Low-Rank Hankel Signal Model: Numerical Results

Lucas Abdalah, Walter Freitas Jr, Pedro Marinho R de Oliveira, Vicente Zarzoso

Abstract—Hankel matrices arise in several applications of signal processing, such as tensor decompositions, biomedical signal processing, etc. In general, these techniques rely on digital signals that can be modeled as a linear combination of exponential polynomials. Hence, the Hankel matrix built from these signals presents full rank, equal to the number of poles, ensured under mild constraints. The present work observes other features of the low-rank Hankel model, using singular value decomposition (SVD) to assess rank deficiency. The effects observed may impact blind source separation (BSS) problems.

Keywords— Vandermonde Decomposition, Hankel Matrix, Singular Value Decomposition, Low-Rank.

I. INTRODUCTION

One of the challenges in signal processing is the estimation of a set of source signals based on the observation of a set of mixed signals. A blind source separation (BSS) approach can be applied aiming to isolate the source signals and usually relies on matrix decompositions. Nevertheless, strong mathematical constraints are necessary to assure uniqueness of the decomposition, such as mutual orthogonality between spatial factors and statistical independence [1], [2]. The best-known matrix methods are based on principal and independent component analysis. In order to overcome these limitations, tensor decompositions are powerful tools, ensuring uniqueness and outperforming matrix-based methods under milder constraints [3].

A celebrated tensor-based method is the block term decomposition (BTD) [4], [5]. This technique takes advantage of discrete-time signals that can be modeled as linear combinations of exponentials (all-pole models). Since the sources can be expressed as low-rank Hankel matrices, the signal separation can be performed via BTD. The Hankel matrix built from such source signals accepts the Vandermonde decomposition [6], a well-known result in problems of matrix analysis and telecommunications. Different fields in signal processing take advantage on low-rank Hankel source models for speech analysis [7], black-box polynomial analysis [8], and tensor decompositions [4], some applied to biomedical signal analysis [5], [9].

The theoretical baseline relies on a strong assumption: since a Vandermonde matrix generated by distinct poles has full rank equal to the number of poles [4], the Hankel matrix mapped from a finite signal with enough samples presents its rank equal to the number of poles. This notion can be used to assess

Lucas Abdalah and Walter Freitas Jr are with Universidade Federal do Ceará, Wireless Telecom Research Group, Fortaleza-CE, {lucas.abdalah, walter}@gtel.ufc.br; Pedro Marinho R de Oliveira is with BioSerenity, Paris, France, pedro.marinho@bioserenity.com; Vicente Zarzoso is with Université Côte d'Azur, CNRS, I3S Laboratory, Sophia Antipolis, France, vicente.zarzoso@univ-cotedazur.fr. This work was partially funded by the National Council for Scientific and Technological Development (CNPq).

signal complexity, since the more poles a signal contains, the more complex it can be considered. The rank of a matrix can be computed with a popular and reliable method based on the singular value decomposition (SVD), that considers the rank as the number of singular values larger than a tolerance parameter [10].

However, the present work shows that numerical features such as the distance between the poles that generate the signal may impact the computed rank. Aiming to observe clearly the behavior described, we propose three different experiments: 1) Vandermonde matrix built from two vectors, to illustrate how the distance between poles hampers the construction of a full-rank matrix, and consequently how it impacts the singular values and the Hankel matrix rank; 2) multiple signals with various distances between poles vs. the amount of samples necessary to obtain the Hankel matrix rank equal to the number of poles contained in the signal; 3) scenario using the previous framework but with noisy signals with different signal-to-noise ratio (SNR).

The rest of this work is structured as follow: Section II presents the problem statement, whereas numerical experiments are presented in Section III. The discussions and conclusions based on the experimental results are given in Sections IV and V, respectively.

II. PROBLEM STATEMENT

A. Low-Rank Hankel Model

Assume that if a discrete time signal $s(n)$ is a linear combination of L damped complex exponentials, say:

$$s(n) = \sum_{l=1}^{L} c_l z_l^n, \quad 0 \le n \le N-1 \qquad (1)$$

where c_l, $z_l \in \mathbb{C}$, are coefficients and poles, respectively.

The sequence $s(n)$ can be mapped onto an $(M \times M)$ Hankel matrix denoted $\mathbf{H}_s$, where each sample is placed along the anti-diagonal of $\mathbf{H}_s$ as shown in Fig. 1, with mapping function $\mathbf{H}_{s_{(i,j)}} = s(i+j-1)$, for ith row, and jth column. We assume N is odd without loss of generality, hence:

$$M = \frac{N+1}{2}. \qquad (2)$$

A major result in signal processing states that $\mathbf{H}_s$ accepts the following Vandermonde decomposition [4], [6]:

$$\mathbf{H}_s = \mathbf{V}_s \, \mathbf{D} \, \mathbf{V}_s^{\mathsf{T}} \qquad (3)$$
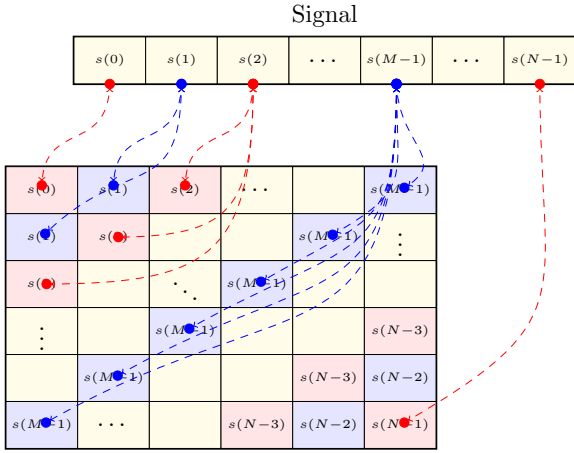
Fig. 1. Visual representation of a Hankel matrix $\mathbf{H}_s$ built from a signal $s(n)$.

where $\mathbf{V}_s$ is the Vandermonde matrix

$$\mathbf{V}_s = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_L \\ \vdots & \vdots & & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_L^{M-1} \end{bmatrix} \in \mathbb{C}^{M \times L} , \quad (4)$$

and $\mathbf{D} = \mathrm{diag}(c_1, c_1, \dots, c_L) \in \mathbb{C}^{L \times L}$ is a diagonal matrix. Symbol $(\cdot)^\intercal$ denotes the matrix transpose operator. We assume, for simplicity, but without loss of generality, that $c_l = 1$, $l = 1, 2, \dots, L$. As result of the Vandermonde decomposition (4), matrix $\mathbf{H}_s$ has rank at most $\min\{L, M\}$. Hence, a first remark is that the more poles a given signal is composed of, the higher the rank of its Hankel matrix. This observation underlies the use of $rank(\mathbf{H}_s)$ as a measure of signal complexity. Nevertheless, a Hankel matrix build from a signal $s(n)$ with L poles needs to map at least N_{min} samples:

$$N_{min} = 2L - 1 , \quad (5)$$

to present its rank R equal to L [4]–[6]. As soon as this minimum threshold is observed, the rank R should be equal to the number of poles L for a signal following model (1), regardless of the sample size.

B. Vandermonde Matrix and Pole Distance

When a signal is a sum of complex exponentials (1), its poles are closely linked to the columns of matrix $\mathbf{V}_s$, according to eqn. (4). If the distance between these poles decreases, the columns of $\mathbf{V}_s$ become closer to each other, having an impact on the Hankel matrix rank.

To illustrate this behavior, we assume two poles $z_1 = e^{j\omega_1}$ and $z_2 = e^{j\omega_2}$, and the corresponding Vandermonde vectors $\mathbf{v}_1 = [1, e^{j\omega_1}, e^{j2\omega_1}, \dots, e^{j(M-1)\omega_1}]^\intercal$ and $\mathbf{v}_2 = [1, e^{j\omega_2}, e^{j2\omega_2}, \dots, e^{j(M-1)\omega_2}]^\intercal$.

We can observe that each vector norm is equal to M, where

$$||\mathbf{v}_1|| = ||\mathbf{v}_2|| = M$$

and the columns scalar product leads to:

$$\mathbf{v}_1^\intercal \mathbf{v}_2 = \sum_{n=0}^{M-1} e^{-jn\omega_1} e^{jn\omega_2} .$$

Finally, replacing $\Delta\omega = (\omega_2 - \omega_1)$ and the exponential identities we have that the scalar product becomes:

$$\cos(\theta) = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{||\mathbf{v}_1|| \, ||\mathbf{v}_2||} = \frac{\sin(M\frac{\Delta\omega}{2})}{M \sin(\frac{\Delta\omega}{2})} . \quad (6)$$

Based on the previous result, we can ensure that when $\Delta\omega$ tends to zero for a fixed M, we have that:

$$\lim_{\Delta\omega \to 0} \cos(\theta) = 1$$

showing that the columns becomes colinear, as expected, if the distance between poles ($\Delta\omega$) become too small. In this case, if we compare the singular values, the Hankel matrix rank is equal to one. This shows mathematically that a minimum distance between poles is required in practical scenarios. In this case, the equivalent Hankel matrix rank (R) is equal to one, built from two damped exponentials, resulting in $R \neq L$.

However, it is relevant to notice the presence of M in the denominator of eqn. (6). We can deduce that increasing the value of M used to build $\mathbf{V}_s$ may compensate for the poles proximity. If we replace M using eqn. (2), we have that:

$$\cos(\theta) = \frac{2\sin((N+1)\Delta\omega)}{(N+1)\sin(\frac{\Delta\omega}{2})} .$$

Therefore, for a small yet nonzero value of $\Delta\omega$, but nonzero, if we increase N then the scalar-product gets closer to zero (due to N in the numerator), in such a way that colinearity between $\mathbf{v}_1$ and $\mathbf{v}_2$ is reduced:

$$\lim_{N \to \infty} \frac{2\sin((N+1)\Delta\omega)}{(N+1)\sin(\frac{\Delta\omega}{2})} = 0 \quad (7)$$

resulting in $R = L$, in agreement with the classical result observed in the literature [6].

C. Singular Value Decomposition

The SVD is a valuable tool in signal processing, a common technique for multivariate data analysis and provides a foundation for many other techniques such as principal component analysis (PCA) [11].

The SVD of $\mathbf{X} \in \mathbb{C}^{I \times J}$ is given by:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H , \quad (8)$$

where $\mathbf{U} \in \mathbb{C}^{I \times I}$ and $\mathbf{V} \in \mathbb{C}^{J \times J}$ are unitary matrices with orthonormal columns, and the symbol $(\cdot)^H$ denotes the Hermitian operator. Diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{I \times J}$ contains real, non-negative entries. It can also be represented as $\mathbf{\Sigma} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_J)$, where each λ_j is a singular value, presented in decreasing order.

In order to compute the rank of $\mathbf{X}$, a reliable SVD-based method may be applied [12], [13]. It consists in performing $\mathrm{SVD}(\mathbf{X})$, then taking the number of singular values $(\lambda_1, \dots, \lambda_J)$ that are larger than a tolerance $\varepsilon = \max(I, J)||\mathbf{X}||\mathcal{E}$, where $\mathcal{E}$ is the minimum distance that the

floating point arithmetic can recognize between two numbers and $||\mathbf{X}||$ represents the frobenius norm of $\mathbf{X}$.

Another possible way to assess rank-deficiency is by the condition number (σ), taking advantage of the relationship between singular values. The criterion used in this work is the equation:

$$\sigma = \frac{\lambda_{max}}{\lambda_{min}} \tag{9}$$

that provides a quotient between its maximum and minimum singular values, respectively. A large condition number is linked to a matrix close to rank-deficiency.

III. EXPERIMENTAL RESULTS

The experiments are divided in 3 parts. The first and second parts consider noiseless signals, whereas the third part is on noisy signals. In the first part we build a Vandermonde matrix from two vectors to illustrate how the distance between poles in the vector hamper the construction of a full-rank matrix, and consequently how it impacts the condition number, as theoretically anticipated in sec. II-B.

In the second part, we test various distances between poles to assess the minimum amount of samples necessary to obtain the Hankel matrix rank equal to the number of poles contained in the signal.

Finally, in the third scenario noisy signals are analyzed with various SNR values.

A. Condition Number

We assess the relationship between the condition number (σ) and the number of samples (N) varying the distance between the poles ($\Delta\omega$). An example inspired on the theoretical problem statement is assuming two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ as presented in Section II-B, with $\omega_1 = 0$, and $\omega_2 = \Delta\omega$, i.e., $\mathbf{v}_1 = [1, 1, 1, \ldots 1]^\mathsf{T}$ and $\mathbf{v}_2 = [1, e^{j\Delta\omega}, e^{j2\Delta\omega}, \ldots, e^{j(M-1)\Delta\omega}]^\mathsf{T}$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the columns of the Vandermonde matrix $\mathbf{V}_s$. Hence, the Hankel matrix $\mathbf{H}_s \in \mathbb{C}^{N \times N}$ associated with this model is built from a simplified form of eqn. (3), where $\mathbf{V}_s \in \mathbb{C}^{N \times 2}$ and $\mathbf{D} \in \mathbb{R}^{2 \times 2}$ is an identity matrix. After constructing the Hankel matrix, we compute the SVD($\mathbf{H}_s$) for 10 different values of N, rounding logarithmically spaced values in the interval $[10^1, 10^2]$, and considering 7 different values of $\Delta\omega$ as shown in Fig. 2.

In the first place, if the distances between poles are so small, then even increasing the number of samples, σ keeps very high. This behavior affects rank computation and may prevent us to obtain $R = L$. Fig. 2 also shows that σ decreases as we increase the number of samples to the columns, which compensates for the poles proximity as expected from the analysis of Section II-B.

Moreover, for $\Delta\omega = [8.91 \times 10^{-9}, 1.96 \times 10^{-9}]$, σ indicates that $\lambda_1 \gg \lambda_2$ for all values of N, but it presents $R \neq L$ until reaching enough samples. Nevertheless, as N increases, σ tends to 1, i.e., the number of samples compensates for the insufficient distance between poles, allowing us to find $R = L$.

Finally, for other values of $\Delta\omega$, the condition number still indicates $\lambda_1 \gg \lambda_2$. However, with very few samples it is possible to find the rank (R) of $\mathbf{H}_s$ equal to the number of
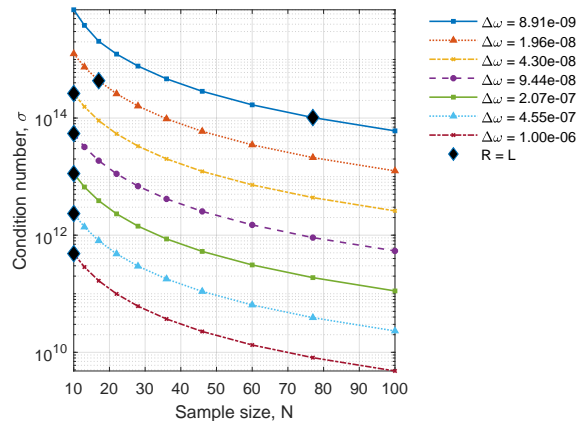

Fig. 2. Computing the condition number (σ) vs. the number of samples (N) to build the matrix using different values of $\Delta\omega$. For each curve, the diamond marker indicates the minimum value of N to have $rank(\mathbf{H}_s) = 2$.

poles L. It indicates that these are sufficient distances between poles.

B. Sample Size for a Full-Rank Hankel Matrix

This scenario aims to assess how the poles proximity ($\Delta\omega$) in a signal $s(n)$ with L poles, shown in eqn. (1), impacts its Hankel matrix rank R in the absence of noise.

The experiment consists in constructing $\mathbf{H}_s$ from 100 samples for each $s(n)$. To compare the numerical and theoretical results, we use a windowed version of $s(n)$ with $N \leq 100$ samples to build its equivalent Hankel structure, i.e., just the window $s(1:N)$ from the first N samples, is used to map the signal onto the matrix. Then, we search through each $\mathbf{H}_s$ mapped from the windowed signal for the first value of N yelding $R = L$.

We generate $s(n)$, where z_l represent equispaced poles, L is in the [2,8] interval, and distance $\Delta\omega$ takes 100 logarithmically spaced values in the $[10^{-4}, 10^{-1}]$ interval. The number of samples used in window N is a positive integer value between 2 and 100. The simulation of Fig. 3 may be summarized as follows. For each value of L in the set of positive integers, we vary $\Delta\omega$ from very close poles gradually moving them away from each other at each iteration, noting the value of N for $R = L$.

The experiment shows that the theoretical values based on eqn. (5) are different from those obtained empirically. One would expect plots with horizontal lines at $2L - 1$ regardless of $\Delta\omega$. Nevertheless, the simulation presents a different result: when the distance gets smaller, the number of samples N necessary in the windowed signal to build $\mathbf{H}_s$ with $R = L$ increases. This outcome reinforces the hypothesis that the increase of N may compensate for the distance between poles to obtain $R = L$ as anticipated by the analysis of Section II-B. This will increase the computational cost, since a reliable SVD algorithm (Golub-Reinsch) performs $8M^3/3$ flops to compute only the singular values [12], where M is the matrix dimension following eqn. (2).

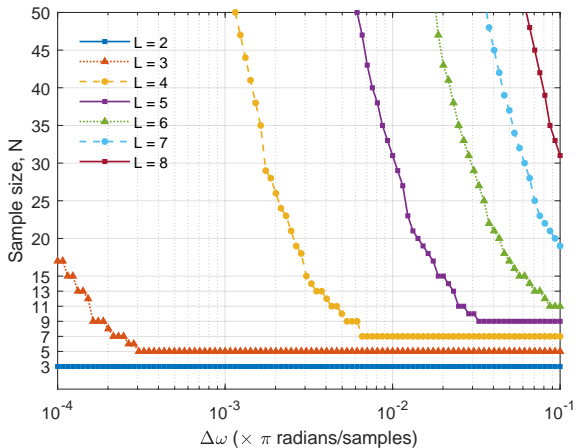Furthermore, one can see that as L increases, the curves move towards top right corner of the figure, which suggests

Fig. 3.   Number of samples N necessary to build a Hankel matrix $\mathbf{H}_s$, with $R = L$, versus the distance between poles ($\Delta\omega$). variation.
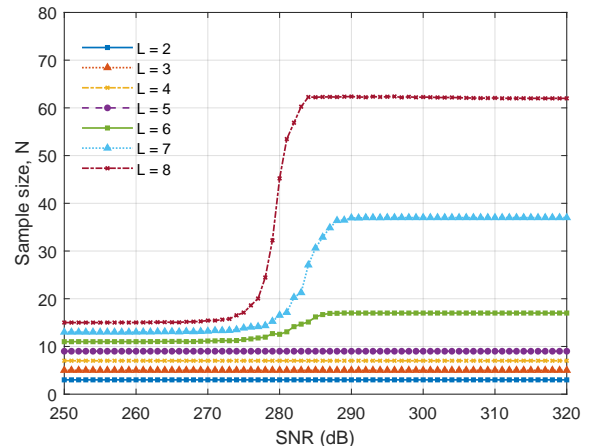

Fig. 4.   Number of samples N necessary in the windowed signal $y(n)$ to build $\mathbf{H}_s$, with $R = L$, versus the SNR values, and fixed distance between poles $\Delta\omega = 0.05$ for 100 Monte Carlo runs.

that rank depends on number of poles and $\Delta\omega$.

C. Noisy Scenario and SNR Threshold

To assess a more realistic case, one may consider a scenario with additive white Gaussian noise (AWGN), according to the following data model:

$$y(n) = s(n) + \beta(n). \qquad (10)$$

The simulation framework setup is similar to the presented in Section III-B, except for the mapped signal $y(n)$, a version of $s(n)$ as shown in eqn. (1), affected by the AWGN term $\beta(n)$, and with a fixed distance between poles $\Delta\omega = 0.05$.

Aiming to compare the theoretical values based in eqn. (5) with the data obtained empirically, we present an experiment that varies high SNR values, along the [250,320] interval in 10 dB steps. The experiment is performed to search a threshold SNR value, where the equivalent Hankel matrix $\mathbf{H}_s$ presents a rank R equal to the number of exponentials L, regardless of $\Delta\omega$. We repeat the procedure described in Section III-B (Fig. 3) for each SNR, and noting the results obtained for each single run. The scenario with 100 Monte Carlo runs is computed varying the noise for each realization.

Fig. 4 shows that for values of SNR less than 270 dB, in general the most commonly encountered in real problems, the theoretical result is respected. The plots illustrate the expected result, discussed in previous sections, with horizontal lines at $2L-1$, despite the $\Delta\omega$ variation with a rank R dependent only on its number of poles. For values greater than 270 dB, the theoretical values are once again different from those obtained empirically, but N starts to increase.

Finally, for this case of study, the three largest values of L present a transition close to 290 dB. We can see that for higher values of SNR, the trend observed in Fig.3 at $\Delta\omega = 0.05$ is exactly the same for all L. This outcome supports the results of Section III-B, obtained in the noiseless case.

IV. Discussion

The present work discusses the numerical features of the low-rank Hankel signal model using Vandermonde decompo-

sition. It explores various parameters as the distance between poles, and number of poles and samples in the construction of Hankel structures, mapped from all-pole signals.

Experiments based on the SVD measure the impact of parameters, taking into consideration two indices: the condition number, and the amount of samples mapped onto a Hankel matrix necessary to present a rank equal to the number poles in the equivalent signal. An unexpected relationship between the parameters appears in the noiseless case due to numerical effects, despite the noisy setup that presents the expected behavior with a limited SNR. Also, evidence was given with the help of simulations to support two remarks:

(I) As the distance between poles decreases, the number required to obtain a full rank matrix increases. Furthermore, in the second noiseless scenario, with a low-rank Hankel signal model, if the poles get too close to each other, a dependency between the rank, number of poles, amount of samples, and the distance between poles is observed.

(II) In the noisy scenario, with SNR values commonly encountered in practice, the distance between poles becomes irrelevant. The theoretical minimum value is observed, *i.e.*, the rank depends only on the number of poles.

V. Conclusion and Further Work

This work has studied the numerical features of Hankel matrices associated with complex exponential signal models. In the noiseless cases, to assure a full rank matrix and avoid the numerical problems presented, a minimum distance between poles is required or the number of mapped samples must increase to compensate for pole proximity and ensure full rank. These observations may impact the computational cost where the Vandermonde decomposition and Hankel structure are applied. In the noisy case, in real applications: with a realistic noise level, the Hankel matrix is always full rank, since the noise acts to balance the linear independence between the columns in the associated Vandermonde matrix. This can

reduce significantly the computational cost for problems with solid ground-truth.

Future work should focus on performing the condition number experiment for larger Vandermonde matrices, in order to compare more than two singular values to provide more relevant statistical information. Furthermore, the experiments should be in the context of tensor decompositions, such as BTD, since the model performs the separation of the noise in a different block of the low-rank Hankel signal model and the remaining blocks could fall within the noiseless scenario.

ACKNOWLEDGMENT

REFERENCES

[1] V. Zarzoso, "Extraction of ECG characteristics using source separation techniques: Exploiting statistical independence and beyond," in *Advanced Biosignal Processing*, A. Nait-Ali, Ed., Springer, 2009, ch. 2, pp. 15–47.

[2] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.

[3] P. M. R. de Oliveira and V. Zarzoso, "Block term decomposition of ecg recordings for atrial fibrillation analysis: Temporal and inter-patient variability," *Journal of Communication and Information Systems*, vol. 34, no. 1, pp. 111–119, 2019.

[4] L. De Lathauwer, "Blind separation of exponential polynomials and the decomposition of a tensor in rank-$(L_r, L_r, 1)$ terms," *SIAM Journal on Matrix Analysis and Applications*, vol. 32, no. 4, pp. 1451–1474, 2011.

[5] J. H. de M. Goulart, P. M. R. de Oliveira, R. C. Farias, V. Zarzoso, and P. Comon, "Alternating group lasso for block-term tensor decomposition and application to ECG source separation," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2682–2696, 2020.

[6] D. Boley, F. Luk, and D. Vandevoorde, "A general Vandermonde factorization of a Hankel matrix," *Int. Lin. Alg. Soc.(ILAS) Symp. on Fast Algorithms for Control, Signals and Image Processing*, 1997.

[7] L. Weruaga and A. Al-Khayat, "All-pole model estimation of vocal tract on the frequency domain," *Proceedings of the Annual Conference of the International Speech Communication Association*, vol. 2, 2006.

[8] B. Beckermann, G. Golub, and G. Labahn, "On the numerical condition of a generalized Hankel eigenvalue problem," *Numerische Mathematik*, vol. 106, pp. 41–68, 2007.

[9] L. Abdalah, P. M. R. de Oliveira, W. Freitas, and V. Zarzoso, "Tensor-based noninvasive atrial fibrillation complexity index for catheter ablation," in *Computing in Cardiology*, Rimni, Italy, Sep. 2020.

[10] Z. Drmac, "SVD of Hankel matrices in Vandermonde-Cauchy product form," *Electronic Transactions on Numerical Analysis*, vol. 44, pp. 593–623, 2015.

[11] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular value decomposition and principal component analysis," in *A Practical Approach to Microarray Data Analysis*, D. Berrar, W. Dubitzky, and M. Granzow, Eds. 2003, pp. 91–109.

[12] G. H. Golub and C. F. van Loan, *Matrix Computations*, Fourth ed. John Hopkins University Press, 2013.

[13] S. L. Brunton and J. N. Kutz, "Singular value decomposition (SVD)," in *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019, pp. 3–46.