ON THE FINITE RANK OF BOUNDED SEMI-INFINITE HANKEL OPERATORS AND THEIR FINITE DIMENSIONAL REPRESENTATION

DRAFT AS OF August 28, 2012

MOODY T. CHU* AND MATTHEW M. LIN†

Abstract. Bounded, semi-infinite Hankel matrices of finite rank over the space ℓ^2 of square-summable sequences occur frequently in classical analysis and engineering applications. The notion of finite rank often appears under different contexts and the literature is diverse. The first part of this paper reviews some elegant, classic criteria and establish connections among the various characterizations of finite rank in terms of rational function, recursion, matrix factorization, and sinusoidal signal. All criteria require $2d$ parameters, though in different meaning, for a matrix of rank d. The Vandermonde factorization, in particular, permits immediately a singular-value preserving, finite dimensional representation of the original semi-infinite Hankel matrix and, hence, makes it possible to retrieve the nonzero singular values of the semi-infinite Hankel matrix. The second part of this paper proposes using the LDL^* decomposition of a specially constructed sample matrix to find the unitarily equivalent finite dimensional representation. This approach enjoys several advantages, including the ease of computation by avoiding infinite dimensional vectors, the ability to reveal rank deficiency, and the established pivoting strategy for stability. No error analysis is given, but several computational issues are discussed.

Key words. Hankel operator, semi-infinite matrix, finite rank, Vandermonde factorization, orthogonalization, generating function, LDL^* decomposition

AMS subject classifications. 47B35, , 65F25, 65F30, 15A23, 15B05

1. Introduction.

The notion of Hankel operators has played an important role in relating many disparate parts of classical mathematics together [1, 25, 39, 40], among which we mention the notable moment problems [2]. In recent years, Hankel operators have also found important applications in wide range of disciplines outside mathematics. See the monographs [25, 38, 40] for some general discussions. One such instance is the H^∞ techniques in control theory and systems theory where Hankel operators are employed to synthesize controllers that minimize sensitivity or achieve stabilization [14, 18].

Hankel operators may appear in different forms. In many applications, it is common to cast the Hankel operators in matrix notation, finite or infinite,

$$H = \begin{bmatrix} h_0 & h_1 & h_2 & \dots \\ h_1 & h_2 & h_3 & \\ h_2 & h_3 & & \\ h_3 & & & \\ \vdots & & & \end{bmatrix}, \tag{1.1}$$

with respect to some preselected bases. That is, a Hankel matrix is characterized by the property that its entries depend only on the sum of indices. We often identify H with the sequence $\{h_n\}_{n\geq 0}$ of its first column (and last row). When the underlying matrix is of finite dimension and of the Hankel structure, many numerical algorithms can be implemented more efficiently. For instance, it is well known that solving a linear system $H\mathbf{x} = \mathbf{b}$ where H is an $n \times n$ Hankel (or Toeplitz) matrix requires only $O(n^2)$ arithmetic operations [24, 36]. Similarly, the singular value decomposition can be accomplished with $O(n^2 \log n)$ operations [23, 31]. For semi-infinite Hankel matrices, however, numerical computation remains challenging. One example is that, despite our theoretical understanding on spectral properties of semi-infinite Hankel matrices [44], it requires serious efforts to retrieve information analogous to the notion of finite dimensional singular value decomposition [28, 46].

*Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu) This research was supported in part by the National Science Foundation under grant DMS-1014666.

†Department of Mathematics, National Chung Cheng University, Chiaya 62012 Taiwan. (mlin@math.ccu.edu.tw). This research was supported in part by the National Science Council of Taiwan under grant 99-2115-M-194-010-MY2.

Indeed, a concern already arises when dealing with matrix-to-vector multiplication that involves a semi-infinite matrix H. It inevitably involves infinite series. For the series to converge, we must impose the condition that H be a bounded linear operator. In most literature, H is considered as an operator over the space ℓ^2 of square summable (semi-infinite) sequences. It is known that a Hankel matrix H represents a bounded operator over ℓ^2 if and only if there exists a function $\psi \in L^\infty$ on the unit circle such that [39, Theorem 2.1]

$$h_n = \frac{1}{2\pi} \int_0^{2\pi} \psi(\theta) e^{-\imath n\theta} d\theta, \quad n = 0, 1, \ldots. \tag{1.2}$$

In other words, whether the sequence$\{h_n\}_{n\geq 0}$ for H determines a bounded operator on ℓ^2 is equivalent to whether the sequence itself represents the Fourier coefficients of an essentially bounded function over the unit disk [35]. Since L^∞ is dense in L^2 over the compact set $[0, 2\pi]$, we may assume without loss of generality that all rows and columns of H are sequences in ℓ^2 themselves [8, 41].

In this paper, we are particularly interested in the case when an semi-infinite H is known a priori to have finite rank. This class of matrices arises in many fields and the literature on this subject is diverse, often under different names. We mention only a number of its applications, such as signal processing [21], system identification [14, 18], model reduction [29, 33], speech and audio processing [30], modal and spectral analysis [32], and image processing [22]. This paper focuses only on the characterization of such a semi-infinite Hankel matrix H in terms of its nonzero singular values and singular vectors. Our main contribution is to provide a theoretical groundwork for the required computational endeavor.

We intend to bring forth two discussions. In Section 2 we review various characterizations of finite rank scattered in the literature with the hope of establishing connections in terms of rational function, recursion, matrix factorization, and sinusoidal signal. In Section 3 we propose to form a finite dimensional matrix that is unitarily equivalent to the original semi-infinite Hankel matrix via the rank-revealing LDL^* algorithm. Numerical challenges are pointed out along the course of discussion.

2. Finite rank. As the phrase "finite rank" plays a key role in many applications, in the first of this paper we explain the cause and effect of such a notion. The following brief review highlights some, but not all, classical results in the literature which is scattered. We assemble them as a collage and show their connections.

2.1. Rational function. We begin with perhaps the earliest theoretical consideration. By a generating function associated with a sequence $\{a_n\}_{n\geq 0}$ of complex numbers, we mean the formal power series defined by

$$G(z; \{a_n\}) := \sum_{n=0}^{\infty} a_n z^n. \tag{2.1}$$

It was Kronecker who first proposed an intriguing way to check the finiteness of the rank of a Hankel matrix by examining the convergence of the associated generating function. More precisely, we state the following theorem which, in fact, also asserts explicitly the corresponding rank [39, Theorem 4.1].

THEOREM 2.1. *The Hankel matrix H has finite rank if and only if the power series (2.1) determines a rational function. In this case,*

$$\operatorname{rank}(H) = \deg(zG(z; \{h_n\})), \tag{2.2}$$

where the degree of a rational function is the maximum of the degrees of its minimal constituent polynomials.

Any calculus student should have learned that determining the convergence of an infinite series usually is easier than finding the limit of its partial sums. To find the limit $h(z)$ analytically of the generating function so as to calculate the rank of a given Hankel matrix as indicated above, therefore, is already posed as a challenging task.

2.2. Recursive formula. An equivalent but more direct observation of finite rank is through the following recursive relationship among elements of H [15, Chapter XV, Theorem 7].

THEOREM 2.2. *The Hankel matrix H is of finite rank d if and only if there exist constants $\gamma_0, \ldots, \gamma_{d-1}$ such that*

$$h_i = \gamma_{d-1} h_{i-1} + \gamma_{d-2} h_{i-2} + \ldots + \gamma_o h_{i-d}, \quad i = d, d+1, \ldots \tag{2.3}$$

and d is the least integer having this property.

Assuming that the semi-infinite Hankel matrix H in Theorem 2.2 is of rank d, then it can be argued that the $d \times d$ leading principal submatrix $\widehat{H}$, called the *trajectory matrix* of the segment $h_0, \ldots, h_{2d-2}$, is necessarily nonsingular. Rewrite the finite difference equation (2.3) in the form of a linear system

$$\begin{bmatrix} h_0 & h_1 & h_2 & \ldots & h_{d-1} \\ h_1 & h_2 & h_3 & & h_d \\ h_2 & h_3 & & & h_{d+1} \\ \vdots & & & & \\ h_{d-1} & h_d & & & h_{2d-2} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{d-1} \end{bmatrix} = \begin{bmatrix} h_d \\ h_{d+1} \\ \vdots \\ h_{2d-1} \end{bmatrix}, \tag{2.4}$$

known as the Yule-Walker equations in signal processing, for the coefficients $\gamma_0, \gamma_1, \ldots, \gamma_{d-1}$. It follows that the parameters $\gamma_0, \ldots, \gamma_{d-1}$ and, consequently, the entire semi-infinite time series, are uniquely determined by the first $2d$ elements $h_0, h_1, \ldots, h_{2d-1}$. In this sense, we may say that a semi-infinite Hankel matrix of finite rank d has a $2d$-dimensional representation $\{h_0, h_1, \ldots, h_{2d-1}\}$. The Yule-Walker equations are ubiquitous in science, finance, and technology. It is said that "every cell phone solves the Yule-Walker equations every 10 microseconds" [11] and, needless to say, many discussions and effective algorithms are available. We mention an interesting web site [42] where further references can be found.

In theory, from a given Hankel matrix H, we could repeatedly try out a sequence of nested linear systems in the form of (2.4) by gradually *decreasing* the value of d from a large initial guess until the trajectory matrix becomes nonsingular for the first time, whence the linear system is solved and the rank is determined. Nonetheless, issues of numerical rank do arise in reality because, even though the Courant-Fischer theorem ensures interlacing of singular values between two nested trajectory matrices, it does not guarantee how far a nonsingular matrix is away from the nearest singular matrix. It is possible, even without the presence of noise, that we call a matrix singular when in fact it is only nearly singular. Such an adjudication often is subjective. Rank revealing algorithms developed thus far in the literature cannot resolve this problem perfectly.

2.3. Sinusoidal signal. In the field of signal processing, there is another popular way to parameterize Hankel matrices of finite rank by means of sinusoidal signals. The forward connection of associating a damped signal of finite component to a Hankel matrix of finite rank is quite straightforward. Consider a noiseless time-domain signal comprising d components of exponentially decaying sinusoids,

$$\mathfrak{s}(t) = \sum_{\ell=1}^{d} a_\ell e^{-\alpha_\ell t} e^{\imath(2\pi\nu_\ell t + \phi_\ell)}, \tag{2.5}$$

where a_ℓ, α_ℓ, ν_ℓ, and ϕ_ℓ are real numbers denoting the magnitude, the decay rate, the frequency, and the phase angle of the ℓ-th sinusoid, respectively. Starting with $t_0 = 0$ and sampling this signal at uniformly spaced nodes $t_0, t_1, \ldots$ with fixed interval length Δt (so $1/\Delta t$ is the so called sampling rate), we obtain an infinite sequence

$$s_k := \mathfrak{s}(t_k) = \sum_{\ell=1}^{d} a_\ell e^{-\alpha_\ell k \Delta t} e^{\imath(2\pi\nu_\ell k \Delta t + \phi_\ell)} = \sum_{\ell=1}^{d} a_\ell e^{\imath\phi_\ell} \left(e^{(-\alpha_\ell + \imath 2\pi\nu_\ell)\Delta t}\right)^k.$$

For simplicity, denote

$$\beta_\ell := a_\ell e^{\imath\phi_\ell}, \tag{2.6}$$

$$\lambda_\ell := e^{(-\alpha_\ell + \imath 2\pi\nu_\ell)\Delta t}. \tag{2.7}$$

Obviously, the time series $\{s_0, s_1, \ldots\}$ enjoys the relationship

$$s_k = \sum_{\ell=1}^{d} \beta_\ell \lambda_\ell^k. \tag{2.8}$$

Such a relationship implies a decomposition of the corresponding trajectory matrix

$$S = \sum_{\ell=1}^{d} \beta_\ell \begin{bmatrix} \lambda_\ell^0 & \lambda_\ell^1 & \lambda_\ell^2 & \cdots \\ \lambda_\ell^1 & \lambda_\ell^2 & & \\ \lambda_\ell^2 & & & \\ \vdots & & & \end{bmatrix}$$

into the form

$$S = \begin{bmatrix} \lambda_1^0 & \lambda_2^0 & \cdots & \lambda_d^0 \\ \lambda_1^1 & \lambda_2^1 & & \lambda_d^1 \\ \lambda_1^2 & & & \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & & \\ \vdots & & \ddots & \\ 0 & & & \beta_d \end{bmatrix} \begin{bmatrix} \lambda_1^0 & \lambda_1^1 & \lambda_1^2 & \cdots \\ \lambda_2^0 & \lambda_2^1 & & \\ \vdots & & & \\ \lambda_d^0 & \lambda_d^1 & & \cdots \end{bmatrix}, \tag{2.9}$$

which is known as the Vandermonde factorization of S. Under the generic assumption that all $\lambda_1, \ldots, \lambda_d$ are distinct, it is seen immediately from (2.9) that the Hankel matrix S is necessarily of rank d. Note also that for strictly decaying signals, we should have $\alpha_\ell > 0$ and, hence, $|\lambda_\ell| < 1$, implying that S is a bounded operator.

It turns out that backward connection of associating a Hankel matrix of finite rank to a damped signal of finite component is also true. The reason is based on the fact that any given Hankel matrix of finite rank d enjoys a Vandermonde factorization of which (2.9) is a special case. We shall establish the details in Section 2.4. When this factorization is in hand, then by (2.6) the polar form of the complex number β_ℓ determines the magnitude a_ℓ and phase angle ϕ_ℓ. In the meantime, by (2.7) the polar form the quantity λ_ℓ determines, up to a scaling by Δt, the decay rate α_ℓ and the frequency ν_ℓ. That there is a dependence on Δt does make sense because the measurement of decay rate and frequency should be relative to the meaning of a unit time that must be defined somewhere. Thus, once Δt is specified, a composite signal is completely determined from a given Vandermonde decomposition. In short, we can interpret interchangeably between a bounded low rank Hankel matrix and a sinusoidal signal through the relationship (2.8). Ultimately, the term "singular values" of a signal $\mathfrak{s}$ can be understood from those of the corresponding trajectory matrix S.

The role of the sampling rate is significant and is worth noting. Out of the very same composite signal $\mathfrak{s}(t)$, a large Δt, i.e., a low sampling rate, will make λ_ℓ's cluster around the origin, resulting the sequence $\{s_k\}$ to quickly decay to zero. The corresponding Hankel matrix S would then carry lots of anti-diagonals whose entries are nearly machine zero. In theory, as we have described above, the Hankel matrix S should still be of rank d. When we calculate its numerical rank, however, the number of calculated singular values that are greater than a specified threshold could be far less than d. Such a mismatch then causes confusion on how the original signal should be represented.

More difficulties arise when noises are present in a signal. Even with a slight perturbation, the above connection usually is lost, that is, the trajectory matrix of the contaminated time series usually is of full rank and the true value of d is smudged. Reconstructing a Hankel low rank approximation to the

contaminated trajectory matrix is a problem of practical importance and has been extensively studied in the literature. Far from being complete, we mention only a few references [9, 17, 30, 32, 33, 37]. See also [3, 5, 10, 16, 34, 43, 45] for the populous singular-value spectral analysis (SSA) techniques. In all, it is imperative to first determine an appropriate rank before any filtering technique can be applied to rid of the noises. Such a task of rank detection remains to be a challenging question to this date [19].

2.4. Vandermonde factorization. In this section we fill in the details that associate a semi-infinite Hankel matrix of finite rank to a sinusoidal signal. We approach this backward connection by exploiting the finite difference equation (2.3) and the corresponding the characteristic polynomial

$$p(\lambda) := \lambda^d - \gamma_{d-1}\lambda^{d-1} - \ldots - \gamma_1\lambda - \gamma_0. \tag{2.10}$$

To fix the idea, let λ_ℓ, $\ell = 1, \ldots, r$, denote the distinct roots of $p(\lambda)$, each of which has multiplicity ρ_ℓ. So $\sum_{\ell=1}^{r} \rho_\ell = d$. A general solution to the difference equation (2.3) can be formulated as follows. Let $[\beta_1^{(0)}, \ldots, \beta_1^{(\rho_1-1)}, \beta_2^{(0)}, \ldots, \beta_2^{(\rho_2-1)}, \ldots \beta_r^{(0)}, \ldots \beta_r^{(\rho_r-1)}]^\top \in \mathbb{C}^d$ be the solution to the confluent Vandermonde system [26]

$$\mathscr{V}_{p(\lambda)} \begin{bmatrix} \beta_1^{(0)} \\ \vdots \\ \beta_1^{(\rho_1-1)} \\ \vdots \\ \beta_r^{(0)} \\ \vdots \\ \beta_r^{(\rho_r-1)} \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{d-1} \end{bmatrix}, \tag{2.11}$$

where $\mathscr{V}_{p(\lambda)} \in \mathbb{C}^{d\times d}$ is partitioned column-wise into r blocks

$$\mathscr{V}_{p(\lambda)} = [\mathscr{V}^{(1)}, \ldots, \mathscr{V}^{(r)}],$$

with the block $\mathscr{V}^{(\ell)}$, $\ell = 1, \ldots, r$, defined by

$$\mathscr{V}^{(\ell)} := \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ \lambda_\ell & 1 & 0 & \ldots & \\ \lambda_\ell^2 & 2\lambda_\ell & 1 & \ldots & \\ \lambda_\ell^3 & 3\lambda_\ell^2 & 3\lambda_\ell & & \\ \vdots & & & \ddots & \\ \lambda_\ell^{\rho_\ell-1} & & & & 1 \\ \lambda_\ell^{\rho_\ell} & & & & \rho_\ell\lambda_\ell \\ \vdots & \vdots & & & \vdots \\ \lambda_\ell^{d-1} & (d-1)\lambda_\ell^{d-2} & \frac{(d-1)(d-2)}{2}\lambda_\ell^{d-3} & \ldots & \begin{pmatrix} d-1 \\ \rho_\ell - 1 \end{pmatrix} \lambda_\ell^{d-\rho_\ell} \end{bmatrix} \in \mathbb{C}^{d\times\rho_\ell}.$$

For notational simplicity, mark entries of the $d \times \rho_\ell$ matrix $\mathscr{V}^{(\ell)} = [v_{ij}^{(\ell)}]$ by

$$v_{ij}^{(\ell)} = c_{ij}\lambda_\ell^{i-j}, \quad i = 0, 1, \ldots, d-1, \quad j = 0, 1, \ldots, \rho_\ell - 1, \tag{2.12}$$

where $c_{ij} = 0$ if $i < j$; and $c_{ij} = \frac{i!}{(i-j)!j!}$ otherwise. Note also that for convenience, the indices of the entries start from $(0, 0)$. The purpose of imposing the system (2.11) is to enforce the initial values

$h_0, \ldots, h_{d-1}$ on the solution to the finite difference equation (2.3). Trivially, the equations in (2.11) can be expressed as

$$h_i = \sum_{\ell=1}^{r} \sum_{j=0}^{\rho_\ell - 1} \beta_\ell^{(j)} c_{ij} \lambda_\ell^{i-j}, \quad i = 0, 1, \ldots, d-1,$$

from which we can verify recursively that

$$\begin{aligned} h_i &= \sum_{s=0}^{d-1} \gamma_s h_{i-d+s} \\ &= \sum_{s=0}^{d-1} \gamma_s \left(\sum_{\ell=1}^{r} \sum_{j=0}^{\rho_\ell - 1} \beta_\ell^{(j)} c_{i-d+s,j} \lambda_\ell^{i-d+s-j} \right) = \sum_{\ell=1}^{r} \sum_{j=0}^{\rho_\ell - 1} \beta_\ell^{(j)} \left(\sum_{s=0}^{d-1} \left(c_{i-d+s,j} \gamma_s \lambda_\ell^{i-d+s-j} \right) \right) \\ &= \sum_{\ell=1}^{r} \sum_{j=0}^{\rho_\ell - 1} \beta_\ell^{(j)} c_{ij} \lambda_\ell^{i-j}, \quad i = d, d+1, \ldots, \end{aligned} \tag{2.13}$$

where the last equality is obtained by using the identity

$$\frac{d^j (\lambda_\ell^{i-d} p(\lambda_\ell))}{d\lambda^j} = 0, \quad \ell = 1, \ldots, r, \quad j = 0, 1, \ldots, \rho_\ell - 1 \tag{2.14}$$

for all $i \geq d$. Let V_∞ denote the confluent Vandermonde matrix of size $\infty \times d$ obtained by extending the matrix $\mathscr{V}_{p(\lambda)}$ defined in (2.12) downward to infinite length. Then the following result, known as the Vandermonde factorization for a semi-infinite Hankel matrix H, is but an rearrangement of the expression (2.13) in a matrix form [7].

THEOREM 2.3. *Suppose H is a semi-infinite Hankel matrix of rank d. Then there exist a $d \times d$ block diagonal matrix D_∞ whose ℓ-th block is of size $\rho_\ell \times \rho_\ell$ and is Hankel and upper anti-triangular such that*

$$H = V_\infty D_\infty V_\infty^\top. \tag{2.15}$$

It is important to point out that the Vandermonde factorization described above is valid for any semi-infinite Hankel matrix of rank d. If H is further required to be bounded (as an operator), then the root conditions must be satisfied, that is, all roots λ_ℓ's of the polynomial (2.10) must have modulus less than or equal to one and those of modulus one must be simple.

Consider the generic case, as we shall assume henceforth, that all roots of $p(\lambda)$ are distinct. The general solution (2.13) can be easily reduced to

$$\begin{aligned} h_i &= \sum_{s=0}^{d-1} \gamma_s h_{i-d+s} = \sum_{s=1}^{d-1} \gamma_s \left(\sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i-d+s} \right) = \sum_{\ell=1}^{d} \beta_\ell^{(0)} \left(\sum_{s=0}^{d-1} \gamma_s \lambda_\ell^{i-d+s} \right) \\ &= \sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i-d} \left(\sum_{s=0}^{d-1} \gamma_s \lambda_\ell^{s} \right) = \sum_{\ell=1}^{d} \beta_\ell^{(0)} \lambda_\ell^{i}, \quad i = d, d+1, \ldots, \end{aligned} \tag{2.16}$$

which now formally agrees with (2.8) and, thus, we can associate the Hankel matrix to a corresponding sinusoidal signal.

The idea of truncation and the extension is worth mentioning. Given a semi-infinite Hankel matrix H with rank d, any of its $m \times k$ submatrix with $\min\{m, k\} \geq d$ is of rank at most d. However, not

all low rank finite Hankel matrices can be embedded as a submatrix of a semi-infinite Hankel matrix with equal rank. It seems intuitive that on could trivially expand a given $m \times k$ Hankel matrix into a semi-infinite Hankel matrix by padding with innocuous zeros along the anti-diagonals, but such an extension usually results in a semi-infinite matrix with higher rank. Such a study actually is in the area of singular extension theory [25, II.9] and has applications such as Gaussian quadratures [20]. In the course of our discussion, we have already suggested an alternative way of extension by embedding a given $d \times d$ nonsingular Hankel matrix as the leading principal submatrix of a nontrivial semi-infinite Hankel matrix H with a specified value for one additional element h_{2d-1}. After solving the corresponding equation (2.4), we will have created the recursive relationship (2.13) which results in a matrix H of rank d. Through this way, any finite dimensional nonsingular Hankel matrix also admits a Vandermonde factorization [7].

Finally, we mention in passing that our discussion thus far has been about "scalar" Hankel matrices. Similar representation can be extended to nonsingular "block" Hankel matrices [12, 29]. For the purpose of conveying our basic idea, we shall concentrate on the scalar Hankel matrices only in this paper. Its generalization to block Hankel matrices deserves further investigation.

3. Finite dimensional representation. Given a bounded Hankel operator H of finite rank d, we see immediately from the Vandermonde factorization (2.15) that

$$\mathfrak{R}(H) = \mathfrak{R}(V_\infty), \tag{3.1}$$

$$\mathfrak{R}(H^*) = \mathfrak{R}(\overline{V}_\infty), \tag{3.2}$$

where $\overline{V}_\infty$ denotes the complex conjugate of V_∞ and $\mathfrak{R}(H)$ stands for the range space of H. By the Fredholm alternative theorem, we see that H and H^* have effects only on finite dimensional spaces. Thus it is natural to expect that H should have a finite dimensional representation. We stress that we are interested in rendering the "action" H as a map, which is more than just a passive characterization of H by finitely many parameters, such as the first $2d$ elements $\{h_0, h_1, \ldots, h_{2d-1}\}$ or the corresponding Vandermonde parameters $\{\beta_1, \ldots, \beta_d, \lambda_1, \ldots, \lambda_d\}$. In particular, we are interested in finding a singular value decomposition of H and representing H via a finite dimensional matrix of size $d \times d$.

For singular value computation, an interesting approach employing the generating function (2.1) of the infinite sequence $\{h_0, h_1, \ldots\}$ in ℓ^2 has already been proposed in [46]. That development was based on the fact asserted by Theorem 2.1 that the associated generating function of H is necessarily a rational function, whence the range space $R(H)$ of H as well as that of its adjoint H^* can be fully characterized in terms of some specially selected finite bases of rational functions. Our approach in this note is similar in spirit, but different from [46] in that we utilize the Vandermonde factorization (2.15) of H to build the representation.

The connections described in the preceding section can be implemented as a numerical procedure to achieve the Vandermonde factorization of a given Hankel matrix H. We caution, however, that many computational issues remain to be thoroughly investigated, including the stability of the algorithm itself that leads to the Vandermonde factorization and the conditioning of the roots of the characterization polynomial (2.10). These are important and rightful questions that, to our knowledge, have only been partially studied in the literature [6, 27]. In order to stay focus on the singular value decomposition, we shall not discuss these issues in this note either, but assume that the factorization is already in hand. We hope that the following discussion serves at least as a theoretic framework of interest.

To characterize the action of H, it suffices to find a matrix representation of the restricted map

$$H|_{\mathfrak{R}(\overline{V}_\infty)} : \mathfrak{R}(\overline{V}_\infty) \to \mathfrak{R}(V_\infty)$$

with respect to some orthonormal bases for the two range spaces. The reason is that, by the Fredholm alternative theorem, the action of H on the orthogonal complement of $\mathfrak{R}(\overline{V}_\infty)$ is zero. For convenience,

denote this restricted map by $\widetilde{H} := H|_{\mathfrak{R}(\overline{V}_\infty)}$ and express the relationship through the commutative diagram in Figure 3.1. It is important to point out that while the images of the restricted map $\widetilde{H}$ are identical to those of H as an operator on nontrivial elements, it no longer holds the Hankel structure in general.

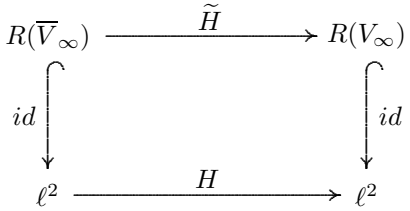

FIG. 3.1. *Commutative diagram between semi-infinite Hankel operator H of finite rank and its finite dimensional representation $\widetilde{H}$.*

We now describe how the finite dimensional matrix $\widetilde{H}$ can be found.

LEMMA 3.1. *Suppose that the matrix V_∞ has a QR decomposition*

$$V_\infty = QR, \tag{3.3}$$

where columns of $Q \in \mathbb{C}^{\infty\times d}$ are mutually orthonormal and $R \in \mathbb{C}^{d\times d}$ is upper triangular. Then the nonzero singular values of H are the same as those of the $d \times d$ matrix

$$\widetilde{H} := Q^* H \overline{Q} = R D_\infty R^\top. \tag{3.4}$$

Proof. The two subspaces $\mathfrak{R}(V_\infty)$ and $\mathfrak{R}(\overline{V}_\infty)$ have columns of Q and $\overline{Q}$ as orthonormal bases, respectively. The first equality follows from standard formula for change of basis. The second equality follows from the substitution $H = QRD_\infty R^\top Q^\top$ and the fact $Q^*Q = Q^\top \overline{Q} = I$. ∎

Both equalities in (3.4) have critical meanings. The first equality ensures that $\widetilde{H}$ and H are unitarily equivalent and, hence, the singular values of both systems are preserved. That is, the nonzero singular values of the semi-infinite matrix H can now be calculated from the finite matrix $\widetilde{H}$. To generate $\widetilde{H}$ by using Q, however, is cumbersome because the multiplication involves infinite series from columns of Q and H. For computation, the second equality becomes handy because both matrices R and D_∞ are of size $d \times d$ only.

We now explain how an orthonormal basis can be found for V_∞. Our ultimate goal is to compute the finite dimensional matrix R without explicitly forming Q so that together with D_∞ we can compute the nonzero singular values of H via $\widetilde{H}$. We divide our approach into three steps.

3.1. Modified Gram-Schmidt orthogonalization. The challenge in computing the QR decomposition for V_∞ is that we have to deal with infinite dimensional vectors. We first propose using the Gram-Schmidt process to carry out the task. One attractive feature of this approach is that, with the Vandermonde structure, all inner products of infinite dimensional vectors can be handled effectively via rational arithmetic if we employ the notion of generating functions over ℓ^2.

We point out two useful facts before proceed. Firstly, observe that the generating function f_ℓ corresponding to the ℓth column of V_∞, $\ell = 1, \ldots, d$, is given by

$$f_\ell(z) := G(z; \{\lambda_\ell^n\}) = \sum_{n=0}^{\infty} \lambda_\ell^n z^n = \frac{1}{1 - \lambda_\ell z}. \tag{3.5}$$

Secondly, a natural way to define inner product between two generating functions is by

$$\langle G(z;\{a_n\}), G(z;\{b_n\}\rangle := \sum_{n=0}^{\infty} a_n \overline{b}_n, \tag{3.6}$$

where $\overline{b}_n$ denotes the complex conjugate of b_n. The relationship

$$\left\langle \frac{1}{1-sz}, G(z;\{b_n\}) \right\rangle = \overline{G(\overline{s},\{b_n\})} \tag{3.7}$$

then holds whenever the evaluation $G(\overline{s},\{b_n\})$ makes sense.

Applying the Gram-Schmidt orthogonalization process to columns of V_∞ is equivalent to applying the process to the corresponding generating functions f_ℓ's. Summarized in Algorithm 1 is the modified Gram-Schmidt scheme which is the same as the Gram-Schmidt process in exact arithmetic.

Algorithm 1 Modified Gram-Schmidt Process on the Generating Functions.

Require: complex numbers $\lambda_1, \ldots, \lambda_d$ {from the Vandermonde decomposition of a Hankel matrix H of rank d}

Ensure: QR decomposition of V_∞ in the form (3.9)

```
for ℓ = 1, ··· , d do
    w_ℓ := f_ℓ
end for
for ℓ = 1, ··· , d do
    r_ℓℓ := ‖w_ℓ‖
    w_ℓ := w_ℓ / r_ℓℓ
    for k = ℓ + 1, ··· , d do
        r_ℓk := ⟨w_k, w_ℓ⟩
        w_k := w_k − r_ℓk w_ℓ
    end for
end for
```

Because w_ℓ's are functions of the indeterminant z, the calculation of $r_{\ell k}$'s in Algorithm 1 is but function evaluations which can be carried out by rational arithmetic. We typify a few steps below. It is obvious by (3.7) that

$$r_{11} = \sqrt{f_1(\overline{\lambda}_1)} = \frac{1}{\sqrt{1-\lambda_1\overline{\lambda}_1}}$$

which is guaranteed to be a positive number. After obtaining the normalized $w_1(z)$, we see by (3.7) again that the first row of R is attainable via

$$r_{1k} = \overline{w_1(\overline{\lambda}_k)}, \quad k = 2, \ldots, d.$$

According to the algorithm, we temporarily should have

$$w_k(z) = f_k(z) - r_{1k} w_1(z), \quad k = 2, \ldots d,$$

each of which is orthogonal to $w_1(z)$ and, hence,

$$r_{22} = \|w_2\| = \sqrt{\langle f_2 - r_{12} w_1, w_2\rangle} = \sqrt{\overline{w_2(\overline{\lambda}_2)}}.$$

The conjugation in the above actually is not needed because the quantity $w_2(\overline{\lambda}_2)$ is guaranteed, in exact arithmetic, real and nonnegative. We keep the conjugation because it is consistent with the general expression that, after normalizing $w_2(z)$, we can generate the second row of R via

$$r_{2k} = \langle w_k, w_2 \rangle = \langle f_k - r_{1k} w_1, w_2 \rangle = \langle f_k, w_2 \rangle = \overline{w_2(\overline{\lambda}_k)}, \quad k = 2, \ldots, d.$$

By recursion, we are ready to generate the ℓ-th row of R via the expression

$$r_{\ell k} = \overline{w_\ell(\overline{\lambda}_k)}, \quad k = \ell, \ldots, d, \tag{3.8}$$

where $w_\ell(z)$ is the normalized generating function as is specified in the algorithm. In this way, the QR factorization for the semi-infinite matrix V_∞ can be represented in terms of a decomposition of functions $f_1, \ldots, f_d$ in the form

$$[f_1, f_2, \ldots, f_d] = [w_1, w_2, \ldots, w_d] \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & \cdots & r_{1d} \\ 0 & r_{22} & r_{23} & r_{24} & & r_{2d} \\ 0 & 0 & r_{33} & r_{34} & & \\ \vdots & & & & & \\ & & & & \ddots & \\ 0 & & & & \cdots & r_{dd} \end{bmatrix}, \tag{3.9}$$

whereas $w_1, \ldots, w_d$ themselves are mutually orthonormal functions (and, hence, infinite sequences), with respect to the inner product (3.6) and the $d \times d$ upper triangular matrix $R = [r_{\ell k}]$ is really the one needed in the finite dimensional representation (3.4).

It is interesting to make the following observation about the roots of the function $w_\ell(z)$.

LEMMA 3.2. *Let w_ℓ denote the generating function corresponding to the ℓ-th column of the orthogonal matrix in the QR decomposition of V_∞. Then the function $w_\ell(z)$ has at least $\ell - 1$ roots at $\overline{\lambda}_1, \ldots, \overline{\lambda}_{\ell-1}$.*

Proof. In the same way as (3.8), for each $j = 1, \ldots, \ell - 1$ the inner product $\langle w_j, w_\ell \rangle = \overline{w_\ell(\overline{\lambda}_j)}$ is just the conjugation of the evaluation. By construction, $w_\ell(z)$ is orthogonal to all w_j. ∎

The implication of Lemma 3.2 is that if $\overline{\lambda}_\ell$ is near to any of these roots, then the value of $r_{\ell\ell}$ will be small, causing numerical instability even for the modified Gram-Schmidt process. We shall offer an alternative way to circumvent this difficulty.

3.2. Implicit calculation. In Algorithm 1, we compute each function $w_\ell(z)$ at its entirety. Such a function is needed only if the corresponding column in Q is needed. Symbolic calculation of the rational function $w_\ell(z)$ is theoretically possible but lengthy and tedious.

On the other hand, a close examination of Algorithm 1 indicates that all we need is the evaluation of $w_\ell(z)$ at discrete points $\overline{\lambda}_\ell, \ldots, \overline{\lambda}_d$ for each $\ell = 1, \ldots, d$. These are the entries for the upper triangular matrix R. Clearly, we have

$$w_1(\overline{\lambda}_k) = \frac{f_1(\overline{\lambda}_k)}{\sqrt{f_1(\overline{\lambda}_1)}}, \quad k = 1, \ldots, d, \tag{3.10}$$

to begin with. The remaining recursion can be written as

$$w_\ell(\overline{\lambda}_k) = \frac{f_\ell(\overline{\lambda}_k) - \sum_{j=1}^{\ell-1} r_{j\ell} w_j(\overline{\lambda}_k)}{\sqrt{f_\ell(\overline{\lambda}_\ell) - \sum_{j=1}^{\ell-1} r_{j\ell} w_j(\overline{\lambda}_\ell)}}, \quad k = \ell, \ldots, d. \tag{3.11}$$

Reorganized in Algorithm 2 is the very same Gram-Schmidt process, except that the generating functions $w_\ell(z)$'s and, thus, the corresponding orthogonal matrix Q are now implicit. Furthermore, we

write the code in the MATLAB syntax for the convenience of vector processing. The algorithm generates internally one Hermitian matrix F, referred to as the sample matrix, and one lower triangular matrix Ω. The (k,ℓ) entries of F and Ω are $f_\ell(\overline{\lambda_k})$ and $w_\ell(\overline{\lambda}_k)$, denoting the evaluation of the generating functions $f_\ell(z)$ and $w_\ell(z)$ at $\overline{\lambda}_1,\ldots,\overline{\lambda}_d$ (and recall Lemma 3.2), respectively. The upper triangular matrix $R := \Omega^*$ returned by the algorithm is precisely the matrix needed in (3.9). By construction, the three matrices satisfy the relationship

$$F = \Omega R. \tag{3.12}$$

Algorithm 2 Reorganized Modified Gram-Schmidt Process (Implicit Q).

Require: complex numbers $\lambda_1,\ldots,\lambda_d$ {from the Vandermonde decomposition of a Hankel matrix H of rank d}

Ensure: Hermitian matrix F, lower triangular matrix Ω, and upper triangular matrix R, all finite dimensional and satisfying (3.12)

$\Lambda := [\lambda_1,\ldots,\lambda_d]$
$F := 1./(1-\Lambda' * \Lambda)$
for $\ell = 1,\cdots,d$ **do**
  $r(\ell,\ell) := \sqrt{F(\ell,\ell) - \omega(\ell,1:\ell-1) * r(1:\ell-1,\ell)}$
  $\omega(\ell:d,\ell) := (F(\ell:d,\ell) - \omega(\ell:d,1:\ell-1) * r(1:\ell-1,\ell))/r(\ell,\ell)$
  $r(\ell,\ell:d) := (\omega(\ell:d,\ell))'$
end for

The expressions (3.9) and (3.12) look similar, but there is an important difference. In (3.9), columns of $[f_1,\ldots f_d]$ and $[w_1,\ldots,w_d]$ are supposedly infinite sequences or, in our representation, generating functions. In (3.12), columns of F and Ω are d-dimensional vectors, representing the evaluations of respective functions at $\overline{\lambda}_1,\ldots,\overline{\lambda}_d$. Also, columns of $[w_1,\ldots,w_d]$ in (3.9) are mutually orthogonal, but it is not the case for columns of Ω.

Algorithm 2 is modified from Algorithm 1 with the desire to avoid the computation of Q. But now it can be realized as the following trivial fact which, nonetheless, fundamentally changes the tactics of computation for the finite representation (3.4) as we shall describe in Section 3.3.

THEOREM 3.3. *The sample matrix F is Hermitian and positive definite and the relationship (3.12) is the unique Cholesky decomposition of F with positive diagonal entries in R.*

Proof. Given the Vandermonde matrix V_∞, it is easy to see that $F = V_\infty^* V_\infty$. With the QR factorization $V_\infty = QR$, we also have $F = R^*R$ which gives rise to the Cholesky decomposition. Since R has positive diagonal entries, the uniqueness follows. □

3.3. LDL^* decomposition. As indicated in (3.4), we only need R in order to compute the finite dimensional representation $\widetilde{H}$. The above observation suggests that we may achieve this goal by computing the Cholesky decomposition of F which can easily be formed as is seen in Algorithm 2. However, we note that the matrix F might be ill-conditioned, especially when $\lambda_1,\ldots,\lambda_d$ are clustered (see Example 2 in Section 4). When this happens, it is known that floating-point errors may cause the Cholesky decomposition to break down.

To avoid taking square roots in the intermediate steps of computing the Cholesky decomposition and, more so, to exploit its ability in handling indefinite matrices, we suggest the called LDL^* decomposition as an alternative way of computing R. The general theory of LDL^* decomposition can be found in [19]. An efficient and high precision implementation is discussed in [4] and is available in MATLAB. In particular, a pivoting strategy to improve stability has been employed in the algorithm for the LDL^* decomposition. Such a pivoting corresponds to a rearrangement of the values $\lambda_1,\ldots,\lambda_d$ in our application.

THEOREM 3.4. *There exist a permutation σ of the integers $1, \ldots, d$, a unit lower triangular matrix L, and a real diagonal matrix D such that*

$$P^\top F P = L D L^* \tag{3.13}$$

where P is the permutation matrix corresponding to σ and the (k, ℓ) entry of the matrix $P^\top F P$ is precisely $f_{\sigma(\ell)}(\overline{\lambda}_{\sigma(k)})$.

Once the LDL^* decomposition of F is achieved, we may use the information provided by D to decide whether some extremely small diagonal entries should be discarded, thus modifying the prescribed rank d to a smaller value. This seems to be a reasonable way of estimating the rank. Numerical experiments seem to support this approach.

Algorithm 3 LDL^* approach.

Require: complex numbers $\lambda_1, \ldots, \lambda_d$ {from the Vandermonde decomposition of a Hankel matrix H of rank d}

Ensure: Hermitian matrix F, unit lower triangular matrix L, diagonal matrix D, and permutation vector P

$\Lambda := [\lambda_1, \ldots, \lambda_d]$
$F := 1./(1 - \Lambda' * \Lambda)$
[L,D,P] = ldl(F,'vector')
{The following has not been implemented in this work}
if D has extremely small entries **then**
  Reduce d
  Recalculate $\{\lambda_1, \ldots, \lambda_d\}$
end if

4. Numerical experiments. In this section, we carry out some numerical experiments with the purpose to demonstrate the theory discussed above. We have already pointed out a few unsettled issues concerning the step from a given Hankel matrix to its Vandermonde factorization. With that advancing, our work here is to furnish the next step when a Vandermonde factorization and the rank d are known. In reality, because a slightly perturbed low rank matrix is easily of full rank, we have to deal with the low rank approximation with Hankel structure only. That is, though we have a substantial understanding about Hankel operators of finite rank in theory, we constantly encounter the uncertainty of numerical rank which is a fairly difficult subject itself in practice. We wish to use examples to bring forth some of the concerns.

Example 1. We fix $d = 6$ and randomly generate the following vibration parameters

$$[\mathbf{a}, \boldsymbol{\alpha}, \boldsymbol{\nu}, \boldsymbol{\phi}] = \begin{bmatrix} 3.8566 & 0.7923 & 0.3948 & 5.7665 \\ 0.1038 & 3.0421 & 51.2192 & 4.4898 \\ 3.1682 & 0.6764 & 81.2621 & 3.4089 \\ 3.7440 & 0.3534 & 61.2526 & 0.8933 \\ 2.4925 & 2.7414 & 72.1755 & 2.3458 \\ 1.1240 & 3.8136 & 29.1876 & 4.2357 \end{bmatrix}$$

whose columns represent the magnitudes a_ℓ, the decay rates α_ℓ, the frequencies ν_ℓ, and the phase angles ϕ_ℓ, $\ell = 1, \ldots, 6$, for six sinusoids, respectively. For simplicity, we set $\Delta t = 1$. We use the resulting infinite dimensional Hankel matrix H from the combined signal as our test data. The corresponding Vandermonde factorization (2.9) is readily known. The corresponding λ_ℓ's arranged in ascending order of their moduli are

$$0.0084{+}0.0204\imath, 0.0092{+}0.0468\imath, 0.0291{+}0.0575\imath, -0.3575{+}0.2779\imath, -0.0386{+}0.5070\imath, -0.0115{+}0.7022\imath.$$

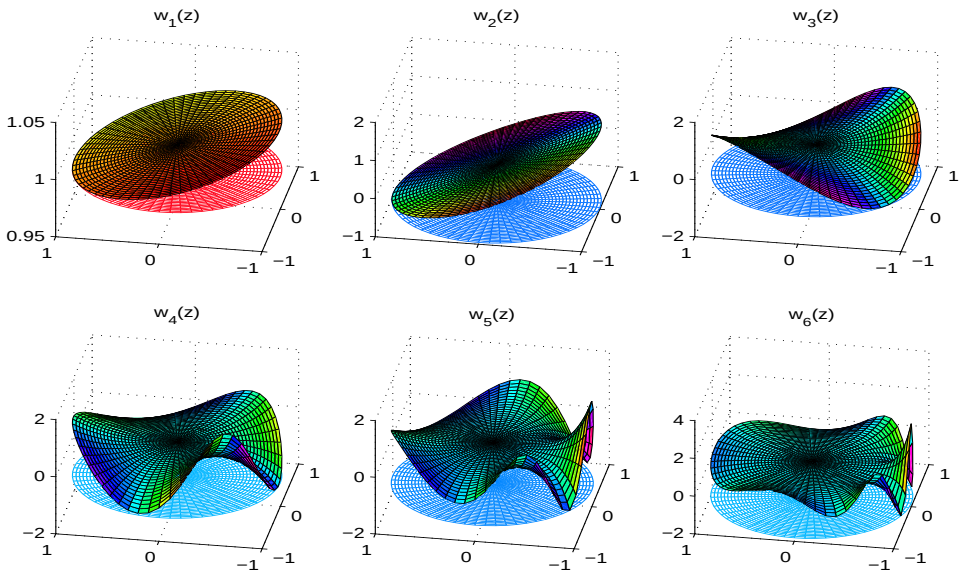

FIG. 4.1. *Sketch of orthonormal functions $w_\ell(z)$ over the unit disc. The height of the surface is the real part, and the color of the surface is the imaginary part varying the hue in the HSV color model.*

We know that the infinite matrix H should be of rank 6, but we want to compute its finite dimensional representation $\widetilde{H}$ and its singular values.

This toy problem is small enough that Algorithm 1 is applicable. For curiosity, we plot in Figure 4.1 the six orthonormal functions $w_\ell(z)$. These plots of complex-valued functions are to be interpreted as follows. Over the domain of the unit disc (displayed underneath the surface in polar coordinates), the real part $\Re(w_\ell(z))$ of $w_\ell(z)$ is represented by the height of the surface, while the color of the surface, varying the hue according to the HSV color model, denotes the imaginary part $\Im(w_\ell(z))$.

We can also use the MATLAB command `ldl` to calculate the LDL^* decomposition of the 6×6 sample matrix F to obtain the upper triangular matrix R needed in (3.4). Built in the routine `ldl` is a pivoting strategy which returns that the permutation $\sigma = [1, 6, 4, 5, 3, 2]$. Rearranging λ_ℓ's in accordance to the suggested permutation σ, we regenerate the orthonormal functions $w_\ell(z)$ by Algorithm 1 and the results are plotted in Figure 4.2. It is interesting to note that the "shapes" of these orthonormal functions are changed, though their practical meaning is not clear at present. In all cases, the six nonzero singular values of H are found to be

$$4.5999e{+}00, 1.8109e{+}00, 2.1203e{-}01, 1.3692e{-}02, 1.2352e{-}04, 7.4630e{-}10.$$

Even though this test is but a small, random example, the effect of clustering λ_ℓ's is evident at the smallest singular value. Had we not known that the original signal is composed of six sinusoids beforehand, it would be a close call to regard the smallest singular value in the above list as nonzero.

Example 2. The real strength of the LDL^* approach is at its efficiency and reliability. First, it is possible to use the SYMBOLIC MATH TOOLBOX in MATLAB to calculate $w_\ell(z)$ via rational arithmetic. But the process is extremely slow. In contrast, the floating-point arithmetic of the MATLAB routine `ldl` applied to the $d \times d$ matrix F is significantly faster. Second, when the rank d is relatively high, we have observed that the value of $r_{\ell\ell}$ obtained by either Algorithms 1 or 2 decreases to nearly zero as ℓ increases, strongly signifying that the modified Gram-Schmidt process is becoming unstable. In contrast, the LDL^* calculation handles this stability issue more smoothly. To demonstrate this advantage, we create a sample matrix F from 100 randomly generated λ_ℓ's over the unit disk. In

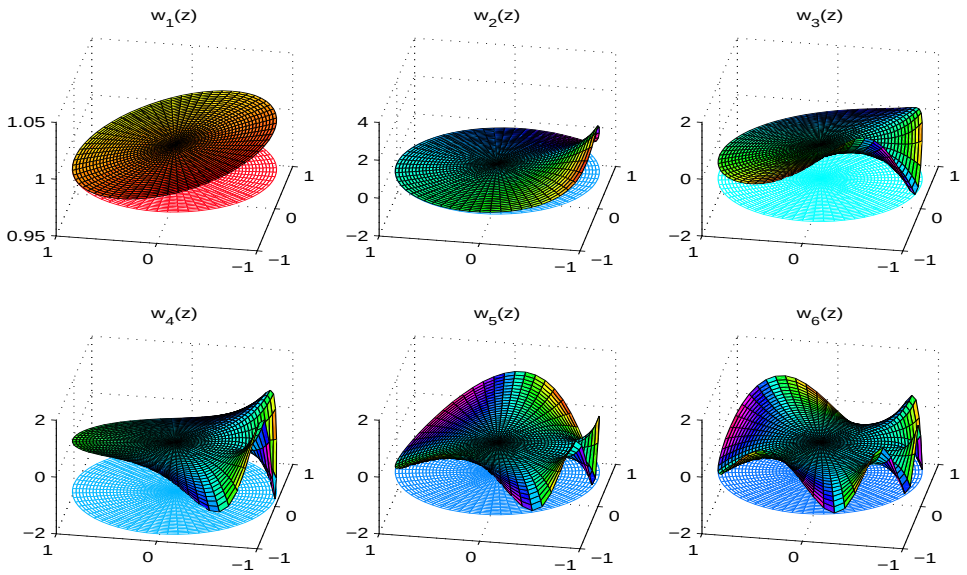

FIG. 4.2. *Sketch of orthonormal functions $w_\ell(z)$ over the unit disc with λ_ℓ rearranged according to LDL^* pivoting.*

theory, F should be of rank 100, and so is the corresponding Hankel matrix. However, when we plot in Figure 4.3 the singular values of F and the absolute values of the diagonal elements of the matrix D from the LDL^* of F, two observations deserve our attention from this experiment. The most noticeable is that there is a considerable agreement between these two sets of values. The other equally noteworthy is that a significant portion of the singular values of F or diagonal elements of D are nearly machine zero. Indeed, numerical experiments indicate that the analogy between singular values of F and the absolute values of diagonal entries of D remains even for low rank d. These observations are not totally surprising because it is known that the LU algorithms, of which the LDL^* is a special case, with suitable pivoting strategies can reveal ranks properly [13].

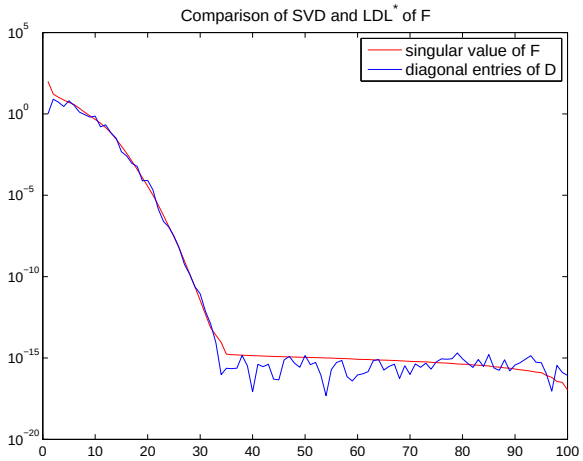

FIG. 4.3. *Comparison*

With this fact in mind, the detection of nearness to zero by diagonal entries of D then suggests

that maybe the initial rank d should be reduced. The implication of this rank reduction is significant because it will cause a chain of reactions. For instance, the equation (2.4) will have to be resolved for a smaller system, which redefines the coefficients $\gamma_0, \ldots, \gamma_{d-1}$ for the characteristic polynomial (2.10) and, consequently, a new Vandermonde factorization (2.15). Equivalently, a new sample matrix F will be generated of which the LDL^* is called again to check the rank condition and the process may be repeated. In terms of the sinusoidal signal (or equivalently its trajectory matrix), rank reduction means that maybe the more complicated initial signal could be approximated with smaller number of components (or a much lower rank Hankel matrix). This rank estimation for Hankel operators, partially related to the subject on singular-value spectral analysis, will be reported in a separate paper.

5. Conclusion. Semi-infinite bounded Hankel matrices of finite rank form an important class of matrices because they occur in applications across multiple fields. This paper intends to make two contributions. First, on the theoretical side, we revisit existing results and establish connection among the various characterizations of finite rank in terms of rational function, recursion, sinusoidal signal, and Vandermonde factorization. Second, on the computational side, we propose to compute the nonzero singular values of the infinite matrix by means of the LDL^* decomposition of a specific finite-dimensional sample matrix.

In all characterizations, though different parameters are used, a semi-infinite bounded Hankel matrix of rank d depends on $2d$ parameters. The Vandermonde factorization enables us to characterize the action of the original infinite dimensional H as an operator via a singular-value preserving transformation to a finite dimensional matrix from which singular values can be calculated.

It is observed that, even though a Hankel is of rank d in exact arithmetic, its numerical rank goes wild as soon as the parameters are replaced by floating-point numbers. The usual modified Gram-Schmidt orthogonalization applied to the generating functions, therefore, becomes increasingly unstable in numerical computation. Under such circumstances, we propose to calculate the upper triangular matrix R needed in the finite dimensional representation, fast and effectively without any orthogonalization, from the LDL^* decomposition of a $d \times d$ sample matrix which is readily available from the Vandermonde factorization. Additionally, the diagonal entries of D might be indicative of rank reduction.

There are many unfinished works to be further studied. A partial list includes the effect on the singular values when the rank d is wrongly assumed, the sensitivity of the computed roots λ_ℓ's subject to perturbations, complexity analysis of the proposed method for singular value computation, and the effectuation of the rank reduction procedure suggested by the LDL^* algorithm.

REFERENCES

[1] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn, *Infinite Hankel matrices and generalized Carathéodory-Fejér and I. Schur problems*, Funkcional. Anal. i Priložen., 2 (1968), pp. 1–17.

[2] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Translated by N. Kemmer, Hafner Publishing Co., New York, 1965.

[3] M. R. Allen and L. A. Smith, *Monte carlo ssa: detecting irregular oscillations in the presence of coloured noise*, Journal of Climate, 9 (1996), pp. 3372–3404.

[4] C. Ashcraft, R. G. Grimes, and J. G. Lewis, *Accurate symmetric indefinite linear equation solvers*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 513–561 (electronic).

[5] M. Auvergne, *Singular value analysis applied to phase space reconstruction of pulsating stars*, Astronomy and Astrophysics, 204 (1988), pp. 341–348.

[6] L. Bezerra, *Vandermonde factorizations of a regular hankel matrix and their application to the computation of bzier curves*, SIAM Journal on Matrix Analysis and Applications, 33 (2012), pp. 411–432.

[7] D. L. Boley, F. T. Luk, and D. Vandevoorde, *Vandermonde factorization of a Hankel matrix*, in Scientific computing (Hong Kong, 1997), Springer, Singapore, 1997, pp. 27–39.

[8] C. M. Bowden, *Boundedness of linear operators in the space ℓ^2*, Int. J. Quantum Chem., 2 (1968), pp. 363–371.

[9] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, *Structured low rank approximation*, Linear Algebra Appl., 366 (2003), pp. 157–172. Special issue on structured matrices: analysis, algorithms and applications (Cortona, 2000).

[10] D. Danilov and A. E. Zhigljavsky, *Principal components of time series: the 'caterpillar' method*, University of St. Petersburg Press. (In Russian), (1997).

[11] T. Dutoit, *Unusual teaching short-cuts to the levinson and lattice algorithms*, in Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on, vol. 5, may 2004, pp. V – 1029–32 vol.5.

[12] S. Feldmann and G. Heinig, *Vandermonde factorization and canonical representations of block Hankel matrices*, in Proceedings of the Fourth Conference of the International Linear Algebra Society (Rotterdam, 1994), vol. 241/243, 1996, pp. 247–278.

[13] L. V. Foster and X. Liu, *Comparison of rank revealing algorithms applied to matrices with well defined numerical rank*, tech. report, San Jose State University, 2012. Notes and Matlab codes available at http://www.math.sjsu.edu/ foster/rankrevealingcode.html.

[14] B. A. Francis, *A course in H_∞ control theory*, vol. 88 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1987.

[15] F. R. Gantmacher, *The theory of matrices. Vols. 1, 2*, Translated by K. A. Hirsch, Chelsea Publishing Co., New York, 1959.

[16] M. Ghil and C. Taricco, *Advanced spectral analysis methods*, In Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models, (1997), pp. 137–159.

[17] J. Gillard and A. Zhigljavsky, *Analysis of structured low rank approximation as an optimization problem*, Informatica (Vilnius), 22 (2011), pp. 489–505.

[18] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, Internat. J. Control, 39 (1984), pp. 1115–1193.

[19] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.

[20] G. H. Golub and J. H. Welsch, *Calculation of Gauss quadrature rules*, Math. Comp. 23 (1969), 221-230; addendum, ibid., 23 (1969), pp. A1–A10.

[21] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of time series structure*, vol. 90 of Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, Boca Raton, FL, 2001. SSA and related techniques.

[22] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, 3rd ed. ed., 2008.

[23] W. B. Gragg and L. Reichel, *On singular values of Hankel operators of finite rank*, Linear Algebra Appl., 121 (1989), pp. 53–70. Linear algebra and applications (Valencia, 1987).

[24] G. Heinig, *Fast and superfast algorithms for Hankel-like matrices related to orthogonal polynomials*, in Numerical analysis and its applications (Rousse, 2000), vol. 1988 of Lecture Notes in Comput. Sci., Springer, Berlin, 2001, pp. 385–392.

[25] I. S. Iohvidov, *Hankel and Toeplitz matrices and forms*, Birkhäuser Boston, Mass., 1982. Algebraic theory, Translated from the Russian by G. Philip A. Thijsse, With an introduction by I. Gohberg.

[26] D. Kalman, *The generalized Vandermonde matrix*, Math. Mag., 57 (1984), pp. 15–21.

[27] A. Y. Kibangou and G. Favier, *Toeplitz ndash;vandermonde matrix factorization with application to parameter estimation of wiener ndash;hammerstein systems*, Signal Processing Letters, IEEE, 14 (2007), pp. 141 –144.

[28] S. Y. Kung, *A new identification and model reduction algorithm via singular value decomposition*, in Proceedings of the 12th Asilomar Conference on Circuits, Systems and Computers, Pacific Grove, CA, November 1978, pp. 705–714.

[29] S. Y. Kung and D. W. Lin, *Recent progress in linear system model-reduction via Hankel matrix approximation*, in Circuit theory and design (The Hague, 1981), North-Holland, Amsterdam, 1981, pp. 222–233.

[30] P. Lemmerling and S. Van Huffel, *Analysis of the structured total least squares problem for Hankel/Toeplitz matrices*, Numer. Algorithms, 27 (2001), pp. 89–114.

[31] F. T. Luk and S. Qiao, *A fast singular value algorithm for Hankel matrices*, in Fast algorithms for structured matrices: theory and applications (South Hadley, MA, 2001), vol. 323 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2003, pp. 169–177.

[32] I. Markovsky, *Low rank approximation*, Communications and Control Engineering Series, Springer, London, 2012. Algorithms, implementation, applications.

[33] I. Markovsky, J. C. Willems, S. Van Huffel, B. De Moor, and R. Pintelon, *Application of structured total least squares for system identification and model reduction*, IEEE Trans. Automat. Control, 50 (2005), pp. 1490–1500.

[34] A. Mineva and D. Popivanov, *Method for single-trial readiness potential identification, based on singular spectrum analysis*, Journal of Neuroscience Methods, 68 (1996), pp. 91 – 99.

[35] Z. Nehari, *On bounded bilinear forms*, Ann. of Math. (2), 65 (1957), pp. 153–162.

[36] V. Y. Pan, *Structured matrices and polynomials*, Birkhäuser Boston Inc., Boston, MA, 2001. Unified superfast algorithms.

[37] H. Park, L. Zhang, and J. B. Rosen, *Low rank approximation of a Hankel matrix by structured total least norm*, BIT, 39 (1999), pp. 757–779.

[38] J. R. Partington, *An introduction to Hankel operators*, vol. 13 of London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 1988.

[39] V. V. Peller, *An excursion into the theory of Hankel operators*, in Holomorphic spaces (Berkeley, CA, 1995), vol. 33 of Math. Sci. Res. Inst. Publ., Cambridge Univ. Press, Cambridge, 1998, pp. 65–120.
[40] ———, *Hankel operators and their applications*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2003.
[41] S. C. Power, *Hankel operators on Hilbert space*, vol. 64 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, Mass., 1982.
[42] J. M. Steele, *The yule-walker equations and algorithms for their solution*, 2012. Class notes and references available at http://www-stat.wharton.upenn.edu/ steele/Courses/956/ResourceDetails/YuleWalkerAndMore.htm.
[43] R. Vautard, P. Yiou, and M. Ghil, *Singular-spectrum analysis: A toolkit for short, noisy chaotic signals*, Physica D: Nonlinear Phenomena, 58 (1992), pp. 95 – 126.
[44] H. Widom, *Hankel matrices*, Trans. Amer. Math. Soc., 121 (1966), pp. 1–35.
[45] P. Yiou, D. Sornette, and M. Ghil, *Data-adaptive wavelets and multi-scale singular-spectrum analysis*, Phys. D, 142 (2000), pp. 254–290.
[46] N. J. Young, *The singular-value decomposition of an infinite Hankel matrix*, Linear Algebra Appl., 50 (1983), pp. 639–656.