Forward-Backward Hankel Matrix Fitting for Spectral Super-Resolution

Zai Yang and Xunmeng Wu
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
yangzai@xjtu.edu.cn

***Abstract*—Hankel-based approaches form an important class of methods for line spectral estimation within the recent spectral super-resolution framework. However, they suffer from the fundamental limitation that their estimated signal poles do not lie on the unit circle in general, causing difficulties of physical interpretation and performance loss. In this paper, we present a modified Hankel approach called forward-backward Hankel matrix fitting (FB-Hankel) that can be implemented by simply modifying the existing algorithms. We show analytically that the new approach has great potential to restrict the estimated poles on the unit circle. Numerical results are provided that corroborate our analysis and demonstrate the advantage of FB-Hankel in improving the estimation accuracy.** [1]

***Index Terms*—Forward-backward Hankel matrix fitting, line spectral estimation, spectral super-resolution, Vandermonde decomposition, Kronecker's theorem.**

I. Introduction

Line spectral estimation refers to the process of estimating the spectral modes of a signal given its discrete samples. In particular, we have the uniformly sampled data sequence $\{y_j\}$ that is given by

$$y_j = \sum_{k=1}^{K} s_k z_k^j + e_j, \quad |z_k| = 1, \tag{1}$$

where the poles $z_k = e^{i2\pi f_k}$, $k = 1, \ldots, K$ lie on the unit circle and have a one-to-one connection to the unknown frequencies $f_k \in [0, 1)$, $k = 1, \ldots, K$, with $i = \sqrt{-1}$. In (1), $s_k \in \mathbb{C}$ denotes the amplitude of the kth component and $e_j \in \mathbb{C}$ is noise.

Line spectral estimation has a long history of research, and a great number of approaches have been developed, see [1]. With the development of compressed sensing and sparse signal processing, Candès and many other researchers have established the framework of spectral super-resolution, known also as gridless sparse methods, for line spectral estimation by utilizing the signal sparsity and optimization theory in the past decade, see [2]. Roughly speaking, these methods follow two lines of research, positive-semidefinite- (PSD-) Toeplitz-based and Hankel-based. In particular, PSD-Toeplitz-based methods are rooted on the exact data model in (1) and a PSDness constraint is introduced to capture the knowledge $|z_k| = 1$, see [3]–[5]. Differently, Hankel-based methods drop the prior knowledge $|z_k| = 1$ so that the PSDness constraint is absent, which eventually results in simple algorithm design [6]–[8]. However, this causes a fundamental limitation of Hankel-based methods, to be specific, the estimated poles do not lie on the unit circle in general, resulting in difficulties of physical interpretation and potential performance loss (see the main context). This paper aims at resolving this fundamental problem.

In this paper, by using the classic forward-backward technique [9], we propose a modified Hankel-based approach for line spectral estimation, named as forward-backward Hankel matrix fitting (FB-Hankel). The core of FB-Hankel is introducing a (conjugated) backward Hankel matrix which is composed of the same spectral modes as and processed jointly with the original (forward) Hankel matrix. We show analytically that the FB-Hankel can substantially shrink the solution space to approach the ground truth and has great potential to push the estimated poles to the unit circle. An algorithm is presented to solve the FB-Hankel optimization problem by simply modifying the algorithm in [8]. Numerical simulations are presented that demonstrate the great capabilities of the proposed approach in restricting the estimated poles on the unit circle and improving the estimation accuracy.

Notation: The set of real and complex numbers are denoted $\mathbb{R}$ and $\mathbb{C}$ respectively. Boldface letters are reserved for vectors and matrices. The complex conjugate and amplitude of scalar a are denoted $\overline{a}$ and $|a|$. The transpose of matrix $\boldsymbol{A}$ is denoted $\boldsymbol{A}^T$. The jth entry of vector $\boldsymbol{x}$ is x_j. The diagonal matrix with vector $\boldsymbol{x}$ on the diagonal is denoted $\mathrm{diag}\,(\boldsymbol{x})$. The rank of matrix $\boldsymbol{A}$ is denoted $\mathrm{rank}\,(\boldsymbol{A})$.

Organization: The data model is restated and some preliminaries are presented in Section II. The FB-Hankel approach is presented in Section III. Numerical simulations are provided in Section IV, and conclusions are drawn in Section V.

II. Preliminaries

A. Hankel-Based Methods

Without loss of generality, we assume the sample size $N = 4n+1$ and the sample index $j \in \{-2n, \ldots, 2n\}$.[2] Let $\boldsymbol{a}(z) = [z^{-n}, \ldots, z^n]^T$ that represents a sampled complex exponential of size $2n+1$ with pole $z \in \mathbb{C}$. It is a sinusoidal wave if $|z| = 1$. In the absence of noise, the samples $\{y_j\}$ can be

[1]The research of the project was supported by the National Natural Science Foundation of China under Grants 11922116 and 61977053.

[2]Otherwise, we can add up to three more samples such that $N = 4n+1$ that then are dealt with as missing data.

used to form the $(2n+1)\times(2n+1)$ square Hankel matrix

$$\begin{aligned}\mathcal{H}(\boldsymbol{y}) &= \begin{bmatrix} y_{-2n} & y_{-2n+1} & \cdots & y_0 \\ y_{-2n+1} & y_{-2n+2} & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_0 & y_1 & \cdots & y_{2n} \end{bmatrix} \\ &= \sum_{k=1}^{K} s_k \boldsymbol{a}(z_k)\boldsymbol{a}^T(z_k) \\ &= \boldsymbol{A}(\boldsymbol{z})\,\mathrm{diag}(\boldsymbol{s})\,\boldsymbol{A}^T(\boldsymbol{z}),\end{aligned} \tag{2}$$

where $\boldsymbol{A}(\boldsymbol{z}) = [\boldsymbol{a}(z_1),\ldots,\boldsymbol{a}(z_k)]$ is an $(2n+1)\times K$ Vandermonde matrix. Therefore, if $K \le 2n$, then $\boldsymbol{A}(\boldsymbol{z})$ has rank K, so does $\mathcal{H}(\boldsymbol{y})$. The Hankel-based methods do line spectral estimation by recovering the low-rank Hankel matrix $\mathcal{H}(\boldsymbol{y})$ from the samples.

B. Forward-Backward Processing

Consider the conjugated backward data sequence $\{\overline{y}_{-j} : j = -2n,\ldots,2n\}$ that in the absence of noise is given by

$$\overline{y}_{-j} = \sum_{k=1}^{K} \overline{s}_k \overline{z}_k^{-j} = \sum_{k=1}^{K} \overline{s}_k z_k^j \tag{3}$$

since $|z_k| = 1$. Therefore, the sequence $\{\overline{y}_{-j}\}$ corresponds to a virtual signal that is also composed of the poles $\{z_k\}$. The joint operation on $\{y_j\}$ and $\{\overline{y}_{-j}\}$ for line spectral estimation is called forward-backward processing that has achieved great success in array signal processing [9]–[12].

III. Forward-Backward Hankel Matrix Fitting

A. A General Framework of Spectral Super-Resolution

In this subsection, we present a general framework of spectral super-resolution by exploiting signal sparsity. Assume that the number of poles K is known. Let

$$S_0 = \left\{ \boldsymbol{x} \in \mathbb{C}^N : \; x_j = \sum_{k=1}^{K} s_k z_k^j,\; s_k \in \mathbb{C},\; |z_k| = 1 \right\} \tag{4}$$

that denotes the set of candidate signals composed of at most K sinusoids. To do line spectral estimation, the framework of spectral super-resolution tells us to find some candidate signal $\boldsymbol{x} \in S_0$ that is nearest to its noisy (or incomplete) measurement $\boldsymbol{y} = [y_{-2n},\ldots,y_{2n}]^T$ and then retrieve the estimated poles from $\boldsymbol{x}$. To be specific, we attempt to solve the following optimization problem:

$$\min_{\boldsymbol{x}} \mathrm{Loss}(\boldsymbol{x}-\boldsymbol{y}), \;\text{ subject to } \boldsymbol{x} \in S_0, \tag{5}$$

where $\mathrm{Loss}(\cdot)$ denotes a loss function that is chosen according to the sampling environment (e.g., noise, missing data, etc). The key underlies how to characterize the constraint $\boldsymbol{x} \in S_0$.

An exact characterization of the constraint $\boldsymbol{x} \in S_0$ is provided in [4], in terms of the atomic ℓ_0 norm, by introducing a PSD Toeplitz matrix. The resulting optimization problem and its variants form the PSD-Toeplitz-based methods.

Let

$$S_1 = \left\{ \boldsymbol{x} \in \mathbb{C}^N : \; \mathrm{rank}\left(\mathcal{H}(\boldsymbol{x})\right) \le K \right\} \tag{6}$$

that includes S_0. By changing S_0 to S_1, the original problem (5) is then relaxed to the following Hankel matrix optimization model:

$$\min_{\boldsymbol{x}} \mathrm{Loss}(\boldsymbol{x}-\boldsymbol{y}), \;\text{ subject to } \boldsymbol{x} \in S_1. \tag{7}$$

The optimization model in (7) and its variants result in the Hankel-based methods.

B. Forward-Backward Hankel Matrix Fitting

Let $\widetilde{\boldsymbol{x}} = [\overline{x}_{2n},\ldots,\overline{x}_{-2n}]^T$ be the conjugated backward version of vector $\boldsymbol{x} \in \mathbb{C}^N$. Let also

$$S_2 = \left\{ \boldsymbol{x} \in \mathbb{C}^N : \; \mathrm{rank}\left([\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})]\right) \le K \right\} \tag{8}$$

that is a subset of S_1. By changing S_0 to S_2, we propose the optimization model

$$\min_{\boldsymbol{x}} \mathrm{Loss}(\boldsymbol{x}-\boldsymbol{y}), \;\text{ subject to } \boldsymbol{x} \in S_2 \tag{9}$$

that is called forward-backward Hankel matrix fitting (FB-Hankel). Correspondingly, the model (7) is referred to as forward-only Hankel matrix fitting (FO-Hankel).

For any $\boldsymbol{x} \in S_0$, by making use of (2) and (3), we have that

$$\mathcal{H}(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{z})\,\mathrm{diag}(\boldsymbol{s})\,\boldsymbol{A}^T(\boldsymbol{z}), \tag{10}$$

$$\mathcal{H}(\widetilde{\boldsymbol{x}}) = \boldsymbol{A}(\boldsymbol{z})\,\mathrm{diag}(\overline{\boldsymbol{s}})\,\boldsymbol{A}^T(\boldsymbol{z}). \tag{11}$$

It follows that

$$[\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})] = \boldsymbol{A}(\boldsymbol{z})\left[\mathrm{diag}(\boldsymbol{s})\,\boldsymbol{A}^T(\boldsymbol{z}) \mid \mathrm{diag}(\overline{\boldsymbol{s}})\,\boldsymbol{A}^T(\boldsymbol{z})\right] \tag{12}$$

and thus $\mathrm{rank}\left([\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})]\right) \le K$. This implies that $\boldsymbol{x} \in S_2$ and thus

$$S_0 \subset S_2 \subset S_1. \tag{13}$$

As a relaxation of (5), consequently, the FB-Hankel model in (9) is at least as tight as the FO-Hankel model in (7) and is expected to have higher accuracy. Indeed, the FB-Hankel model is a tighter relaxation, which is shown in the ensuing subsection.

C. Analysis

We show in this subsection that, unlike the PSD-Toeplitz-based methods, the estimated poles of FO-Hankel and FB-Hankel may not lie on the unit circle. Compared to FO-Hankel, however, the proposed FB-Hankel has better capability in restricting the estimated poles on the unit circle. In particular, we let

$$S_1' = \left\{ \boldsymbol{x} \in \mathbb{C}^N : \; x_j = \sum_{k=1}^{K} s_k z_k^j,\; s_k, z_k \in \mathbb{C} \right\} \tag{14}$$

that is obtained by removing the constraint $|z_k| = 1$ in S_0. Evidently, $S_0 \subset S_1' \subset S_1$. Moreover, the Kronecker's theorem [13] says that if $\boldsymbol{x} \in S_1$ and $K \le 2n$, then $\boldsymbol{x} \in S_1'$ except for degenerate cases. Therefore, S_1 and S_1' are approximately identical if $K \le 2n$. This implies that the FO-Hankel in (7) can be viewed as a relaxation of (5) by removing the constraint $|z_k| = 1$.

We now consider the FB-Hankel. Assume $\boldsymbol{x} \in S_2$ and $K \leq 2n$, i.e., rank $([\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})]) \leq K \leq 2n$. It follows that rank $(\mathcal{H}(\boldsymbol{x})) \leq K$ and rank $(\mathcal{H}(\widetilde{\boldsymbol{x}})) \leq K$. The Kronecker's theorem implies that $\boldsymbol{x}, \widetilde{\boldsymbol{x}} \in S_1'$ almost surely, to be specific, there exist $\{s_k, z_k \in \mathbb{C}\}_{k=1}^K$ and $\{s_k', z_k' \in \mathbb{C}\}_{k=1}^K$ with $z_k, z_k' \neq 0$ such that for $j = -2n, \ldots, 2n$,

$$x_j = \sum_{k=1}^{K} s_k z_k^j, \tag{15}$$

$$\overline{x}_{-j} = \sum_{k=1}^{K} s_k' {z'}_k^j, \tag{16}$$

yielding that

$$x_j = \sum_{k=1}^{K} s_k z_k^j = \sum_{k=1}^{K} \overline{s}_k' {z'}_k^{\star j}, \quad j = -2n, \ldots, 2n, \tag{17}$$

where $z^\star = \overline{z}^{-1}$ for $z \in \mathbb{C}$. Using (17) and the fact that an $m \times n$ Vandermonde matrix has rank $\min(m, n)$, it can be shown by simple arguments that $\{(z_k, s_k) : s_k \neq 0\}$ and $\left\{\left({z'}_k^\star, \overline{s}_k'\right) : s_k' \neq 0\right\}$ must be identical. Without loss of generality, we assume that

$$\begin{aligned} z_k &= {z'}_k^\star, \quad s_k = \overline{s}_k', \quad k = 1, \ldots, K_0 \leq K, \\ s_k &= \overline{s}_k' = 0, \quad k = K_0 + 1, \ldots, K. \end{aligned} \tag{18}$$

Using (15), (16), (2) and (18), we obtain

$$\begin{aligned} &[\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})] \\ &= \left[\boldsymbol{A}(\boldsymbol{z}) \operatorname{diag}(\boldsymbol{s}) \boldsymbol{A}^T(\boldsymbol{z}) \mid \boldsymbol{A}(\boldsymbol{z}') \operatorname{diag}(\boldsymbol{s}') \boldsymbol{A}^T(\boldsymbol{z}')\right] \\ &= \left[\boldsymbol{A}(\boldsymbol{z}) \operatorname{diag}(\boldsymbol{s}) \boldsymbol{A}^T(\boldsymbol{z}) \mid \boldsymbol{A}(\boldsymbol{z}^\star) \operatorname{diag}(\overline{\boldsymbol{s}}) \boldsymbol{A}^T(\boldsymbol{z}^\star)\right] \\ &= [\boldsymbol{A}(\boldsymbol{z}) \mid \boldsymbol{A}(\boldsymbol{z}^\star)] \begin{bmatrix} \operatorname{diag}(\boldsymbol{s}) & \\ & \operatorname{diag}(\overline{\boldsymbol{s}}) \end{bmatrix} \begin{bmatrix} \boldsymbol{A}(\boldsymbol{z}) & \\ & \boldsymbol{A}(\boldsymbol{z}^\star) \end{bmatrix}^T. \end{aligned} \tag{19}$$

It then can be shown that rank $([\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})])$ equals the number of distinct poles in $\{z_k\}_{k=1}^{K_0} \cup \{z_k^\star\}_{k=1}^{K_0}$ that is no greater than K by assumption. Since $z_k = z_k^\star$ if and only if $|z_k| = 1$, we have $\boldsymbol{x} \in S_2'$, where

$$S_2' = \{\boldsymbol{x} \in S_1' \text{ where either } |z_k| = 1 \text{ or } z_k, z_k^\star \text{ appear in pair}\}. \tag{20}$$

It can easily be shown that $S_2' \subset S_2$. Therefore, S_2' and S_2 are approximately identical. This implies that the FB-Hankel can be viewed as a relaxation of (5) by allowing appearance of pairs of poles $\{z_k, z_k^\star\}$ if they are not on the unit circle, which substantially shrinks the solution space of the FO-Hankel. Note that if we write $z_k = r_k e^{i2\pi f_k}$ in a polar coordinate, then $z_k^\star = r_k^{-1} e^{i2\pi f_k}$ and thus they share an identical frequency, implying that they correspond to a single frequency component. Intuitively, this occurs only when two frequencies are closely located so that they cannot be separated. If the frequencies $\{f_k\}$ are properly separated, it is expected that all of the estimated poles of FB-Hankel lie on the unit circle.

Finally, note that if $K = 1$, it can easily be shown that

$$S_0 = S_2' = S_2 \subset S_1' \subset S_1, \tag{21}$$

and thus the FB-Hankel is equivalent to solving the original model (5) but FO-Hankel is not.

D. Algorithm

In this subsection, we show that the proposed FB-Hankel model can be solved similarly to the FO-Hankel by simply modifying existing algorithms for the latter. As an illustration, we consider the noisy, full data observation model and the least square loss function and thus the two Hankel matrix rank minimization problems that we need to solve are given by:

$$\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2, \text{ subject to rank}(\mathcal{H}(\boldsymbol{x})) \leq K, \tag{22}$$

$$\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2, \text{ subject to rank}(\mathcal{H}_{\text{FB}}(\boldsymbol{x})) \leq K, \tag{23}$$

where $\mathcal{H}_{\text{FB}}(\boldsymbol{x}) = [\mathcal{H}(\boldsymbol{x}) \mid \mathcal{H}(\widetilde{\boldsymbol{x}})]$.

We consider the iterative hard thresholding (IHT) algorithm that has been widely used for compressed sensing and low-rank matrix recovery [14], [15]. It was adapted in [6], [8] to the FO-Hankel model in (22) and an accelerated version of it was also presented in [8]. The IHT for FO-Hankel is illustrated in Algorithm 1. In particular, the signal is initialized by the noisy measurement $\boldsymbol{y}$ and so we have $\boldsymbol{x}_0$. Then gradient descent is used to update the signal estimate in each iteration with step size $\alpha_t = \frac{1}{\sqrt{t}}$. After the Hankel matrix $\boldsymbol{D}_t$ is formed, the hard thresholding operator $\boldsymbol{\Gamma}_K$ is applied to obtain the best rank-K approximation, $\boldsymbol{G}_{t+1}$, of $\boldsymbol{D}_t$ by setting its singular values but the largest K to zero. Eventually we get the new estimate $\boldsymbol{x}_{t+1}$ by applying the pseudoinverse of the Hankel operator, $\mathcal{H}^\dagger$, on $\boldsymbol{G}_{t+1}$.

Note that the above IHT algorithm for the FO-Hankel (22) can be adapted to the FB-Hankel (23) by simply changing the Hankel operator $\mathcal{H}$ to the FB-Hankel operator $\mathcal{H}_{\text{FB}}$.

Once the algorithm is converged and produces the signal estimate $\hat{\boldsymbol{x}}$, it is expected that $\hat{\boldsymbol{x}}$ is a feasible solution and thus rank $(\mathcal{H}(\hat{\boldsymbol{x}})) \leq K$. Then, the ESPRIT algorithm [16] can be applied on $\mathcal{H}(\hat{\boldsymbol{x}})$ to obtain its Vandermonde decomposition as in (2) which provides the estimated poles $\{\hat{z}_k\}$.

Algorithm 1 The IHT algorithm

Initialize: $\boldsymbol{x}_0 = \boldsymbol{y}$ and $\boldsymbol{G}_0 = \mathcal{H}(\boldsymbol{y})$
for $t = 1, \ldots$ **do**
  $\boldsymbol{D}_t = \mathcal{H}(\boldsymbol{x_t} + \alpha_t(\boldsymbol{y} - \boldsymbol{x_t}))$
  $\boldsymbol{G}_{t+1} = \boldsymbol{\Gamma}_K(\boldsymbol{D}_t)$
  $\boldsymbol{x}_{t+1} = \mathcal{H}^\dagger(\boldsymbol{G_{t+1}})$
end for

IV. Numerical Simulations

In this section, we provide numerical results to demonstrate the performance of the proposed FB-Hankel model and algorithm as compared to the FO-Hankel. In our simulation, we consider $K = 3$ sinusoids that are randomly generated without or with a minimum separation $\frac{4}{N}$ (referred to as *random* and *spaced* frequencies respectively), with amplitudes $0.5 + |w|$ and random phases, where w follows the standard

TABLE I
The three failures of FB-Hankel with random frequencies in Fig.1.

Ground truth of poles	Estimated poles of FB-Hankel	Estimated poles of FO-Hankel
$e^{i2\pi\{0.7568,0.5952,0.6019\}}$	$\{1.0000e^{i2\pi 0.7583}, \mathbf{1.0155}e^{i2\pi\mathbf{0.5970}}, \mathbf{1.0155}^{-1}e^{i2\pi\mathbf{0.5970}}\}$	$\{0.9948e^{i2\pi 0.7583}, 1.0466e^{i2\pi 0.5951}, 0.9900e^{i2\pi 0.3786}\}$
$e^{i2\pi\{0.0733,0.2271,0.2420\}}$	$\{1.0000e^{i2\pi 0.0730}, \mathbf{1.0291}e^{i2\pi\mathbf{0.2331}}, \mathbf{1.0291}^{-1}e^{i2\pi\mathbf{0.2331}}\}$	$\{0.9989e^{i2\pi 0.0729}, 0.9739e^{i2\pi 0.2241}, 1.0774e^{i2\pi 0.2449}\}$
$e^{i2\pi\{0.5708,0.9943,0.9819\}}$	$\{1.0000e^{i2\pi 0.2233}, \mathbf{1.0331}e^{i2\pi\mathbf{0.9862}}, \mathbf{1.0331}^{-1}e^{i2\pi\mathbf{0.9862}}\}$	$\{0.9501e^{i2\pi 0.7540}, 0.9604e^{i2\pi 0.9883}, 1.0352e^{i2\pi 0.9831}\}$

normal distribution. A number of $N = 65$ uniform samples are acquired with random noise added that follows a complex normal distribution. The IHT algorithm is terminated if $\|\boldsymbol{x}_{t+1} - \boldsymbol{x}_t\|_2 / \|\boldsymbol{x}_t\|_2 < 10^{-5}$ or maximally 3000 iterations are reached. We say that an algorithm successfully produces poles $\{\hat{z}_k\}$ on the unit circle in a single trial if their average distance to the unit circle $\frac{1}{K}\sum_{k=1}^{K} ||\hat{z}_k| - 1| < 10^{-4}$.

In our first simulation, we consider the noiseless case and the noisy case with SNR = 0dB and run 1000 trials for each. The number of successful trials in which the estimated poles lie on the unit circle is shown in Fig. 1. In the absence of noise, both the FO- and FB-Hankel approaches produce the signal estimate $\hat{\boldsymbol{x}} = \boldsymbol{y}$ and thus the real poles, as shown in Fig. 1. In the presence of noise, it can be seen that the FO-Hankel cannot produce poles on the unit circle, as expected. In contrast to this, the proposed FB-Hankel approach rarely fails. In fact, if the frequencies are properly separated, the FB-Hankel succeeds in all trials to restrict the estimated poles on the unit circle.

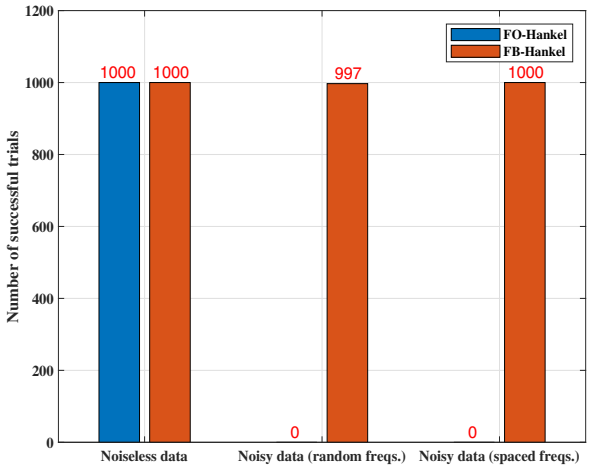

Fig. 1. Histogram of the number of successful trials of FO- and FB-Hankel with SNR = ∞ and 0dB and a total number of 1000 trials.

To study what happens when the FB-Hankel fails to produce poles on the unit circle, we present in Table I the estimated poles of the three failure trials shown in Fig. 1. It can be seen that all the failures occur when two frequencies are very close to each other with distance less than $\frac{1}{N}$. Their estimated poles are located symmetrically on two sides of the unit circle, verifying our analysis.

Finally, we study the average distance of the estimated poles to the unit circle with respect to the SNR that ranges from 0 to 40dB. The simulation results averaged over 1000 trials at each SNR level are presented in Fig. 2. It can be seen that at all SNR levels the proposed FB-Hankel can restrict the estimated poles on the unit circle (within numerical precision) with few failures that occur only when closely located frequencies are present and pull upwards the curve of FB-Hankel (with random frequencies). We also compute the relative signal reconstruction error measured as $\frac{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^o\|_2^2}{\|\boldsymbol{x}^o\|_2^2}$ that is then averaged over all trials at each SNR level, where $\boldsymbol{x}^o$ denotes the ground truth of noiseless signal. At all SNR levels, the FB-Hankel has consistently smaller signal reconstruction error than the FO-Hankel, with an improvement of 1.3dB on average in both cases of random and spaced frequencies.

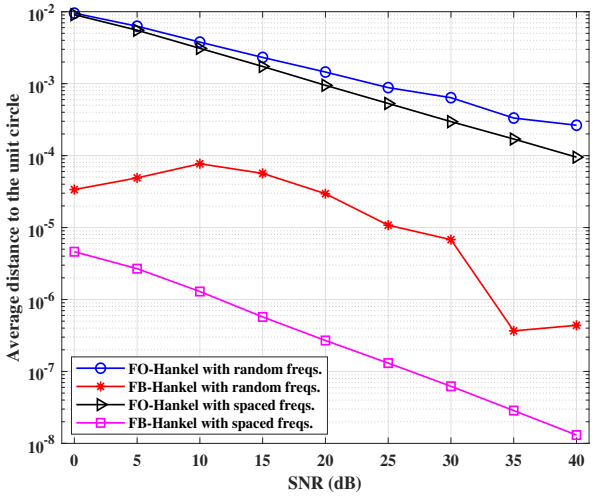

Fig. 2. Average distance of the estimated poles of FO- and FB-Hankel to the unit circle with respect to the SNR.

V. Conclusion

In this paper, the FB-Hankel approach was proposed by modifying the existing Hankel methods for line spectral estimation. We showed that, unlike the existing Hankel methods, the FB-Hankel has great potential to push the estimated poles to the unit circle. Extensive numerical simulations were provided that verify our analysis and demonstrate the good performance of FB-Hankel in restricting the poles on the unit circle and improving the estimation accuracy.

REFERENCES

[1] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Upper Saddle River, NJ, US: Pearson/Prentice Hall, 2005.

[2] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," *Academic Press Library in Signal Processing Volume 7 (R. Chellappa and S. Theodoridis,* Eds.*)*, pp. 509–581, 2018.

[3] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[4] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[5] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3139–3153, 2015.

[6] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6576–6601, 2014.

[7] F. Andersson, M. Carlsson, J.-Y. Tourneret, and H. Wendt, "A new frequency estimation method for equally and unequally spaced data," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5761–5774, 2014.

[8] J.-F. Cai, T. Wang, and K. Wei, "Fast and provable algorithms for spectrally sparse signal reconstruction via low-rank Hankel matrix completion," *Applied and Computational Harmonic Analysis*, vol. 46, no. 1, pp. 94–121, 2019.

[9] J. E. Evans, J. R. Johnson, and D. Sun, "Application of advanced signal processing techniques to angle of arrival estimation in ATC navigation and surveillance systems," Lincoln Laboratory, Tech. Rep., 1982.

[10] R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul, "An improved spatial smoothing technique for bearing estimation in a multipath environment," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 425–432, 1988.

[11] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 8–15, 1989.

[12] Z. Yang, P. Stoica, and J. Tang, "Source resolvability of spatial-smoothing-based subspace methods: A Hadamard product perspective," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2543–2553, 2019.

[13] R. Rochberg, "Toeplitz and Hankel operators on the Paley-Wiener space," *Integral Equations and Operator Theory*, vol. 10, no. 2, pp. 187–235, 1987.

[14] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[15] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[16] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1340–1342, 1986.