algorithm 5.1 obtained as follows.

First, let
\[ \Phi_{kl} = \begin{bmatrix} e^{j \psi_k} & \cdots & e^{j \psi_n} \\ \vdots & \ddots & \vdots \\ e^{j \psi_1} & \cdots & e^{j \psi_N} \end{bmatrix} \mathcal{O}^{-1}, \] (49)

where the columns of \( \Sigma_{2n}^2 \) and \( \Sigma_n \) contain the 2\( n \) largest singular values. Next, calculate
\[ \mathcal{A}_c^p = (J_d U_{2n}^{\prime p})^t J_d U_n^{\prime p} \] (50)

where the columns of \( U_n^{\prime p} \in \mathbb{R}^{p \times n} \) and \( U_n^{\prime n} \in \mathbb{R}^{2p \times n} \) are the first \( n \) columns of \( U_{2n}^{\prime p} \) and \( U_n^{\prime p} \), respectively. Denote the eigenvalues of \( \mathcal{A}_c^p \) and \( \mathcal{A}_c^{p-1} \) respectively by \( \mu_k^c \), \( k = 1, \ldots, n \). Calculate an estimator of \( f(z) \) as
\[ f^\sharp(z) = \prod_{k=1}^n \frac{z - \mu_k^c}{z - \mu_k^{-c}}. \] (53)

As \( N \to \infty \), observe that (37) satisfies
\[ \tilde{H} F \to \begin{bmatrix} \mathcal{C}_c \\ \vdots \\ \mathcal{C}_{n-1} \mathcal{C}_c^{p-1} \end{bmatrix} [B_c \cdots \mathcal{A}_c^{p-1} B_c]. \] (54)

Likewise, from (39) we have as \( N \to \infty \),
\[ \tilde{H} F \to \begin{bmatrix} \mathcal{C}_{ac} \\ \vdots \\ \mathcal{C}_{n-1} \mathcal{C}_{ac}^{p-1} \end{bmatrix} [B_{ac} \cdots \mathcal{A}_{ac}^{p-1} B_{ac}]. \] (55)

From (54) and (55), we conclude that \( f^\sharp(z) \) is a strongly consistent estimator of \( f(z) \) over the additive and the multiplicative noise models in (7) and (5) despite it is biased for finite \( N \) even if data are noise-free. However, this bias error quickly disappears as \( N \) increases.

If \( f(z) \) were calculated by taking \( n \) eigenvalues of
\[ \tilde{A} = (J_d U_{2n}^{\prime p})^t J_d U_n^{\prime p} \] (56)
inside \( \mathbf{T} \) as the poles and \( n \) eigenvalues outside \( \mathbf{T} \) as the zeros, this estimator would be interpolatory. However, due to the convergence in (54), it is not a reliable estimator of \( f(z) \) when \( N \) is large even if data are noise-free.
VI. ILLUSTRATIVE EXAMPLE

We consider a fourth-order system described by the state-space model:

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

\[ B = \begin{bmatrix}
0.2247 & 0.8989 & 0.0323 & 0.1290
\end{bmatrix}^T, \]

\[ C = \begin{bmatrix}
0.4902 & 0.1168 & 10.0534 & 1.6756
\end{bmatrix}, \quad D = 1 \]

and assume that the corruptions \( \psi_k \) in (5) are independent and uniformly distributed random variables in \([-0.2, 0.2]\). We applied the curve fit algorithm \texttt{lsqnonlin} in MATLAB to (2) with corrupted phase data on 513 uniformly spaced frequencies and found the algorithm sensitive to parameter initializations even in the absence of noise.

Next, we applied Algorithm 4.1 with \( p = 9 \), \( R_k = 1 \) for all \( k \). In Figure 1, the frequency responses of the 10 estimated models generated by Monte Carlo simulations using different noise realizations and fixed frequencies are plotted. Our purpose was to illustrate variability of the estimates in the absence of noise covariance information. As noted earlier, noise covariance information is not available since it depends on the unknown transfer function. However, in the uniformly spaced case, it is possible to reduce this noise induced variability as will be demonstrated next.

![Fig. 1](image1.png)

Fig. 1. The frequency responses of the system and the results from Monte Carlo simulations for the 10 estimated models using Algorithm 4.1 with \( p = 9 \) and \( R_k = 1 \).

Lastly, we applied the modified form of Algorithm 5.1 in Section V-A to the same data set with \( p = r = N/2 = 256 \). In Figure 2, the frequency responses of the 10 estimated models computed using (53) are plotted. The variability of the estimates due to noise is reduced by the data preprocessing although this reduction is not easily visible, which may be attributed to the large ratios of \( p \) and \( r \) to \( N \).

Next, we verify interpolation property of Algorithm 5.1. Instead of \( A^T \) in (47), we equivalently used \( \tilde{A} \) in (56) to compute the poles and the zeros of \( f^T(z) \). Based on 10 measurements, by letting \( M = p = 2n + 1 = 9 \), and \( r = 2n = 8 \) we computed the amplitude of the mismatch between the true and the estimated transfer functions as 6.5703 \times 10^{-13}. However, this interpolatory estimator has poor consistency properties even for modest data records and small noise amplitudes in contrast to \( f^T \). For example, with \( M = 256 \) and \( \varepsilon = 0.02 \), the transfer function estimate has rather erratic behavior, which is omitted for space considerations.

VII. CONCLUSIONS

In this paper, we developed two simple identification algorithms to identify stable linear-time-invariant systems from phase samples of the frequency response function. The developed algorithms were used in a simulation study and the results showed that the subspace-based algorithms outperform the curve fit method on the same data.

REFERENCES