A prediction perspective on the Wiener-Hopf equations for discrete time series

Suhasini Subba Rao*[1] and Junho Yang†‡[1]

[1]Texas A&M University, College Station, TX 77845, U.S.A.

July 11, 2021

Abstract

The Wiener-Hopf equations are a Toeplitz system of linear equations that naturally arise in several applications in time series. These include the update and prediction step of the stationary Kalman filter equations and the prediction of bivariate time series. The celebrated Wiener-Hopf technique is usually used for solving these equations, and is based on a comparison of coefficients in a Fourier series expansion. However, a statistical interpretation of both the method and solution is opaque. The purpose of this note is to revisit the (discrete) Wiener-Hopf equations and obtain an alternative solution that is more aligned with classical techniques in time series analysis. Specifically, we propose a solution to the Wiener-Hopf equations that combines linear prediction with deconvolution.

The Wiener-Hopf solution requires the spectral factorisation of the underlying spectral density function. For general spectral density functions this is infeasible. Therefore, it is usually assumed that the spectral density is rational, which allows one to obtain a computationally tractable solution. However, this leads to an approximation error when the underlying spectral density is not a rational function. We use the proposed solution together with Baxter's inequality to derive an error bound for the rational spectral density approximation.

Keywords and phrases: Deconvolution, linear prediction, semi-infinite Toeplitz matrices, stationary time series and Wiener-Hopf equations.

*`suhasini@stat.tamu.edu`

†`junhoyang.stat@gmail.com`

‡Authors ordered alphabetically.

1 Introduction

The Wiener-Hopf technique (Wiener and Hopf, 1931; Hopf, 1934) was first proposed in the 1930s as a method for solving an integral equation of the form

$$g(\tau) = \int_0^\infty h(t)c(\tau - t)dt \qquad \text{for} \qquad \tau \in [0, \infty)$$

in terms of $h(\cdot)$, where $c(\cdot)$ is a known difference kernel and $g(\cdot)$ is a specified function. The above integral equation and the Weiner-Hopf technique have been used widely in many applications in applied mathematics and engineering (see Lawrie and Abrahams (2007) for a review). In the 1940s, Wiener (1949) reformulated the problem within discrete "time", which is commonly referred to as the Wiener (causal) filter. The discretization elegantly encapsulates several problems in time series analysis. For example, the best fitting finite order autoregressive parameters fall under this framework. The autoregressive parameters can be expressed as a finite interval Wiener-Hopf equations (commonly referred to as the FIR Wiener filter), for which Levinson (1947) and Durbin (1960) proposed a $O(n^2)$ method for solving these equations. More broadly, the best linear predictor of a causal stationary time series naturally gives rise to the Wiener filter. For example, the prediction of hidden states based on the observed states in a Kalman filter model. The purpose of this paper is to revisit the discrete-time Wiener-Hopf equations (it is precisely defined in (1.2)) and derive an alternative solution using the tools of linear prediction. Below we briefly review some classical results on the Wiener filter.

Suppose that $\{\underline{Z}_t = (X_t, Y_t)' : t \in \mathbb{Z}\}$ is a real-valued, zero mean bivariate weakly stationary time series where $\{X_t\}$ and $\{Y_t\}$ are defined on the same probability space $(\Omega, \mathcal{F}, P)$. Let $c(\ell) = \mathrm{cov}(X_0, X_{-\ell})$ and $c_{YX}(\ell) = \mathrm{cov}(Y_0, X_{-\ell})$ be the autocovariance and cross-covariance function respectively. Let $\mathcal{M}$ denote the real Hilbert space in $L^2(\Omega, \mathcal{F}, P)$ spanned by $\{X_t : t \in \mathbb{Z}\} \cup \{Y_t : t \in \mathbb{Z}\}$. The inner product and norm on $\mathcal{M}$ is $\langle U, V\rangle = \mathrm{cov}[U, V]$ and $\|U\| = \langle U, U\rangle^{1/2}$ respectively. For $t \in \mathbb{Z}$, let $\mathcal{H}_t = \overline{\mathrm{sp}}(X_j : j \le t)$ be the closed subspace of $\mathcal{M}$ spanned by $\{X_j : j \le t\}$ and $P_{\mathcal{H}_t}$ is the orthogonal projection of $\mathcal{M}$ onto $\mathcal{H}_t$. The orthogonal projection of Y_0 onto $\mathcal{H}_0$ is

$$P_{\mathcal{H}_0}(Y_0) = \sum_{j=0}^{\infty} h_j X_{-j}, \tag{1.1}$$

were $P_{\mathcal{H}_0}(Y_0) = \arg\min_{U \in \mathcal{H}_0} \|Y_0 - U\|$. To evaluate $\{h_j : j \ge 0\}$, we rewrite (1.1) as a system of linear equations. By using that $P_{\mathcal{H}_0}(Y_0)$ is an orthogonal projection of Y_0 onto $\mathcal{H}_0 = \overline{\mathrm{sp}}(X_\ell : \ell \le 0)$ it is easily shown that (1.1) leads to the system of normal equations

$$c_{YX}(\ell) = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \text{for} \qquad \ell = 0, 1, ... \tag{1.2}$$

The above set of equations is typically referred to as the Wiener-Hopf equations (or semi-infinite

Toeplitz equations). There are two well-known methods for solving this equation in the frequency domain; the Wiener-Hopf technique (sometimes called the gapped function, see Wiener (1949)) and the prewhitening method proposed in Bode and Shannon (1950) and Zadeh and Ragazzini (1950). Both solutions solve for $H(\omega) = \sum_{j=0}^{\infty} h_j e^{ij\omega}$ (see Kailath (1974), Kailath (1980) and Orfanidis (2018), Sections 11.3-11.8). The Wiener-Hopf technique is based on the spectral factorization and a comparison of Fourier coefficients corresponding to negative and positive frequencies. The prewhitening method, as the name suggests, is more in the spirit of time series where the time series $\{X_t\}$ is whitened using an autoregressive filter. We assume the spectral density $f(\omega) = \sum_{r\in\mathbb{Z}} c(r)e^{ir\omega}$ satisfies the condition $0 < \inf_\omega f(\omega) \leq \sup_\omega f(\omega) < \infty$. Then, $\{X_t\}$ admits an infinite order MA and AR representation (see Pourahmadi (2001), Sections 5-6 and Krampe et al. (2018), page 706)

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}, \qquad X_t - \sum_{j=1}^{\infty} \phi_j X_{t-j} = \varepsilon_t \qquad t \in \mathbb{Z} \tag{1.3}$$

where $\sum_{j=1}^{\infty} \psi_j^2 < \infty$, $\sum_{j=1}^{\infty} \phi_j^2 < \infty$, and $\{\varepsilon_t\}$ is a white noise process with $\mathbb{E}\varepsilon_t^2 = \sigma^2 > 0$.

From (1.3), we immediately obtain the spectral factorization $f(\omega) = \sigma^2|\phi(\omega)|^{-2}$, where $\phi(\omega) = 1 - \sum_{j=1}^{\infty} \phi_j e^{ij\omega}$. Given $A(\omega) = \sum_{j=-\infty}^{\infty} a_j e^{ij\omega}$, we use the notation $[A(\omega)]_+ = \sum_{j=0}^{\infty} a_j e^{ij\omega}$ and $[A(\omega)]_- = \sum_{j=-\infty}^{-1} a_j e^{ij\omega}$. Both the Wiener-Hopf technique and prewhitening method yield the solution

$$H(\omega) = \sigma^{-2}\phi(\omega)[\phi(\omega)^* f_{YX}(\omega)]_+, \tag{1.4}$$

where $f_{YX}(\omega) = \sum_{\ell\in\mathbb{Z}} c_{YX}(\ell)e^{i\ell\omega}$ and $\phi(\omega)^*$ is a complex conjugate of $\phi(\omega)$.

The normal equations in (1.2) belong to the general class of Wiener-Hopf equations of the form

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \text{for} \qquad \ell \geq 0, \tag{1.5}$$

where $\{c(r) : r \in \mathbb{Z}\}$ is a symmetric, positive definite sequence. The Wiener-Hopf technique yields the solution

$$H(\omega) = \sigma^{-2}\phi(\omega)[\phi(\omega)^* G_+(\omega)]_+, \tag{1.6}$$

where $G_+(\omega) = \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}$ (the derivation is well-known, but for completeness we give a short proof in Section 2.3). An alternative method for solving for $\{h_j : j \geq 0\}$ is within the time domain. This is done by representing (1.5) as the semi-infinite Toeplitz system

$$\mathbf{g}_+ = T(f)\mathbf{h}_+, \tag{1.7}$$

where $\mathbf{g}_+ = (g_0, g_1, g_2, \ldots)'$ and $\mathbf{h}_+ = (h_0, h_1, \ldots)'$ are semi-infinite (column) sequences and $T(f)$ is a Toeplitz matrix of the form $T(f) = (c(t-\tau); t, \tau \geq 0)$. Let $\{\phi_j : j \geq 0\}$ (setting $\phi_0 = -1$) denote the autoregressive coefficients corresponding to f defined as in (1.3), $\phi(\cdot)$ be its Fourier transform. By letting $\phi_j = 0$ for $j < 0$, we define the lower triangular Toeplitz matrix $T(\phi) = (\phi_{t-\tau}; t, \tau \geq 0)$. Provided that $0 < \inf_\omega f(\omega) \leq \sup_\omega f(\omega) < \infty$, it is well-known that $T(f)$ is invertible on $\ell_2^+ = \{(v_0, v_1, ...) : \sum_{j=0}^{\infty} |v_j|^2 < \infty\}$, and the inverse is $T(f)^{-1} = \sigma^{-2} T(\phi) T(\phi)^*$ (see, for example, Theorem III of Widom (1960)), thus the time domain solution to (1.5) is $\mathbf{h}_+ = T(f)^{-1} \mathbf{g}_+ = \sigma^{-2} T(\phi) T(\phi)^* \mathbf{g}_+$.

In this paper, we study the Wiener-Hopf equations from a time series perspective, combining linear prediction methods developed in the time domain with the deconvolution method in the frequency domain. Observe that (1.5) is semi-infinite convolution equations (since the equations only hold for non-negative index ℓ), thus the standard deconvolution approach is not possible. In Subba Rao and Yang (2021), we used the tools of linear prediction to rewrite the Gaussian likelihood of a stationary time series within the frequency domain. We transfer some of these ideas to solving the Wiener-Hopf equations. In Section 2.2, we show that we can circumvent the constraint $\ell \geq 0$, by using linear prediction to yield the normal equations in (1.2) for all $\ell \in \mathbb{Z}$. In Section 2.3, we show that there exists a stationary time series $\{X_t\}$ and random variable $Y \in \overline{\text{sp}}(X_j : j \leq 0)$ where Y and $\{X_t\}$ induce the general Wiener-Hopf equations of the form (1.5). This allows us to use the aforementioned technique to reformulate the Wiener-Hopf equations as a bi-infinite Toeplitz system, and thus obtain a solution to $H(\omega)$ as a deconvolution. The same technique is used to obtain an expression for entries of the inverse Toeplitz matrix $T(f)^{-1}$.

In practice, evaluating $H(\omega)$ in (1.4) for a general spectral density is infeasible. Typically, it is assumed that the spectral density is rational, which allows one to obtain a computationally tractable solution for $H(\omega)$. Of course, this leads to an approximation error in $H(\omega)$ when the underlying spectral density is not a rational function. In Section 3 we show that Baxter's inequality can be utilized to obtain a bound between $H(\omega)$ and its approximation based on a rational approximation of the general spectral density. The proofs of results in Sections 2 and 3 can be found in the Appendix.

2 A prediction approach

2.1 Notation and Assumptions

In this section, we collect together the notation introduced in Section 1 and some additional notation necessary for the paper.

Let $L_2([0, 2\pi))$ be the space of all square integral complex functions on $[0, 2\pi)$ and ℓ_2 is a space of all bi-infinite complex sequences $\mathbf{v} = (\ldots, v_{-1}, v_0, v_1, \ldots)$ where $\sum_{j\in\mathbb{Z}} |v_j|^2 < \infty$. Similarly, we denote $\ell_2^+ = \{\mathbf{v}_+ = (v_0, v_1, ...) : \sum_{j=0}^{\infty} |v_j|^2 < \infty\}$, a space of all semi-infinite square summable sequences. To connect the time and frequency domain through an isomorphism, we define the

Fourier transform $F : \ell_2 \to L_2([0, 2\pi))$

$$F(\mathbf{v})(\omega) = \sum_{j \in \mathbb{Z}} v_j e^{ij\omega}.$$

We define the semi- and bi-infinite Toeplitz matrices (operators) $T(f) = (c(t-\tau); t, \tau \geq 0)$ and $T_{\pm}(f) = (c(t-\tau); t, \tau \in \mathbb{Z})$ on ℓ_2^+ and ℓ_2 respectively. In this paper will make frequent use of the convolution theorem; if $\mathbf{h} \in \ell_2$, then $F(T_{\pm}(f)\mathbf{h})(\omega) = f(\omega)H(\omega)$, where $f(\omega) = \sum_{r \in \mathbb{Z}} c(r)e^{ir\omega}$ and $H(\omega) = F(\mathbf{h})(\omega)$. We will use the following assumptions.

Assumption 2.1 *Let $\{c(r) : r \in \mathbb{Z}\}$ be a symmetric positive definite sequence and $f(\omega) = \sum_{r \in \mathbb{Z}} c(r)e^{ir\omega}$ be its Fourier transform. Then,*

(i) $0 < \inf_\omega f(\omega) \leq \sup_\omega f(\omega) < \infty$.

(ii) For some $K > 1$ we have $\sum_{r \in \mathbb{Z}} |r^K c(r)| < \infty$.

Under Assumption 2.1(i), we have the unique factorization

$$f(\omega) = \sigma^2 |\psi(\omega)|^2 = \sigma^2 |\phi(\omega)|^{-2}, \tag{2.1}$$

where $\sigma^2 > 0$, $\psi(\omega) = 1 + \sum_{j=1}^{\infty} \psi_j e^{ij\omega}$ and $\phi(\omega) = (\phi(\omega))^{-1} = 1 - \sum_{j=1}^{\infty} \phi_j e^{ij\omega}$. We note that the characteristic polynomials $\Psi(z) = 1 + \sum_{j=1}^{\infty} \psi_j z^j$ and $\Phi(z) = 1 - \sum_{j=1}^{\infty} \phi_j z^j$ do not have zeroes in $|z| \leq 1$ thus the AR(∞) parameters are causal or equivalently are said to have minimum phase (see Szegö (1921) and Inoue (2000), pages 68-69).

We mention that Assumption 2.1(i) is used in all the results in this paper, whereas, Assumption 2.1(ii) is only required in the approximation theorem in Section 3.

2.2 Bivariate time series and the Wiener-Hopf equations

We now give an alternative formulation for the solution of (1.2) and (1.5), which utilizes properties of linear prediction to solve it using a standard deconvolution method. To integrate our derivation within the Wiener causal filter framework, we start with the classical Wiener filter. For $\{X_t : t \in \mathbb{Z}\}$ and $Y \in \mathcal{M}$, let

$$P_{\mathcal{H}_0}(Y) = \sum_{j=0}^{\infty} h_j X_{-j}. \tag{2.2}$$

We observe that by construction, (2.2) gives rise to the normal equations

$$\text{cov}(Y, X_{-\ell}) = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell \geq 0. \tag{2.3}$$

Since (2.3) only holds for positive ℓ, this prevents one using deconvolution to solve for $H(\omega)$. Instead, we define a "proxy" set of variables for $\{X_{-\ell} : \ell < 0\}$ such that (2.3) is valid for $\ell < 0$. By using the property of orthogonal projections, we have

$$\mathrm{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell})) = \mathrm{cov}(P_{\mathcal{H}_0}(Y), X_{-\ell}) \qquad \text{for} \qquad \ell < 0.$$

This gives

$$\mathrm{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell})) = \sum_{j=0}^{\infty} h_j \mathrm{cov}(X_{-j}, X_{-\ell}) = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell < 0. \tag{2.4}$$

Equations (2.3) and (2.4) allow us to represent the solution of $H(\omega)$ as a deconvolution. We define the semi- and bi-infinite sequences $\mathbf{c}_- = (\mathrm{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell})); \ell < 0)$, $\mathbf{c}_+ = (\mathrm{cov}(Y, X_{-\ell}); \ell \geq 0)$, and $\mathbf{c}_\pm = (\mathbf{c}_-, \mathbf{c}_+)$. Taking the Fourier transform of $\mathbf{c}_\pm$ and using the convolution theorem gives $F(\mathbf{c}_\pm)(\omega) = H(\omega)f(\omega)$. Thus

$$H(\omega) = \frac{F(\mathbf{c}_\pm)(\omega)}{f(\omega)} = \frac{\sum_{\ell=0}^{\infty} \mathrm{cov}(Y, X_{-\ell})e^{i\ell\omega} + \sum_{\ell=1}^{\infty} \mathrm{cov}\left(Y, P_{\mathcal{H}_0}(X_\ell)\right) e^{-i\ell\omega}}{f(\omega)}. \tag{2.5}$$

This forms the key to the following theorem.

Theorem 2.1 *Suppose $\{X_t\}$ is a stationary time series whose spectral density satisfies Assumption 2.1(i) and $Y \in \mathcal{M}$. Let $P_{\mathcal{H}_0}(Y) = \sum_{j=0}^{\infty} h_j X_{-j}$, then $(h_j; j \geq 0) \in \ell_2^+$, $(c_{YX}(\ell) = \mathrm{cov}(Y, X_{-\ell}); \ell \geq 0) \in \ell_2^+$ and*

$$H(\omega) = \frac{\sum_{\ell=0}^{\infty} c_{YX}(\ell)\big(e^{i\ell\omega} + \psi(\omega)^* \phi_\ell(\omega)^*\big)}{f(\omega)}, \tag{2.6}$$

where $\phi(\cdot)$ and $\psi(\cdot)$ are defined in (2.1) and $\phi_\ell(\omega) = \sum_{s=1}^{\infty} \phi_{\ell+s} e^{is\omega}$ for $\ell \geq 0$.

PROOF. See Appendix A. □

Remark 2.1 *It is clear that $\sum_{\ell=1}^{\infty} X_\ell e^{i\ell\omega}$ is not a well-defined random variable. However, it is interesting to note that under Assumption 2.1(ii) (for $K = 1$) $\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell)e^{-i\ell\omega}$ is a well defined random variable, where $\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell)e^{-i\ell\omega} \in \mathcal{H}_0$ and*

$$\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell)e^{-i\ell\omega} = \psi(\omega) \sum_{j=0}^{\infty} X_{-j}\phi_j(\omega). \tag{2.7}$$

In other words, despite $\sum_{\ell=1}^{\infty} X_\ell e^{i\ell\omega}$ not be well defined, informally its projection onto $\mathcal{H}_0$ does exist.

2.3 General Wiener-Hopf equations

We now generalize the above prediction approach to general Wiener-Hopf linear equations which satisfy

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell \geq 0, \tag{2.8}$$

where $\{g_\ell : \ell \geq 0\}$ and $\{c(r) : r \in \mathbb{Z}\}$ (which is assumed to be a symmetric, positive definite sequence) are known. We will obtain a solution similar to (2.6) but for the normal equations in (2.8). We first describe the classical Wiener-Hopf method to solve (2.8). Since $\{c(r)\}$ is known for all $r \in \mathbb{Z}$, we extend (2.8) to the negative index $\ell < 0$, and define $\{g_\ell : \ell < 0\}$ as

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \text{for} \qquad \ell < 0. \tag{2.9}$$

Note that $\{g_\ell : \ell < 0\}$ is not given, however it is completely determined by $\{g_\ell : \ell \geq 0\}$ and $\{c(r) : r \in \mathbb{Z}\}$ (this can be seen from equation (2.13), below). The Wiener-Hopf technique evaluates the Fourier transform of the above and isolates the positive frequencies to yield the solution for $H(\omega)$. Specifically, evaluating the Fourier transform of (2.8) and (2.9) gives

$$H(\omega)f(\omega) = G_-(\omega) + G_+(\omega) \tag{2.10}$$

where $G_-(\omega) = \sum_{\ell=-\infty}^{-1} g_\ell e^{i\ell\omega}$ and $G_+(\omega) = \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}$. Replacing $f(\omega)$ with $\sigma^2|\psi(\omega)|^2$ and dividing the above with $\sigma^2\psi(\omega)^*$ yields

$$H(\omega)\psi(\omega) = \frac{G_-(\omega)}{\sigma^2\psi(\omega)^*} + \frac{G_+(\omega)}{\sigma^2\psi(\omega)^*} = \sigma^{-2}\phi(\omega)^*G_-(\omega) + \sigma^{-2}\phi(\omega)^*G_+(\omega). \tag{2.11}$$

Isolating the positive frequencies in (2.11) gives the solution

$$H(\omega) = \sigma^{-2}\phi(\omega)[\phi(\omega)^*G_+(\omega)]_+, \tag{2.12}$$

this proves the result stated in (1.6). Similarly, by isolating the negative frequencies, we obtain $G_-(\omega)$ in terms of f and $G_+(\omega)$

$$G_-(\omega) = \sum_{\ell=-\infty}^{-1} g_\ell e^{i\ell\omega} = -\psi(\omega)^*[\phi(\omega)^*G_+(\omega)]_-. \tag{2.13}$$

Thus (2.12) and (2.13) yield an explicit solution for $H(\omega)$ and $G_-(\omega)$ respectively. However, from a time series perspective, it is difficult to interpret these formulas. We now obtain an alternative expression for these solutions based on a linear prediction of random variables.

We consider the matrix representation, $T(f)\mathbf{h}_+ = \mathbf{g}_+$, in (1.7). We solve $T(f)\mathbf{h}_+ = \mathbf{g}_+$ by embedding the semi-infinite Toeplitz matrix $T(f)$ on ℓ_2^+ into the bi-infinite Toeplitz system on ℓ_2. We divide the bi-infinite Toeplitz matrix $T_\pm(f)$ into four sub-matrices $C_{00} = (c(t-\tau); t, \tau < 0)$, $C_{10} = (c(t-\tau); t < 0, \tau \geq 0)$, $C_{01} = (c(t-\tau); t \geq 0, \tau < 0)$, and $C_{11} = (c(t-\tau); t, \tau \geq 0)$. We observe that $C_{11} = T(f)$. Further, we let $\mathbf{h}_\pm = (\mathbf{0}', \mathbf{h}_+')' = (\ldots, 0, 0, h_0, h_1, h_2, \ldots)'$ and $\mathbf{g}_\pm = (\mathbf{g}_-', \mathbf{g}_+')' = (\ldots, g_{-2}, g_{-1}, g_0, g_1, g_2, \ldots)'$ where $\mathbf{g}_- = C_{00}C_{11}^{-1}\mathbf{g}_+$. Then, we obtain the following bi-infinite Toeplitz system on ℓ_2

$$T_\pm(f)\mathbf{h}_\pm = \begin{pmatrix} C_{00} & C_{01} \\ C_{10} & C_{11} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{h}_+ \end{pmatrix} = \begin{pmatrix} C_{01}\mathbf{h}_+ \\ C_{11}\mathbf{h}_+ \end{pmatrix} = \begin{pmatrix} C_{01}C_{11}^{-1}\mathbf{g}_+ \\ \mathbf{g}_+ \end{pmatrix} = \begin{pmatrix} \mathbf{g}_- \\ \mathbf{g}_+ \end{pmatrix} = \mathbf{g}_\pm. \tag{2.14}$$

We note that the positive indices in the sequence $\mathbf{g}_\pm$ are $\{g_\ell : \ell \geq 0\}$, but for the negative indices, where $\ell < 0$, it is $g_\ell = [C_{01}C_{11}^{-1}\mathbf{g}_+]_\ell$ which is identical to g_ℓ defined in (2.9). The Fourier transform on both sides in (2.14) gives the deconvolution $f(\omega)H(\omega) = F(\mathbf{g}_\pm)(\omega)$, which is identical to (2.10). We now reformulate the above equation through the lens of prediction. To do this we define a stationary process $\{X_t\}$ and a random variable Y on the same probability space which yields (2.8) as their normal equations.

We first note that since $\{c(r) : r \in \mathbb{Z}\}$ is a symmetric, positive definite sequence, there exists a stationary time series $\{X_t\}$ with $\{c(r) : r \in \mathbb{Z}\}$ as its autocovariance function (see Brockwell and Davis (2006), Theorem 1.5.1). Using this we define the random variable

$$Y = \sum_{j=0}^{\infty} h_j X_{-j}. \tag{2.15}$$

Provided that $\mathbf{h}_+ \in \ell_2^+$, then $\mathbb{E}[Y^2] < \infty$ and thus Y belongs to the Hilbert space spanned by $\{X_j : j \leq 0\}$ (we show in Theorem 2.2 that this is true if $\mathbf{g}_+ \in \ell_2^+$). By (2.8), we observe that $\mathrm{cov}(Y, X_{-\ell}) = \sum_{j=0}^{\infty} h_j c(\ell - j) = g_\ell$ for all $\ell \geq 0$. We now show that for $\ell < 0$

$$\mathrm{cov}(Y, X_{-\ell}) = [C_{01}C_{11}^{-1}\mathbf{g}_+]_\ell = g_\ell.$$

First, since $Y \in \overline{\mathrm{sp}}(X_j : j \leq 0)$, then $\mathrm{cov}(Y, X_{-\ell}) = \mathrm{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell}))$. Further, for $\ell < 0$, the ℓth row (where we start the enumeration of the rows from the bottom) of $C_{01}C_{11}^{-1}$ contains the coefficients of the best linear predictor of $X_{-\ell}$ given $\{X_j : j \leq 0\}$ i.e.

$$P_{\mathcal{H}_0}(X_{-\ell}) = \sum_{j=0}^{\infty} [C_{01}C_{11}^{-1}]_{\ell,j} X_{-j} \qquad \ell < 0.$$

Using the above, we evaluate $\text{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell}))$ for $\ell < 0$

$$\begin{aligned}\text{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell})) &= \text{cov}\left(Y, \sum_{j=0}^{\infty}[C_{01}C_{11}^{-1}]_{\ell,j}X_{-j}\right) \\ &= \sum_{j=0}^{\infty}[C_{01}C_{11}^{-1}]_{\ell,j}\text{cov}(Y, X_{-j}) \qquad \text{(from (2.15), } g_j = \text{cov}(Y, X_{-j})) \\ &= \sum_{j=0}^{\infty}[C_{01}C_{11}^{-1}]_{\ell,j}g_j = [C_{01}C_{11}^{-1}\mathbf{g}_+]_\ell = g_\ell.\end{aligned}$$

Thus the entries of $\mathbf{g}'_\pm = (\mathbf{g}'_-, \mathbf{g}'_+)$ are indeed the covariances: $\mathbf{g}'_- = (\text{cov}(Y, P_{\mathcal{H}_0}(X_{-\ell})); \ell < 0)$ and $\mathbf{g}'_+ = (\text{cov}(Y, X_{-\ell}); \ell \geq 0)$. This allows us to use Theorem 2.1 to solve general Wiener-Hopf equations. Further, it gives an intuitive meaning to (2.9) and (2.14).

Theorem 2.2 *Suppose that $\{c(r) : r \in \mathbb{Z}\}$ is a symmetric, positive definite sequence and its Fourier transform $f(\omega) = \sum_{r\in\mathbb{Z}} c(r)e^{ir\omega}$ satisfies Assumption 2.1(i). We define the (semi) infinite system of equations*

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell \geq 0,$$

where $(g_\ell; \ell \geq 0) \in \ell_2^+$. Then, $\mathbf{h}'_+ \in \ell_2^+$ and

$$H(\omega) = \frac{\sum_{\ell=0}^{\infty} g_\ell\big(e^{i\ell\omega} + \psi(\omega)^*\phi_\ell(\omega)^*\big)}{f(\omega)}. \tag{2.16}$$

PROOF. See Appendix A. □

It is interesting to observe that the solution for $H(\omega)$ given in (2.12) was obtained by comparing the frequencies in a Fourier transform. Whereas the solution in Theorem 2.2 was obtained using linear prediction. The two solutions are algebraically different. We now show that they are the same by direct verification. Comparing the solutions (2.12) and (2.16) we have two different expressions for $H(\omega)$

$$H(\omega) = \sigma^{-2}\phi(\omega)[\phi(\omega)^*G_+(\omega)]_+ = \frac{F(\mathbf{g}_\pm)(\omega)}{f(\omega)} = \sigma^{-2}F(\mathbf{g}_\pm)(\omega)|\phi(\omega)|^2.$$

Therefore, the above are equivalent if

$$[\phi(\omega)^*G_+(\omega)]_+ = F(\mathbf{g}_\pm)(\omega)\phi(\omega)^*. \tag{2.17}$$

In the following lemma we prove the claim above by direct verification.

Lemma 2.1 *Suppose the same set of assumptions and notations in Theorem 2.2 hold. Then*

$$[\phi(\omega)^* G_+(\omega)]_+ = \phi(\omega)^* \sum_{\ell=0}^{\infty} g_\ell \left(e^{i\ell\omega} + \psi(\omega)^* \phi_\ell(\omega)^*\right)$$

where $G_+(\omega) = \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}$.

An interesting application of Theorem 2.2 is that it can be used to obtain an expression for $T(f)^{-1}$. As mentioned in Section 1, it is well-known that $T(f)^{-1} = \sigma^{-2}T(\phi)T(\phi)^*$. We show below that an alternative expression for the entries of $T(f)^{-1} = (d_{k,j}; k, j \geq 0)$ can be deduced using the deconvolution method described in Theorem 2.2.

Corollary 2.1 *Suppose the same set of assumptions and notations in Theorem 2.2 hold. Let* $\mathbf{d}_k = (d_{k,j}; j \geq 0)$ *denote the kth row of* $T(f)^{-1}$. *Then,* $\mathbf{d}_k \in \ell_2^+$ *for all* $k \geq 0$ *and the Fourier transfrom* $D_k(\omega) = F(\mathbf{d}_k)(\omega) = \sum_{j=0}^{\infty} d_{k,j} e^{ij\omega}$ *is*

$$D_k(\omega) = \frac{e^{ik\omega} + \psi(\omega)^* \phi_k(\omega)^*}{f(\omega)} \qquad k \geq 0.$$

Therefore,

$$d_{j,k} = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{e^{ik\omega} + \psi(\omega)^* \phi_k(\omega)^*}{f(\omega)} \right) e^{-ij\omega} d\omega \qquad j, k \geq 0.$$

PROOF. See Appendix A. □

Remark 2.2 (Multivariate extension) *The case that the (autocovariance) sequence* $\{\mathbf{C}(r) : r \in \mathbb{Z}\}$ *is made up of* $d \times d$*-dimensions, has not been considered in this paper. However, if* $\mathbf{\Sigma}(\omega) = \sum_{r\in\mathbb{Z}} \mathbf{C}(r) e^{ir\omega}$ *is a positive definite matrix with Vector MA*(∞) *and Vector AR*(∞) *representations, then it is may be possible to extend the above results to the multivariate setting.*

3 Finite order autoregressive approximations

In many applications it is often assumed the spectral density is rational (Cadzow (1982); Ahlén and Sternad (1991), and Ge and Kerrigan (2016)). Obtaining the spectral factorization of a rational spectral density (such as that given in (2.1)) is straightforward, and is one of the reasons that rational spectral densities are widely used. However, a rational spectral density is usually only an approximation of the underlying spectral density. In this section, we obtain a bound for the approximation when the rational spectral density corresponds to a finite order autoregressive process. We mention that using the expression in (2.16) easily lends itself to obtaining a rational approximation and for bounding the difference using Baxter's inequality.

We now use the expression in (2.16) to obtain an approximation of $H(\omega)$ in terms of a best fitting AR(p) coefficients. In particular, using that $\psi(\omega)^* = [\phi(\omega)^*]^{-1}$, we replace the infinite order AR coefficients in

$$H(\omega) = \frac{\sum_{\ell=0}^{\infty} g_\ell \big(e^{i\ell\omega} + [\phi(\omega)^*]^{-1}\phi_\ell(\omega)^*\big)}{f(\omega)} \tag{3.1}$$

with the best fitting AR(p) coefficients. More precisely, suppose that $(\phi_{p,1}, ..., \phi_{p,p})'$ are the best fitting AR(p) coefficients in the sense that it minimizes the mean squared error

$$(\phi_{p,1}, ..., \phi_{p,p})' = \arg\min_{\mathbf{a}} \|X_0 - \sum_{j=1}^{p} a_j X_{-j}\|^2 = \arg\min_{\mathbf{a}} \frac{1}{2\pi}\int_0^{2\pi} |1 - \sum_{j=1}^{p} a_j e^{ij\omega}|^2 f(\omega) d\omega \tag{3.2}$$

where $\mathbf{a} = (a_1, ..., a_p)'$. Let $\phi_p(\omega) = 1 - \sum_{j=1}^{p} \phi_{p,j} e^{ij\omega}$ and $f_p(\omega) = \sigma^2|\phi_p(\omega)|^{-2}$ where $\sigma^2 = \exp((2\pi)^{-1}\int_0^{2\pi} \log f(\omega) d\omega)$. We note that the zeros of $\phi_p(z) = 1 - \sum_{j=1}^{p} \phi_{p,j} z^j$ lie outside the unit circle. Then, we define the approximation of $H(\omega)$ as

$$H_p(\omega) = \frac{\sum_{\ell=0}^{\infty} g_\ell \big(e^{i\ell\omega} + [\phi_p(\omega)^*]^{-1}\phi_{p,\ell}(\omega)^*\big)}{f_p(\omega)}, \tag{3.3}$$

where $\phi_{p,\ell}(\omega) = \sum_{s=1}^{p-\ell} \phi_{p,\ell+s} e^{is\omega}$ for $0 \le \ell < p$ and 0 if $\ell \ge p$. We observe that the Fourier coefficients of $H_p(\omega)$ are the solution of $T(f_p)\mathbf{h}_p = \mathbf{g}_+$ where $\mathbf{h}_p = (h_{p,0}, h_{p,1}, \ldots)'$ with $h_{p,j} = (2\pi)^{-1}\int_0^{2\pi} H_p(\omega) e^{-ij\omega} d\omega$. Thus $T(f_p)$ and $T(f_p)^{-1}$ are approximations of $T(f)$ and $T(f)^{-1}$. Observe that by using Lemma 2.1 and (2.12) we can show that

$$H_p(\omega) = \sigma^{-2}\phi_p(\omega)[\phi_p(\omega)^* G_+(\omega)]_+ \tag{3.4}$$

Below we obtain a bound for $H(\omega) - H_p(\omega)$. It is worth noting that to obtain the bound we use the expressions for $H(\omega)$ and $H_p(\omega)$ given in (3.1) and (3.3), as these expression are easier to study than their equivalent expressions in (2.12) in (3.4).

Theorem 3.1 (Approximation theorem) *Suppose that $\{c(r) : r \in \mathbb{Z}\}$ is a symmetric, positive definite sequence that satisfies Assumption 2.1(ii) and its Fourier transform $f(\omega) = \sum_{r\in\mathbb{Z}} c(r) e^{ir\omega}$ satisfies Assumption 2.1(i). We define the (semi) infinite system of equations*

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell \ge 0,$$

where $(g_\ell; \ell \ge 0) \in \ell_2^+$. Let $H(\omega)$ and $H_p(\omega)$ be defined as in (3.1) and (3.3). Then

$$\big|H(\omega) - H_p(\omega)\big| \le C\left[p^{-K+1}\sup_s |g_s| + p^{-K}|G_+(\omega)|\right],$$

where $G_+(\omega) = \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}$.

PROOF. See Appendix A. $\square$

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgements

SSR and JY gratefully acknowledge the partial support of the National Science Foundation (grant DMS-1812054).

A Proofs

The purpose of this appendix is to give the technical details behind the results stated in the main section.

PROOF of Theorem 2.1 To prove that $\mathbf{h}_+ = (h_j; j \geq 0)' \in \ell_2^+$, we note that since $\mathbb{E}[Y^2] < \infty$, then $P_{\mathcal{H}_0}(Y) = \sum_{j=0}^{\infty} h_j X_{-j}$ is a well defined random variable where $P_{\mathcal{H}_0}(Y) \in \mathcal{H}_0$ with

$$\mathrm{var}[P_{\mathcal{H}_0}(Y)] = \langle \mathbf{h}_+, T(f)\mathbf{h}_+ \rangle.$$

Furthermore, $\lambda_{\inf}\|\mathbf{h}_+\|_2 \leq \langle \mathbf{h}_+, T(f)\mathbf{h}_+ \rangle \leq \lambda_{\sup}\|\mathbf{h}_+\|_2$ where $\lambda_{\inf} = \inf_{\|v\|_2=1, v\in\ell_2^+} T(f)v \geq \inf_\omega f(\omega)$ and $\lambda_{\sup} = \sup_{\|v\|_2=1, v\in\ell_2^+} T(f)v \leq \sup_\omega f(\omega)$. Since $\inf_\omega f(\omega) > 0$, this implies $\|\mathbf{h}_+\|_2 < \infty$.

To prove that $(c_{YX}(\ell); \ell \geq 0) \in \ell_2^+$, we recall that (2.2) leads to the matrix equation $\mathbf{c}_{YX} = T(f)\mathbf{h}_+$ where $\mathbf{c}_{YX} = (c_{YX}(0), c_{YX}(1), \ldots)'$. We define the operator norm $\|A\|_{spec} = \sup_{\|v\|_2=1, v\in\ell_2^+} \|Av\|_2$ and use the result that since $\sup_\omega f(\omega) < \infty$, then $\|T(f)\|_{spec} \leq \sup_\omega f(\omega)$. Thus $\|\mathbf{c}_{YX}\|_2 \leq \|T(f)\|_{spec}\|\mathbf{h}_+\|_2 \leq \sup_\omega f(\omega) \cdot \|\mathbf{h}_+\|_2 < \infty$, as required.

From (2.5), we have $H(\omega) = F(\mathbf{c}_\pm)(\omega)/f(\omega)$. We next express $F(\mathbf{c}_\pm)(\omega)$ in terms of an infinite order AR and MA coefficients of $\{X_t\}$. To do this we observe

$$F(\mathbf{c}_\pm)(\omega) = \sum_{\ell=0}^{\infty} c_{YX}(\ell)e^{i\ell\omega} + \sum_{\ell=1}^{\infty} \mathrm{cov}\big(Y, P_{\mathcal{H}_0}(X_\ell)\big)e^{-i\ell\omega}. \tag{A.1}$$

The second term on the right hand side of (A.1) looks quite wieldy. However, we show below that it can be expressed in terms of the AR(∞) coefficients associated with f. It is well-known that the ℓ-step ahead forecast $P_{\mathcal{H}_0}(X_\ell)$ ($\ell > 0$) has the representation $P_{\mathcal{H}_0}(X_\ell) = \sum_{j=0}^{\infty} \phi_j(\ell)X_{-j}$, where for $\ell > 0$, the ℓ-step prediction coefficients are

$$\phi_j(\ell) = \sum_{s=1}^{\ell} \phi_{j+s}\psi_{\ell-s} \tag{A.2}$$

and $\{\phi_j : j \geq 1\}$ and $\{\psi_j : j \geq 0\}$ are AR(∞) and MA(∞) coefficients defined in (2.1) (setting $\psi_0 = 1$). Since $Y \in \mathcal{M}$ and by the projection theorem $P_{\mathcal{H}_0}(X_\ell) \in \mathcal{H}_0$, we have $\left|\text{cov}(Y, P_{\mathcal{H}_0}(X_\ell))e^{-i\ell\omega}\right| < \infty$. We now obtain an expression for $\text{cov}(Y, P_{\mathcal{H}_0}(X_\ell))e^{-i\ell\omega}$ using (A.2)

$$\begin{aligned}\text{cov}(Y, P_{\mathcal{H}_0}(X_\ell))e^{-i\ell\omega} &= \text{cov}\left(Y, \sum_{j=0}^{\infty} X_{-j} \sum_{s=1}^{\ell} \phi_{j+s}\psi_{\ell-s}e^{-i\ell\omega}\right) \\ &= \sum_{s=1}^{\ell}\sum_{j=0}^{\infty} c_{YX}(j)\phi_{j+s}\psi_{\ell-s}e^{-i\ell\omega}.\end{aligned}$$

Note that we can exchange the summands in the last equation above due to Fubini's theorem:

$$\sum_{s=1}^{\ell} |\psi_{\ell-s}e^{-i\ell\omega}| \sum_{j=0}^{\infty} |c_{YX}(j)\phi_{j+s}| \leq \left(\sum_{s=1}^{\ell} |\psi_{\ell-s}|\right) \left(\sum_{j=0}^{\infty} c(j)^2\right)^{1/2} \left(\sum_{j=1}^{\infty} \phi_j^2\right)^{1/2} < \infty.$$

Therefore

$$\sum_{\ell=1}^{\infty} \text{cov}(Y, P_{\mathcal{H}_0}(X_\ell))e^{-i\ell\omega} = \sum_{\ell=1}^{\infty} \left(\sum_{s=1}^{\ell} \psi_{\ell-s} \sum_{j=0}^{\infty} c_{YX}(j)\phi_{j+s}\right) e^{-i\ell\omega}.$$

The Fourier coefficients of the right hand side of above has a convolution form, thus, we use the convolution theorem and rewrite

$$\begin{aligned}\sum_{\ell=1}^{\infty} \text{cov}(Y, P_{\mathcal{H}_0}(X_\ell))e^{-i\ell\omega} &= \sum_{\ell=1}^{\infty} \left(\sum_{s=1}^{\ell} \psi_{\ell-s} \sum_{j=0}^{\infty} c_{YX}(j)\phi_{j+s}\right) e^{-i\ell\omega} \\ &= \left(\sum_{\ell=0}^{\infty} \psi_\ell e^{-i\ell\omega}\right) \left(\sum_{s=1}^{\infty}\sum_{j=0}^{\infty} c_{YX}(j)\phi_{j+s}e^{-is\omega}\right) \\ &= \psi(\omega)^* \sum_{j=0}^{\infty} c_{YX}(j) \sum_{s=1}^{\infty} \phi_{j+s}e^{-is\omega} \\ &= \psi(\omega)^* \sum_{j=0}^{\infty} c_{YX}(j)\phi_j(\omega)^*\end{aligned}$$

where for $j \geq 0$, $\phi_j(\omega) = \sum_{s=1}^{\infty} \phi_{j+s} e^{is\omega}$. Substituting the above into $F(\mathbf{c}_{\pm})(\omega)$ gives

$$
\begin{aligned}
F(\mathbf{c}_{\pm})(\omega) &= \sum_{\ell=0}^{\infty} c_{YX}(\ell) e^{i\ell\omega} + \psi(\omega)^* \sum_{j=0}^{\infty} \phi_j(\omega)^* c_{YX}(j) \\
&= \sum_{\ell=0}^{\infty} c_{YX}(\ell) \big(e^{-i\ell\omega} + \psi(\omega)^* \phi_\ell(\omega)^* \big). \qquad \text{(A.3)}
\end{aligned}
$$

Note it is easily seen that $F(\mathbf{c}_{\pm})(\omega) \in L_2([0, 2\pi))$ since $\psi(\omega)^*$ is bounded. Finally, substituting the above into $H(\omega) = F(\mathbf{c}_{\pm})(\omega)/f(\omega)$ proves the result. □

Proof of Remark 2.1 By substituting (A.2) into $\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell) e^{i\ell\omega}$ gives

$$
\begin{aligned}
\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell) e^{i\ell\omega} &= \sum_{\ell=1}^{\infty} \left(\sum_{j=0}^{\infty} X_{-j} \sum_{s=1}^{\ell} \phi_{j+s} \psi_{\ell-s} \right) e^{i\ell\omega} \\
&= \sum_{j=0}^{\infty} X_{-j} \sum_{s=1}^{\infty} \phi_{j+s} e^{is\omega} \sum_{\ell=s}^{\infty} \psi_{\ell-s} e^{i(\ell-s)\omega} \\
&= \psi(\omega) \sum_{j=0}^{\infty} X_{-j} \sum_{s=1}^{\infty} \phi_{j+s} e^{is\omega} = \psi(\omega) \sum_{j=0}^{\infty} X_{-j} \phi_j(\omega).
\end{aligned}
$$

We show that if Assumption 2.1(ii) is satisfied for $K = 1$, then the right hand side $\psi(\omega) \sum_{j=0}^{\infty} X_{-j} \phi_j(\omega)$ converges in $\mathcal{H}_0$. To show this, we define the partial sum

$$
S_n = \psi(\omega) \sum_{j=0}^{n} X_{-j} \phi_j(\omega) \in \mathcal{H}_0.
$$

Then, for any $n < m$

$$
\|S_m - S_n\|^2 = \| \sum_{j=n}^{m} \psi(\omega) \phi_j(\omega) X_{-j} \|^2 = \text{var} \left(\sum_{j=n}^{m} \psi(\omega) \phi_j(\omega) X_{-j} \right) = |\psi(\omega)|^2 (\phi_n^m(\omega))' T_{m-n}(f) (\phi_n^m(\omega))
$$

where $(\phi_n^m(\omega)) = (\phi_n(\omega), ..., \phi_m(\omega))'$ and $T_{m-n}(f) = (c(t-\tau); 0 \leq t, \tau \leq m-n)$. Therefore,

$$
\begin{aligned}
\|S_m - S_n\|^2 &= |\psi(\omega)|^2 (\phi_n^m(\omega))' T_{m-n}(f) (\phi_n^m(\omega)) \\
&\leq |\psi(\omega)|^2 \|\phi_n^m(\omega)\|_2^2 \|T_{m-n}(f)\|_{spec} \\
&\leq |\psi(\omega)|^2 (\sup_{\omega} f(\omega)) \|\phi_n^m(\omega)\|_2^2
\end{aligned}
$$

If Assumption 2.1(ii) is satisfied for $K = 1$, then it is easy to show $\sum_{j=0}^{\infty} |\phi_j(\omega)|^2 < \infty$. Therefore, by Cauchy's criterion, $\|\phi_n^m(\omega)\|_2 \to 0$ as $n, m \to \infty$, which implies $\| \sum_{j=n}^{m} \psi(\omega) \phi_j(\omega) X_{-j} \|^2 \to 0$ as $n, m \to \infty$. Thus, using the Cauchy's criterion again, $\psi(\omega) \sum_{j=0}^{\infty} X_{-j} \phi_j(\omega)$ converges in $\mathcal{H}_0$.

This shows $\sum_{\ell=1}^{\infty} P_{\mathcal{H}_0}(X_\ell)e^{i\ell\omega}$ is well-defined in $\mathcal{H}_0$ and satisfies (2.7). □

PROOF of Theorem 2.2 We first prove that $\mathbf{h}_+ \in \ell_2^+$. Under Assumption 2.1(i), it is known that $T(f)$ is invertible on ℓ_2^+ (see Widom (1960), Theorem III). Using that $\|T(f)^{-1}\|_{sp} \leq [\inf_\omega f(\omega)]^{-1}$ we have

$$\|\mathbf{h}_+\|_2 \leq \|T(f)^{-1}\|_{sp}\|\mathbf{g}_+\|_2 \leq [\inf_{\omega} f(\omega)]^{-1}\|\mathbf{g}_+\|_2 < \infty$$

since $\mathbf{g}_+ \in \ell_2^+$. Thus $\mathbf{h}_+ \in \ell_2^+$ and its Fourier transform $H(\omega)$ is well-defined. We use this result to show that there always exists a random variable Y and time series $\{X_t\}$ whose normal equations satisfy

$$g_\ell = \sum_{j=0}^{\infty} h_j c(\ell - j) \qquad \ell \geq 0.$$

This allows us to use Theorem 2.1 to prove the result.

First, since $\{c(r)\}$ is a symmetric, positive definite sequence (or equivalently $f(\omega) = \sum_{r\in\mathbb{Z}} c(r)e^{ir\omega}$ is positive), by Kolmogorov's extension theorem, there exists a weakly stationary random process $\{X_t\}$ which has the autocovariance function $\{c(r)\}$ (see Brockwell and Davis (2006), Theorem 1.5.1). We define the Hilbert space $\mathcal{H}_0 = \overline{\text{sp}}(X_j : j \leq 0)$ and random variable $Y = P_{\mathcal{H}_0}(Y) = \sum_{j=0}^{\infty} h_j X_{-j}$. Since $\{h_j : j \geq 0\} \in \ell_2^+$ then $Y \in \mathcal{H}_0$. Then, by definition, $\text{cov}(Y, X_{-\ell}) = \sum_{j=0}^{\infty} h_j c(\ell - j) = g_\ell$ for all $\ell \geq 0$. This connects the equation to the random variables.

Finally, we follow the proof of Theorem 2.1 to obtain the result. □

PROOF of Lemma 2.1 By (2.16), $F(\mathbf{g}_\pm)(\omega)\phi(\omega)^* = \sum_{\ell=0}^{\infty} g_\ell \left(e^{i\ell\omega}\phi(\omega)^* + \phi_\ell(\omega)^*\right)$. Expanding the right hand side gives

$$\begin{aligned} F(\mathbf{c}_\pm)(\omega)\phi(\omega)^* &= \sum_{\ell=0}^{\infty} g_\ell \left(e^{i\ell\omega}\phi(\omega)^* + \phi_\ell(\omega)^*\right) \\ &= \phi(\omega)^* \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega} + \sum_{\ell=0}^{\infty} g_\ell \phi_\ell(\omega)^* \\ &= -\left(\sum_{j=0}^{\infty} \phi_j e^{-ij\omega}\right)\left(\sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}\right) + \sum_{\ell=0}^{\infty} g_\ell \left(\sum_{s=1}^{\infty} \phi_{s+\ell} e^{-is\omega}\right), \end{aligned}$$

where we set $\phi_0 = -1$. Therefore, the k-th Fourier coefficient of $F(\mathbf{c}_\pm)(\omega)\phi(\omega)^*$ is

$$-\sum_{a,b\geq 0,\ b-a=k} \phi_a g_b + \delta_{k<0} \cdot \sum_{\ell=0}^{\infty} g_\ell \phi_{\ell-k}, \tag{A.4}$$

where $\delta_{k<0} = 1$ if $k < 0$ and 0 otherwise. Depending on the sign of k, (A.4) has different expressions:

$\boldsymbol{k<0}$: Using that $\{(a,b): a,b \geq 0, b-a=k\} = \{(b-k,b): b\geq 0\}$, we get

$$-\sum_{a,b\geq 0,\ b-a=k} \phi_a g_b + \delta_{k<0}\cdot \sum_{\ell=0}^{\infty} g_\ell \phi_{\ell-k} = -\sum_{b=0}^{\infty}\phi_{b-k}g_b + \sum_{\ell=0}^{\infty} g_\ell\phi_{\ell-k} = 0. \tag{A.5}$$

$\boldsymbol{k\geq 0}$: The second term of (A.4) vanishes and using that $\{(a,b): a,b\geq 0, b-a=k\} = \{(a,a+k): a\geq 0\}$, we get

$$-\sum_{a,b\geq 0,\ b-a=k} \phi_a g_b + \delta_{k<0}\cdot \sum_{\ell=0}^{\infty} g_\ell \phi_{\ell-k} = -\sum_{a=0}^{\infty}\phi_a g_{a+k}. \tag{A.6}$$

By inspection, we observe that the left hand side of (A.6) is a lag k cross-correlation between $(...,0,-\phi_0,-\phi_1,...)$ and $(...,0,g_0,g_1,...)$. Therefore, using the convolution theorem, it is the k-th Fourier coefficient of $(\sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega})(\sum_{\ell=0}^{\infty} -\phi_\ell e^{-i\ell\omega}) = G_+(\omega)\phi(\omega)^*$ (for $k\geq 0$). Lastly, combining (A.5) and (A.6), we get

$$\begin{aligned} F(\mathbf{g}_\pm)(\omega)\phi(\omega)^* &= \sum_{k=-\infty}^{\infty}\left(-\sum_{a,b\geq 0,\ b-a=k}\phi_a g_b + \delta_{k<0}\cdot\sum_{\ell=0}^{\infty} g_\ell\phi_{\ell-k}\right)e^{ik\omega} \\ &= \sum_{k=0}^{\infty}\left(-\sum_{a=0}^{\infty}\phi_a g_{a+k}\right)e^{ik\omega} = [G_+(\omega)\phi(\omega)^*]_+. \end{aligned}$$

This proves the lemma. □

Proof of Corollary 2.1 Let $\delta_{\ell,k}$ denote the indicator variable where $\delta_{\ell,k}=1$ if $\ell=k$ and zero otherwise. Since $T(f)^{-1} = (d_{j,k}; j,k\geq 0)$ is the inverse of $T(f) = (c(j-k); j,k\geq 0)$, $\{d_{j,k}\}$ and $\{c(r)\}$ satisfy the normal equations

$$\delta_{\ell,k} = \sum_{j=0}^{\infty} d_{j,k}c(\ell-j) \qquad \ell,k\geq 0. \tag{A.7}$$

Thus for each k, we have a system of Wiener-Hopf equations. To find $d_{j,k}$ we apply Theorem 2.2. Thus for each (fixed) $k\geq 0$ we obtain

$$D_k(\omega) = \frac{1}{f(\omega)}\sum_{\ell=0}^{\infty}\delta_{\ell,k}\big(e^{i\ell\omega} + \psi(\omega)^*\phi_\ell(\omega)^*\big) = \frac{e^{ik\omega}+\psi(\omega)^*\phi_k(\omega)^*}{f(\omega)}, \tag{A.8}$$

where $D_k(\omega) = \sum_{j=0}^{\infty} d_{j,k} e^{ij\omega}$. Inverting the Fourier transform yields the entries

$$d_{j,k} = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{e^{ik\omega} + \psi(\omega)^*\phi_k(\omega)^*}{f(\omega)}\right)e^{-ij\omega}d\omega. \tag{A.9}$$

Thus proving the result.

As an aside it is interesting to construct the random variable Y_{-k} which yields the Wiener-Hopf equation (A.7). Since $\{c(r)\}$ forms a (symmetric) positive definite sequence, there exists a stationary time series $\{X_t\}$ with $\{c(r)\}$ as its autocovariance. We define a sequence of random variables $\{\varepsilon_{-k} : k \geq 0\}$ where for $k \geq 0$

$$\varepsilon_{-k} = X_{-k} - P_{(-k)^c}(X_{-k})$$

and $P_{(-k)^c}$ denotes the orthogonal projection onto the closed subspace $\overline{\text{sp}}(X_r : r \leq 0 \text{ and } r \neq -k)$. We standardize ε_{-k}, where $Y_{-k} = \varepsilon_{-k}/\sqrt{\text{var}[\varepsilon_{-k}]}$, noting that $\text{var}[\varepsilon_{-k}] = \text{cov}[\varepsilon_{-k}, X_{-k}]$. Thus by definition $\text{cov}[Y_{-k}, X_\ell] = \delta_{\ell,k}$ and $Y_{-k} = \sum_{j=0}^{\infty} d_{j,k} X_{-j}$. □

PROOF of Theorem 3.1 We note that under Assumption 2.1(ii), $\sum_{r\in\mathbb{Z}} |r^K c(r)| < \infty$. This condition implies $\sum_{j=1}^{\infty} |j^K \phi_j| < \infty$ (see Kreiss et al. (2011), Lemma 2.1).

To prove the result, we use Baxter's inequality, that is for the best fitting AR(p) approximation of $f(\omega)$ (see equation (3.2)) we have

$$\sum_{j=1}^{p} |\phi_{p,j} - \phi_j| \leq C \sum_{j=p+1}^{\infty} |\phi_j| \tag{A.10}$$

where C is a constant that depends on $f(\omega) = \sigma^2|\phi(\omega)|^{-2}$.

Returning to the proof, the difference $H(\omega) - H_p(\omega)$ can be decomposed as

$$\begin{aligned} H(\omega) - H_p(\omega) &= \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}\left(\frac{1}{f(\omega)} - \frac{1}{f_p(\omega)}\right) + \sum_{\ell=0}^{\infty} g_\ell\left(\frac{[\phi(\omega)^*]^{-1}\phi_\ell(\omega)^*}{f(\omega)} - \frac{[\phi_p(\omega)^*]^{-1}\phi_{p,\ell}(\omega)^*}{f_p(\omega)}\right) \\ &= \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}\left(\frac{1}{f(\omega)} - \frac{1}{f_p(\omega)}\right) + \sigma^{-2}\sum_{\ell=0}^{\infty} g_\ell\left[\phi(\omega)\phi_\ell(\omega)^* - \phi_p(\omega)\phi_{p,\ell}(\omega)^*\right] \\ &= A(\omega) + B(\omega) + C(\omega) \end{aligned}$$

where

$$\begin{aligned} A(\omega) &= \left(\frac{1}{f(\omega)} - \frac{1}{f_p(\omega)}\right)\sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega} \\ B(\omega) &= \sigma^{-2}\left[\phi(\omega) - \phi_p(\omega)\right]\sum_{\ell=0}^{\infty} g_\ell \phi_\ell(\omega)^* \\ C(\omega) &= \sigma^{-2}\phi_p(\omega)\sum_{\ell=0}^{\infty} g_\ell\left[\phi_\ell(\omega)^* - \phi_{p,\ell}(\omega)^*\right]. \end{aligned}$$

We bound each term above. First, we bound $A(\omega)$. Note that

$$\begin{aligned} \frac{1}{f(\omega)} - \frac{1}{f_p(\omega)} &= \sigma^{-2}\left(|\phi(\omega)|^2 - |\phi_p(\omega)|^2\right) \\ &= \sigma^{-2}\left[(\phi(\omega) - \phi_p(\omega))\phi(\omega)^* + (\phi(\omega)^* - \phi_p(\omega)^*)\phi_p(\omega)\right]. \end{aligned}$$

Using (A.10), we have $|\phi(\omega)| \leq \sum_{j=1}^{\infty}|\phi_j| < \infty$,

$$\begin{aligned} |\phi(\omega) - \phi_p(\omega)| &= |\sum_{j=1}^{p}(\phi_j - \phi_{p,j})e^{ij\omega} + \sum_{j=p+1}^{\infty}\phi_j e^{ij\omega}| \\ &\leq \sum_{j=1}^{p}|\phi_j - \phi_{p,j}| + \sum_{j=p+1}^{\infty}|\phi_j| \leq C\sum_{j=p+1}^{\infty}|\phi_j|, \end{aligned} \tag{A.11}$$

and

$$|\phi_p(\omega)| \leq |\phi_p(\omega) - \phi(\omega)| + |\phi(\omega)| \leq C\sum_{j=p+1}^{\infty}|\phi_j| + \sum_{j=1}^{\infty}|\phi_j| < \infty.$$

Therefore,

$$\left|\frac{1}{f(\omega)} - \frac{1}{f_p(\omega)}\right| \leq \sigma^{-2}\left(|\phi(\omega) - \phi_p(\omega)||\phi(\omega)^*| + |\phi(\omega)^* - \phi_p(\omega)^*||\phi_p(\omega)|\right) \leq C\sum_{j=p+1}^{\infty}|\phi_j|.$$

Furthermore, by Assumption 2.1(ii),

$$\sum_{j=p+1}^{\infty}|\phi_j| \leq p^{-K}\sum_{j=p+1}^{\infty}|j^K\phi_j| < Cp^{-K}. \tag{A.12}$$

Therefore, substituting (A.12) into $A(\cdot)$ gives

$$|A(\omega)| \leq C\left(\sum_{j=p+1}^{\infty}|\phi_j|\right)\cdot|\sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}| \leq Cp^{-K}\big|G_+(\omega)\big|$$

where $G_+(\omega) = \sum_{\ell=0}^{\infty} g_\ell e^{i\ell\omega}$. To bound $B(\omega)$ we note that

$$\sum_{\ell=0}^{\infty} |g_\ell \phi_\ell(\omega)| \leq \sum_{\ell=0}^{\infty} |g_\ell| \sum_{s=1}^{\infty} |\phi_{\ell+s}| = \sum_{u=1}^{\infty} \sum_{\ell=0}^{u-1} |g_\ell||\phi_u| \leq \sup_s |g_s| \cdot \sum_{u=1}^{\infty} |u\phi_u|.$$

Thus by using the above, (A.11), and (A.12) we have

$$\sup_\omega |B(\omega)| \leq C \sup_s |g_s| \cdot \sum_{u=1}^{\infty} |u\phi_u| \cdot \sum_{j=p+1}^{\infty} |\phi_j| = O\left(\sup_s |g_s| \cdot p^{-K}\right).$$

Finally, to bound $C(\omega)$, by using (A.10) we have

$$\begin{aligned}\sum_{\ell=0}^{\infty} |g_\ell| \, |\phi_\ell(\omega) - \phi_{p,\ell}(\omega)| &\leq C \sum_{\ell=0}^{\infty} \sum_{s=p+1}^{\infty} |g_\ell||\phi_{s+\ell}| \leq C \sup_s |g_s| \sum_{\ell=0}^{\infty} |\ell\phi_{p+\ell}| \\ &\leq C \sup_s |g_s| \sum_{\ell=p+1}^{\infty} |\ell\phi_\ell| = O\left(\sup_s |g_s| \cdot p^{-K+1}\right).\end{aligned}$$

Altogether, this yields the bound

$$\left|H(\omega) - H_p(\omega)\right| \leq C \left[p^{-K+1} \cdot \sup_s |g_s| + p^{-K} \cdot \left|G_+(\omega)\right|\right].$$

This proves the result. □

References

A. Ahlén and M. Sternad. Wiener filter design using polynomial equations. *IEEE Transactions on Signal Processing*, 39(11):2387–2399, 1991.

H. W. Bode and C. E. Shannon. A simplified derivation of linear least square smoothing and prediction theory. *Proceedings of the IRE*, 38(4):417–425, 1950.

Peter J. Brockwell and Richard A. Davis. *Time series: theory and methods*. Springer Series in Statistics. Springer, New York, NY, 2006. Reprint of the second (1991) edition.

J. A. Cadzow. Spectral estimation: An overdetermined rational model equation approach. *Proceedings of the IEEE*, 70(9):907–939, 1982.

J. Durbin. The fitting of time-series models. *Revue de l'Institut International de Statistique*, 28 (3):233–244, 1960.

M. Ge and E. C. Kerrigan. Short-term ocean wave forecasting using an autoregressive moving average model. In *2016 UKACC 11th International Conference on Control (CONTROL)*, pages 1–6. IEEE, 2016.

Eberhard Hopf. *Mathematical problems of radiative equilibrium.* Canbridge University Press, Cambridge, 1934.

A. Inoue. Asymptotics for the partial autocorrelation function of a stationary process. *Journal d'Analyse Mathematique*, 81(1):65–109, 2000.

T. Kailath. A view of three decades of linear filtering theory. *IEEE Transactions on information theory*, IT-20:146–181, 1974.

Thomas Kailath. *Linear systems.* Prentice-Hall Englewood Cliffs, NJ, 1980. Prentice-Hall Information and System Sciences Series.

J. Krampe, J.-P. Kreiss, and E. Paparoditis. Estimated Wold representation and spectral-density-driven bootstrap for time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80:703–726, 2018.

J.-P. Kreiss, E. Paparoditis, and D. N. Politis. On the range of validity of the autoregressive sieve bootstrap. *Annals of Statistics*, 39(4):2103–2130, 2011.

J. B. Lawrie and I. D. Abrahams. A brief historical perspective of the Wiener-Hopf technique. *Journal of Engineering Mathematics*, 59:351–358, 2007.

N. Levinson. The Wiener (root mean square) error criterion in filter design and prediction. *Journal of Mathematics and Physics*, 25(1-4):261–278, 1947.

Sophocles J. Orfanidis. *Optimum Signal Processing: An introduction.* McGraw-Hill, New York, NY, 2018.

Mohsen Pourahmadi. *Foundations of time series analysis and prediction theory*, volume 379. John Wiley & Sons, Hoboken, NJ, 2001.

S. Subba Rao and J. Yang. Reconciling the Gaussian and Whittle likelihood with an application to estimation in the frequency domain. *To appear in the Annals of Statistics*, 2021.

G. Szegö. Über die randwerte einer analytischen funktion. *Mathematische Annalen*, 84:232–244, 1921.

H. Widom. Inversion of Toeplitz matrices II. *Illinois Journal of Mathematics*, 4(1):88–99, 1960.

N. Wiener and E. Hopf. Über eine klasse singulärer integralgleichungen. *Sitzungsber. Preuss. Akad. Wiss. Berlin*, 31:696–706, 1931.

Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press, Cambridge, MA, 1949.

L. A. Zadeh and J. R. Ragazzini. An extension of Wiener's theory of prediction. *Journal of Applied Physics*, 21(7):645–655, 1950.