Frequency-Selective Vandermonde Decomposition of Toeplitz Matrices With Applications

Zai Yang, Member, IEEE, and Lihua Xie, Fellow, IEEE

Abstract—The classical result of Vandermonde decomposition of positive semidefinite Toeplitz matrices can date back to the early twentieth century. It forms the basis of modern subspace and recent atomic norm methods for frequency estimation. In this paper, we study the Vandermonde decomposition in which the frequencies are restricted to lie in a given interval, referred to as frequency-selective Vandermonde decomposition. The existence and uniqueness of the decomposition are studied under explicit conditions on the Toeplitz matrix. The new result is connected by duality to the positive real lemma for trigonometric polynomials nonnegative on the same frequency interval. Its applications in the theory of moments and line spectral estimation are illustrated. In particular, it provides a solution to the truncated trigonometric \( K \)-moment problem. It is used to derive a primal semidefinite program formulation of the frequency-selective atomic norm in which the frequencies are known \textit{a priori} to lie in a certain frequency band. Numerical examples are also provided.

Index Terms—Frequency-selective Vandermonde decomposition, Toeplitz matrix, truncated trigonometric \( K \)-moment problem, line spectral estimation, atomic norm.

I. INTRODUCTION

The classical Vandermonde decomposition result discovered by Carathéodory and Fejér in 1911 \[1\] states that, if an \( N \times N \) Toeplitz matrix \( T \) is positive semidefinite (PSD) and has rank of \( r \leq N \), then it can be factorized as

\[
T = APA^H,
\]

where \( P \) is an \( r \times r \) positive definite diagonal matrix and \( A \) is an \( N \times r \) Vandermonde matrix whose columns correspond to uniformly sampled complex sinusoids with different frequencies. Moreover, the decomposition is unique if \( r < N \). In the theory of moments, the Vandermonde decomposition provides a solution to the truncated trigonometric moment problem (a.k.a., the moment problem on the unit circle given a finite moment sequence) \[2\]. It is also important in operator theory and has application in system theory \[3\], \[4\]. It has become important in the area of data analysis and signal processing since the 1970s when it was rediscovered by Pisarenko and used for frequency retrieval from the data covariance matrix \[5\]. From then on, the Vandermonde decomposition has formed the basis of a prominent subset of methods for frequency estimation designated as subspace methods, e.g., multiple signal classification (MUSIC) and estimation of parameters by rotational invariant techniques (ESPRIT) (see the review in \[6\]).

The new interest in Vandermonde decomposition of Toeplitz matrices has recently been evoked due to the spectral super-resolution and continuous compressed sensing framework, in which the frequency estimation problem is solved by exploiting sparsity that arises from the fact that the number of frequencies is small and optimization methods are formulated based on the Vandermonde decomposition \[7\]–\[9\]. To be specific, the atomic norm \[10\], which can be viewed as a continuous counterpart of the \( \ell_1 \) norm utilized in previous compressed sensing based methods \[11\], \[12\], is utilized and cast as semidefinite programming (SDP) by applying the Vandermonde decomposition. The atomic norm method is superior in the sense that, unlike the compressed sensing methods, it does not require to grid/discretize the frequency domain, completely resolves the grid mismatch problem, and provides strong theoretical guarantees. The Vandermonde decomposition has also been generalized to high dimensions and used for multidimensional frequency estimation \[13\].

We note that the frequencies in the Vandermonde decomposition in \[1\] may take any value in the normalized frequency domain \([0, 1]\) in which the starting and ending points are identified. In this paper, we ask the following question:

Can the frequencies in the Vandermonde decomposition of \( T \) be restricted to lie in a given interval \( \mathcal{I} \subset [0, 1] \), instead of the entire domain \([0, 1]\), under explicit conditions on \( \mathcal{I} \)?

The resulting decomposition is referred to as frequency-selective (FS) Vandermonde decomposition. The problem is challenging in that by \[1\] \( T \) is a highly nonlinear function of the frequencies and it is unclear how to link \( T \) to a frequency interval \( \mathcal{I} \).

In mathematics, similar questions are asked in a class of moment problems known as truncated \( K \)-moment problems, a.k.a., truncated moment problems on a semialgebraic set \( K \), instead of on an entire domain \([0, 1]\). The truncated \( K \)-moment problems have been solved in the real and complex domains \[15\], \[16\]; however, to the best of our knowledge, it is still open on the unit circle, known as the truncated trigonometric \( K \)-moment problem. In this paper, we show that the study of FS Vandermonde decomposition can provide a solution to this open problem.

Besides mathematical interest, the investigation of FS Vandermonde decomposition is important and has great potentials in practical frequency estimation problems in which the frequencies are known \textit{a priori} to lie in a given interval. For
example, the frequencies will be known to lie in a narrower band when oversampling happens. A radar engineer might be able to estimate the maximum range/delay of a detectable aircraft due to path loss. The maximum Doppler frequency can be obtained given the aircraft’s characteristic speed. In underwater channel estimation, the frequency parameters of interest can reside in a known small interval \([17]\). Similar prior knowledge might also be available given weather observations \([18]\). It has recently been shown in \([19, 20]\) that the frequency estimation performance can be improved by exploiting such prior knowledge based on atomic norm techniques.

In this paper, we provide an affirmative answer to the question asked above. Specifically, we show that a PSD Toeplitz matrix \(T\) admits a FS Vandermonde decomposition on a given interval if and only if \(T\) satisfies another linear matrix inequality (LMI). Interestingly, this FS Vandermonde decomposition result is linked by duality to the positive real lemma (PRL) for trigonometric polynomials \([21]\). The usefulness of the new result is demonstrated. In the theory of moments, it satisfies another linear matrix inequality. Similar prior interest can reside in a known small interval \([17]\). Similar prior knowledge can be obtained given the aircraft’s characteristic speed. In underwater channel estimation, the frequency parameters of interest can reside in a known small interval \([17]\). Similar prior knowledge might also be available given weather observations \([18]\). It has recently been shown in \([19, 20]\) that the frequency estimation performance can be improved by exploiting such prior knowledge based on atomic norm techniques.

In this paper, we provide an affirmative answer to the question asked above. Specifically, we show that a PSD Toeplitz matrix \(T\) admits a FS Vandermonde decomposition on a given interval if and only if \(T\) satisfies another linear matrix inequality (LMI). Interestingly, this FS Vandermonde decomposition result is linked by duality to the positive real lemma (PRL) for trigonometric polynomials \([21]\). The usefulness of the new result is demonstrated. In the theory of moments, it satisfies another linear matrix inequality. Similar prior interest can reside in a known small interval \([17]\). Similar prior knowledge can be obtained given the aircraft’s characteristic speed. In underwater channel estimation, the frequency parameters of interest can reside in a known small interval \([17]\). Similar prior knowledge might also be available given weather observations \([18]\). It has recently been shown in \([19, 20]\) that the frequency estimation performance can be improved by exploiting such prior knowledge based on atomic norm techniques.

In this paper, we provide an affirmative answer to the question asked above. Specifically, we show that a PSD Toeplitz matrix \(T\) admits a FS Vandermonde decomposition on a given interval if and only if \(T\) satisfies another linear matrix inequality (LMI). Interestingly, this FS Vandermonde decomposition result is linked by duality to the positive real lemma (PRL) for trigonometric polynomials \([21]\). The usefulness of the new result is demonstrated. In the theory of moments, it satisfies another linear matrix inequality. Similar prior interest can reside in a known small interval \([17]\). Similar prior knowledge can be obtained given the aircraft’s characteristic speed. In underwater channel estimation, the frequency parameters of interest can reside in a known small interval \([17]\). Similar prior knowledge might also be available given weather observations \([18]\). It has recently been shown in \([19, 20]\) that the frequency estimation performance can be improved by exploiting such prior knowledge based on atomic norm techniques.
Therefore, following from the result in the case of \( r \leq N - 1 \) that we just proved, \( T(t') \) admits a Vandermonde decomposition as in (11) with \( r = N - 1 \). It then follows from (11) that \( T \) admits an \( N \)-atomic Vandermonde decomposition.

We finally show the uniqueness in the case of \( r \leq N - 1 \). Write \( T = A(f)PA^H(f) \), where \( P = \text{diag}(p_1, \ldots, p_r) \) and \( A(f) = [a(f_1), \ldots, a(f_r)] \). Suppose there exists another decomposition: \( T = A(f')P'A^H(f') \) with, similarly, \( p'_j > 0 \), \( j = 1, \ldots, r \) and \( f'_j \in \mathbb{T} \) are distinct. It follows from the equation \( A(f')P'A^H(f') = A(f)P^2A^H(f) \) that there exists an \( r \times r \) unitary matrix \( U' \) such that \( A(f')P^2A^H(f) \) and \( A(f)P^2A^H(f) \) are identical.

By the construction of \( g(z) \) we know that \( g(z) \) has two single roots \( z_L \) and \( z_H \), or equivalently, \( g(f) \) has two single roots \( f_L \) and \( f_H \). It follows that \( g(f) \) changes its sign around \( f_L \) and around \( f_H \). So there are two possibilities: \( g(f) \) is positive on \( (f_L, f_H) \) and negative on \( (f_H, f_L) \), or \( g(f) \) is negative on \( (f_L, f_H) \) and positive on \( (f_H, f_L) \). This can be checked by studying the value of \( g \) at \( f = \frac{1}{2}(f_L + f_H) \). In particular, we have

\[
g\left(\frac{1}{2}(f_L + f_H)\right) = r_0 \pm 2\Re \left( r_1 e^{-i\pi(f_L + f_H)} \right) = \{2 - 2 \cos |\pi(f_L - f_H)|\} \text{sgn}(f_H - f_L).
\]

This means that at \( f = \frac{1}{2}(f_L + f_H) \), the sign of \( g(f) \) is identical to that of \( f_H - f_L \). As a result, we conclude that \( g(f) \) is always positive on \( (f_L, f_H) \) and negative on \( (f_H, f_L) \) whenever \( f_L < f_H \) or \( f_L > f_H \).

B. FS Vandermonde Decomposition

The FS Vandermonde Decomposition of Toeplitz matrices is summarized in the following theorem.

**Theorem 2:** Given \( \mathcal{I} \subset \mathbb{T} \), a Toeplitz matrix \( T \in \mathbb{C}^{N \times N} \) admits a FS Vandermonde decomposition as in (11) but with \( f_k \in \mathcal{I} \), if and only if

\[
T \geq 0, \quad T_g \geq 0,
\]

where \( g \) is defined by (14)-(16) and \( T_g \) by (3). Moreover, the decomposition is unique if either \( T \) or \( T_g \) is rank-deficient.

**Proof:** For the “if” part, we first consider the case of \( r \leq N - 1 \). Under (18), it follows from Theorem 1 that \( T \) admits a unique Vandermonde decomposition as in (11). It suffices to show that \( f_k \in \mathcal{I}, k = 1, \ldots, r \) under (19). To do so, by (11) we have that

\[
t_{n-m} = T_{mn} = \sum_{k=1}^{r} p_k e^{i2\pi(m-n)f_k}.
\]

It follows that

\[
[T_g]_{mn} = \sum_{j=-\infty}^{\infty} r_je^{f_{n-m+j}} = \sum_{j=-\infty}^{\infty} r_j \sum_{k=1}^{r} p_k e^{i2\pi(m-n-j)f_k}
\]

\[
= \sum_{k=1}^{r} p_k e^{i2\pi(m-n)f_k} \sum_{j=1}^{r} r_j e^{-i2\pi j f_k}
\]

\[
= \sum_{k=1}^{r} p_k g(f_k) e^{i2\pi(m-n)f_k}.
\]

Hence, we have that

\[
T_g = \sum_{k=1}^{r} p_k g(f_k) a(N-1, f_k) a^H(N-1, f_k)
\]

\[
= A(N-1, f) \text{diag}(p_1 g(f_1), \ldots, p_r g(f_r)) \times A^H(N-1, f),
\]
where \( A (N-1, f) := \{ a (N-1, f_1), \ldots, a (N-1, f_r) \} \in \mathbb{C}^{(N-1) \times r} \) and diag \((p_1 g (f_1), \ldots, p_r g (f_r))\) denotes a diagonal matrix with \(p_k g (f_k), k = 1, \ldots, r\) on the diagonal. Note that \( A (N-1, f) \) has full column rank since \( r \leq N-1 \). It then follows from \((22)\) and \((19)\) that
\[
\text{diag} (p_1 g (f_1), \ldots, p_r g (f_r)) = A^\dagger (N-1, f) T g A^H (N-1, f) \geq 0,
\]
where \(^\dagger\) denotes the matrix pseudo-inverse. This means that \(p_k g (f_k) \geq 0\) and thus \( g (f_k) \geq 0\), \( k = 1, \ldots, r\). So we conclude that \(f_k \in \mathbb{I}, k = 1, \ldots, r\) by the property of \(g (f)\).

We now consider the case of \( r = N \) in which \( T > 0 \). Let \( f_N = f_L = 0 \) and \( p_N = (a^H (f_N) T^{-1} a (f_N))^{-1} > 0 \). Similarly to the proof of Theorem \([1]\) we define a new sequence \( t' = [t'_j], |j| \leq N-1 \) by \((7)\), followed by \((8), (9)\) and \((10)\). Moreover, similar to \((21)\) we have that
\[
[T_g (t')]_{mn} = \sum_{j=-1}^{1} r_j t'_m - n + j
\]
and hence,
\[
T_g (t') = T_g - p_N g (f_N) a (N-1, f_N) a^H (N-1, f_N).
\]
Using the fact that \( g (f_N) = g (f_L) = 0 \) and by \((19)\), we have that
\[
T_g (t') = T_g \geq 0.
\]
Now consider \( T (t') \) that satisfies \((9), (10)\) and \((20)\). Using the “if” part of Theorem \([2]\) in the case of \( r \leq N-1 \) that we just proved, \( T (t') \) admits a unique decomposition as in \([4]\), with \(f_k \in \mathbb{I}, k = 1, \ldots, r = N-1\). Therefore, it follows from \((8)\) that
\[
T = T (t') + p_N g (f_N) a^H (f_N)
\]
has a decomposition as in \([4]\), with \(f_k \in \mathbb{I}, k = 1, \ldots, r = N\). So we complete the “if” part.

The “only if” part can be shown by similar arguments. In particular, given \( T \) as in \([4]\), it is obvious that \((18)\) holds. Moreover, we still have \((22)\), in which \( g (f_k) \geq 0\), \( k = 1, \ldots, r\) since \(f_k \in \mathbb{I}, k = 1, \ldots, r\). So \((19)\) holds as well.

We finally prove the uniqueness under the additional assumption that \( T \) or \( T_g \) is rank-deficient. In the case when \( T \) is rank-deficient this is a direct consequence of the standard Vandermonde decomposition result. In the other case when \( T \) has full rank and \( T_g \) is rank-deficient, we first note that there are at least 2 distinct \( f_k \)'s in the FS Vandermonde decomposition of \( T \). Now let us recall \((22)\). Since \( A (N-1, f) \) has full row rank, by the property of \( g (f)\), \( T_g \) is rank-deficient only if \( g (f_k) \neq 0\) holds for maximally \( N-2\), \( f_k \)'s and the other \( f_k \)'s must be either \( f_L \) or \( f_H \). This means that there are exactly \( N \) distinct atoms in the decomposition with two frequencies fixed at \( f_L \) and \( f_H \). It follows that the other \( N-2 \) frequencies are also fixed. Therefore, the FS Vandermonde decomposition is unique.

We now discuss how to obtain the FS Vandermonde decomposition when the assumptions of Theorem \([2]\) are satisfied. Note first that in the case when \( T \) is rank-deficient, it admits a unique decomposition that can be computed as in Section \([1]\). When \( T \) has full rank, by Theorem \([2]\) an \( N \)-atomic decomposition is guaranteed to exist that can be computed following from the proof of Theorem \([2]\) to be specific, fix \( f_N \) and \( p_N \) first and then obtain the decomposition following the proof.

### IV. Duality

Using the FS Vandermonde decomposition result shown in the previous section, we can explicitly characterize the cone of Toeplitz matrices admitting such decompositions. Due to the interest in optimization problems, we naturally look at the dual cone, which, as we will see, enables us to link the FS Vandermonde decomposition to the theory on trigonometric polynomials, to be specific, the PRL given in \([24], [25]\) (see also \([21]\)).

For a sequence \( t = [t_j], |j| \leq N-1 \) with \( t_{-j} = \bar{t}_j \), let \( t_R = \begin{bmatrix} R t_N, \ldots, R t_1, \frac{\sqrt{2}}{\sqrt{\pi}} t_0, \Im t_1, \ldots, \Im t_{N-1} \end{bmatrix}^T \in \mathbb{R}^{2N-1} \) be a representation of \( t \) in the real domain, where the coefficient \( \sqrt{\frac{2}{\pi}} \) for \( t_0 \) is chosen for convenience. It is obvious that all \( N \times N \) Toeplitz matrices admitting a FS Vandermonde decomposition on a given interval \( I \subset \mathbb{T} \) form a cone that can be identified with the cone
\[
\mathcal{K}_{\text{VDF}} := \left\{ t_R : T = \sum p_k a (f_k) a^H (f_k), p_k \geq 0, f_k \in I \right\}
\]
(28) Define
\[
\mathcal{K}_{\text{VDM}} := \left\{ t_R : T \geq 0, T_g \geq 0 \right\}
\]
(29) where \( g \) is defined as in Theorem \([2]\). A direct consequence of Theorem \([2]\) is that
\[
\mathcal{K}_{\text{VDF}} = \mathcal{K}_{\text{VDM}}.
\]
(30)

We next consider the dual cone of \( \mathcal{K}_{\text{VDF}} \) defined as \([26]\)
\[
\mathcal{K}_{\text{VDF}}^* := \left\{ \alpha \in \mathbb{R}^{2N-1} : t_R^T \alpha \geq 0 \text{ for any } t_R \in \mathcal{K}_{\text{VDF}} \right\}
\]
(31) Before proceeding to the main result of this section, we first introduce some notations. Let
\[
\mathcal{K}_{\text{PDL}} := \left\{ \gamma_R : \frac{1}{\gamma_j} = \sum_{j=1}^{N-1} \gamma_j e^{i2\pi j f} \geq 0, f \in I \right\}
\]
(32) denote the cone of trigonometric polynomials of order \( N-1 \) and nonnegative on \( I \), where \( \gamma_R \) is similarly defined as \( t_R \). Also let \( \Theta_j, |j| \leq N-1 \) be an \( N \times N \) elementary Toeplitz matrix with ones on its \( j \)th diagonal and zeros elsewhere. With respect to \( \Theta_j \) and the trigonometric polynomial \( g \) defined by \([14], [16]\), we define the \((N-1) \times (N-1)\) Toeplitz matrix \( \Theta_g \), like \( T_g \) with respect to \( T \). By definition, it is easy to
verify that
\[ T = \sum_{j=1}^{N-1} \Theta_j t_j, \quad (33) \]
\[ T_g = \sum_{j=1}^{N-1} \Theta_{g_j} t_j. \quad (34) \]

We also define the cone
\[ K_{\text{PolM}} := \left\{ \gamma_R : \gamma_{-j} = \text{tr} (\Theta_j Q_0) + \text{tr} [\Theta_{g_j} Q_1], \right. \]
\[ |j| \leq N - 1, \quad Q_0 \in \mathbb{C}^{N \times N}, Q_1 \in \mathbb{C}^{(N-1) \times (N-1)}, \quad Q_0 \geq 0, Q_1 \geq 0 \}. \quad (35) \]

The main result of this section is given in the following theorem.

**Theorem 3:** We have the following identities:
\[ K_{\text{VDF}} = K_{\text{PolM}}, \quad (36) \]
\[ K_{\ast \text{PolM}} = K_{\text{VDM}}. \quad (37) \]

Therefore, provided that \( K_{\text{VDF}} = K_{\text{VDM}} \) we can conclude that \( K_{\text{PolF}} = K_{\text{PolM}} \) and vice versa.

**Proof:** We first show (36). Note that \( t_R \in K_{\text{VDF}} \) if and only if
\[ t_j = \sum_k p_k e^{-i2\pi j f_k}, \quad j = 1, N, \ldots, N - 1, \quad (38) \]
where \( p_k \geq 0 \) and \( f_k \in \mathcal{I} \). For any \( \alpha = [\alpha_1, \ldots, \alpha_N, \ldots, \alpha_{N-1}]^T \in \mathbb{R}^{2N-1} \), we define \( \gamma \in \mathbb{C}^{2N-1} \) such that \( \gamma_0 = \sqrt{2} \alpha_0 \), \( \gamma_j = \alpha_{-j} + i \alpha_j \) and \( \gamma_{-j} = \alpha_{-j} - i \alpha_j \), \( j = 1, \ldots, N - 1 \). It follows that \( \alpha = \gamma_R \) and
\[ t_R^T \alpha = \frac{\sqrt{2}}{2} t_0 \cdot \frac{\sqrt{2}}{2} \gamma_0 + \sum_{j=1}^{N-1} t_j \gamma_j = 1/2 \sum_{j=1}^{N-1} t_j \gamma_j. \quad (39) \]

Inserting (38) into (39), we have that
\[ t_R^T \alpha = \frac{1}{2} \sum_k p_k \sum_{j=1}^{N-1} \gamma_j e^{i2\pi j f_k}. \quad (40) \]

By (40) and the definition of the dual cone, \( \alpha = \gamma_R \in K_{\ast \text{VDF}} \) if and only if \( h(f) := \sum_{j=1}^{N-1} \gamma_j e^{i2\pi j f} \) is nonnegative on \( \mathcal{I} \), or equivalently, \( \alpha \in K_{\text{PolF}} \) by (32).

To show (37), we can similarly define \( t \) for \( \alpha \in \mathbb{R}^{2N-1} \) such that \( \alpha = t_R \). For \( \gamma_R \in K_{\ast \text{PolM}} \),
\[ \gamma_R^T \alpha = \frac{1}{2} \sum_{j=1}^{N-1} \gamma_j t_j = \frac{1}{2} \sum_{j=1}^{N-1} \gamma_j t_j = \frac{1}{2} \sum_{j=1}^{N-1} t_j \{ \text{tr} (\Theta_j Q_0) + \text{tr} [\Theta_{g_j} Q_1] \}. \quad (41) \]

Using the identities in (33) and (34), we have that
\[ \gamma_R^T \alpha = \frac{1}{2} \text{tr} (T Q_0) + \frac{1}{2} \text{tr} (T_g Q_1). \quad (42) \]

Therefore, \( \alpha = t_R \in K_{\ast \text{PolM}} \) if and only if \( \text{tr} (T Q_0) + \text{tr} (T_g Q_1) \geq 0 \) for any \( Q_0 \geq 0 \) and \( Q_1 \geq 0 \), which holds if and only if \( T \geq 0 \) and \( T_g \geq 0 \), or equivalently, \( \alpha \in K_{\text{VDM}} \). Finally, provided that \( K_{\text{VDF}} = K_{\text{VDM}} \) using (36) and (37), we have that
\[ K_{\text{PolF}} = K_{\ast \text{VDF}} = K_{\ast \text{PolM}} = K_{\ast \text{PolM}}. \quad (43) \]

Using the identity that \( K_{\ast \text{PolM}} = K_{\text{PolM}} \), which follows from the fact that \( K_{\text{PolM}} \) is convex and closed (29), we conclude that \( K_{\text{PolF}} = K_{\text{PolM}} \). By similar arguments we can also show that \( K_{\text{VDF}} = K_{\text{VDM}} \) provided that \( K_{\text{PolF}} = K_{\text{PolM}} \).

By Theorem 3, the FS Vandermonde decomposition on \( \mathcal{I} \) is linked via duality to the trigonometric polynomials nonnegative on the same interval. Moreover, the identity that \( K_{\text{PolF}} = K_{\text{PolM}} \) provides a matrix form parametrization of the coefficients of these polynomials. In fact, this is nothing but the Gram matrix parametrization concluded by the PRL in [24], [25] (see also [21]). This means that the PRL in [24], [25] can be obtained from the FS Vandermonde decomposition; conversely, the PRL also provides an alternative way to characterize the set of Toeplitz matrices admitting a FS Vandermonde decomposition.

**Remark 1:** The trigonometric polynomial \( g(z) = r_{-1} z + r_0 + r_1 z^{-1} \) that is nonnegative on \( \mathcal{I} \) and negative on its complement plays an important role in both the FS Vandermonde decomposition of Toeplitz matrices and the Gram matrix parametrization of trigonometric polynomials. It is worth noting that the polynomial defined in the present paper (recall (14)-(16)) is different from those in [21], [24], [25]. As a matter of fact, while the polynomial we define applies uniformly to all intervals \( \mathcal{I} \in \mathcal{T} \), certain modifications to the polynomial or additional operations such as sliding the interval have to be taken in other papers when \( \mathcal{I} \) contains certain critical points such as 0 (or 1) and \( 1/2 \).

V. APPLICATION IN THE THEORY OF MOMENTS

A. Problem Statement

For a given sequence \( t_j, |j| \leq N - 1 \) and a given domain \( F \), a truncated moment problem entails determining whether there exists a positive Borel measure \( \mu \) on \( F \) such that (2)
\[ t_j = \int_F z^j d\mu (z), \quad |j| \leq N - 1. \quad (44) \]

1Note that Theorem 2 is stronger in the sense that it concludes that all such Toeplitz matrices always admit a decomposition containing \( N \) atoms or less.
The problem is further referred to as a truncated K-moment problem if μ is constrained to be supported on a semialgebraic set $K \subset \mathbb{T}$, i.e.,
$$\text{supp}(\mu) \subset K.$$  \hspace{1cm} (45)
A measure $\mu$ satisfying (44) is a representing measure for $t$; $\mu$ is a $K$-representing measure if it satisfies (44) and (45).

The truncated moment and K-moment problems have been solved when $F$ is the real or the complex domain (note that the complex moment problem is defined slightly differently from (44) [15], [16], [27]). The truncated moment problem is also solved when $F$ is the unit circle, known as the truncated trigonometric moment problem [3], [27]. In fact, the solution is also solved when $F$ admits a Vandermonde decomposition of Toeplitz matrices: A representing measure $\mu$ exists if and only if the Toeplitz matrix $T_0^*$ formed using $\mu$-representing measure $\mu$ exists if and only if there exist sequences $t_l$, $l = 1, \ldots, J$ such that
$$\text{supp}(\mu) \subset K = \bigcup_{l=1}^{J} [f_{L_l}, f_{H_l}] \subset \mathbb{T}.$$ \hspace{1cm} (47)

Note that a semialgebraic set $K$ on the unit circle $\mathbb{T}$ can be identified with the union of finite disjoint subintervals $[f_{L_l}, f_{H_l}] \subset \mathbb{T}$, $l = 1, \ldots, J$. Therefore, the moment problem of interest can be restated as follows. For a given sequence $t_j$, $|j| \leq N - 1$, the truncated trigonometric $K$-moment problem entails determining whether there exists a $K$-representing measure $\mu$ on $\mathbb{T}$ satisfying that
$$t_j = \int_{\mathbb{T}} e^{i2\pi j f} d\mu(f), \quad |j| \leq N - 1,$$
$$\text{supp}(\mu) \subset K = \bigcup_{l=1}^{J} [f_{L_l}, f_{H_l}] \subset \mathbb{T}.$$ \hspace{1cm} (46)

**Lemma 1:** An $r$-atomic $K$-representing measure $\mu$ for $t$ exists if and only if $T$ admits an $r$-atomic FS Vandermonde decomposition on $K$.

We next provide explicit conditions on $T$ by applying Theorem 4. In the case when $K$ is a single interval, the following theorem is a direct consequence by combining Lemma 1 and Theorem 2.

**Theorem 4:** Given $K = [f_{L_l}, f_{H_l}]$, an $r$-atomic $K$-representing measure $\mu$ for $t$ exists if and only if if (18) and (19) hold, where $r = \text{rank}(T)$, and $g$ is defined by (14)-(16). Moreover, $\mu$ can be found by computing the FS Vandermonde decomposition of $T$ on $K$, and it is unique if $T$ is rank-deficient.

In the case of $K = \bigcup_{l=1}^{J} [f_{L_l}, f_{H_l}]$, where $[f_{L_l}, f_{H_l}] \subset \mathbb{T}$, $l = 1, \ldots, J \geq 2$ are disjoint, we have the following corollary of Theorem 4, the proof of which is straightforward and thus is omitted.

**Corollary 1:** Given $K = \bigcup_{l=1}^{J} [f_{L_l}, f_{H_l}]$, a $K$-representing measure $\mu$ for $t$ exists if and only if there exist sequences $t_l$, $l = 1, \ldots, J$ such that
$$\sum_{l=1}^{J} t_l = t,$$ \hspace{1cm} (51)
$$T(t_l) \geq 0,$$ \hspace{1cm} (52)
$$T_{g_l}(t_l) \geq 0, \quad l = 1, \ldots, J,$$ \hspace{1cm} (53)
where $g_l$, $l = 1, \ldots, L$ are defined with respect to $[f_{L_l}, f_{H_l}]$, respectively.

**Remark 2:** By solving (53), the number of atoms in the obtained representing measure, if it exists, can be as large as $NJ$. To possibly reduce the number of atoms, we can solve (53) for $T_0^*$, the sum of which finally form a $K$-representing measure for $t$. If (53) is infeasible, then no $K$-representing measure for $t$ exists.

**Remark 3:** By solving (53), the number of atoms in the obtained representing measure, if it exists, can be as large as $NJ$. To possibly reduce the number of atoms, we can find the one minimizing certain convex function of $t_l$, $l = 1, \ldots, J$ among the solutions satisfying (51)-(53). As an example, the function may be chosen as $\pm \text{tr}(T_0^*)$.

Finally, it is interesting to note that the dual problem of (53) can be easily obtained using the result in Section IV. Using the cone notations (54) can be written as:

Find $t_{l, 1} \in \mathcal{K}_{VDM, 1}$, \hspace{1cm} (54)
$$l = 1, \ldots, J,$$
subject to
$$\sum_{l=1}^{J} t_{l, 1} = t_R,$$ \hspace{1cm} (55)

This means that $T$ admits an $r$-atomic FS Vandermonde decomposition on $K$. It is easy to show that the above arguments also hold conversely. So we conclude the following result.

**Lemma 1:** An $r$-atomic $K$-representing measure $\mu$ for $t$ exists if and only if $T$ admits an $r$-atomic FS Vandermonde decomposition on $K$.
where \( t_{R,l} := [t_l]_{R^*} \) and \( K_{\text{VDM,l}} \) denotes \( K_{\text{VDM,l}} \) in (29) with \( g \) being \( g_l \). The Lagrangian function is given by:

\[
\mathcal{L}(t_{R,1}, \ldots, t_{R,J}, \alpha) = \left( \sum_{l=1}^{J} t_{R,l} - t_R \right)^T \alpha = \sum_{l=1}^{J} t_{R,l}^2 \alpha - t_R^2 \alpha,
\]

(56)

where \( t_{R,l} \in K_{\text{VDM,l}}, l = 1, \ldots, J, \) and \( \alpha \) is the Lagrangian multiplier. Using the knowledge of dual cone, we have that

\[
\min_{t_{R,l} \in K_{\text{VDM,l}}} \mathcal{L} = \begin{cases} -t_{R,l}^2 \alpha, & \text{if } \alpha \in K_{\text{VDM,l}}, l = 1, \ldots, J; \\
-\infty, & \text{otherwise}. \end{cases}
\]

Therefore, the dual problem is given by:

\[
\max \mathcal{P}_R^T \alpha, \quad \text{subject to } \alpha \in \bigcap_{l=1}^{J} K_{\text{PolM,l}},
\]

(57)

where we have used the identity that \( K_{\text{VDM,l}} = \bigcap_{l=1}^{J} K_{\text{PolM,l}} \) given by Theorem 3. Note that (58) can be cast as SDP following from (35).

Example 1: Suppose that the moment sequence \( t_j, |j| \leq N-1 \) is generated from its 3-atomic representing measure

\[
\mu_1 = 0.7\delta_{0,1} + 2\delta_{0,25} + \delta_{0.7}.
\]

1) In the case of \( N \geq 4 \), we can form the Toeplitz matrix \( T \) using \( t \), having that rank \( (T) = 3 < N \). By Theorem 1, \( \mu_1 \) is the unique measure for \( T \).

2) Suppose that \( N = 3 \) and \( K = [0.2, 0.3] \cup [0.6, 0.8] \). We solve (54) using SDPT3 [29] in Matlab and successfully find a solution. This means that a \( K \)-representing measure exists for \( t \) by Corollary 1. Applying FS Vandermonde decomposition to the solution, a 6-atomic \( K \)-representing measure is given by:

\[
\mu_2 = 1.9614\delta_{0.2} + 0.1296\delta_{0.2290} + 0.4456\delta_{0.2891} + 0.2437\delta_{0.6} + 0.3637\delta_{0.6467} + 0.5561\delta_{0.7962}.
\]

(60)

3) With the same settings as in 2), instead of solving (54), we find the one maximizing \( \text{tr}(T(t^*_1)) \) among the solutions satisfying (51)-(53). The obtained solution \( \{t^*_1, t^*_2\} \) satisfies that rank \( (T(t^*_1)) = \text{rank}(T(t^*_2)) = 2 < N \), resulting in the following 4-atomic representing measure:

\[
\mu_3 = 2.0837\delta_{0.2} + 0.4726\delta_{0.3} + 0.6218\delta_{0.6382} + 0.5219\delta_{0.8}.
\]

(61)

4) Suppose that \( N = 3 \) and \( K = [0.2, 0.3] \cup [0.6, 0.75] \). Then (54) is infeasible. This means that no \( K \)-representing measure for \( t \) exists by Corollary 1.

VI. APPLICATION IN LINE SPECTRAL ESTIMATION

A. Problem Statement

Line spectral estimation can be found in wide applications such as communications, radar, sonar, and so on [6]. In particular, we have the following data model:

\[
y^\ast = \sum_{k=1}^{r} a(f_k) s_k = A(f) s,
\]

(62)

where \( y^\ast \in \mathbb{C}^N \) is a uniformly sampled signal (at the Nyquist sampling rate), \( f_k \in \mathbb{T} \) and \( s_k \in \mathbb{C} \) are the normalized frequency and the complex amplitude of the \( k \)-th sinusoid respectively, and \( r \) is the number of sinusoids. To estimate the frequencies, we are given a part of entries of \( y^\ast \) that form the subvector \( y^\ast_{\Omega} \in \mathbb{C}^M \), where \( \Omega \) denotes the set of sampling indexes and is of cardinality \( M < N \). This frequency estimation problem is referred to as off-grid/continuous compressed sensing in [8] in the sense that we have compressive data as in the pioneering work of compressed sensing [29], but differently, the frequencies can take any value in \( \mathbb{T} \) as opposed to the discrete setting in [29].

In this section, we consider the case when the frequencies are known \textit{a priori} to lie in an interval \( \mathbb{I} \subset \mathbb{T} \). Inspired by the recent atomic norm techniques [7]-[9], the paper [19] proposed a FS atomic norm approach (or constrained atomic norm in the language of [19]) that was shown theoretically and empirically to achieve better performance than the standard atomic norm by exploiting the prior knowledge. In particular, define the (FS) set of atoms

\[
A(\mathbb{I}) := \{a(f_k, \phi_k) = a(f) : f \in \mathbb{I}, |\phi| = 1 \}.
\]

(63)

The FS atomic norm is the norm induced by \( A(\mathbb{I}) \):

\[
\|y\|_{A(\mathbb{I})} := \inf_{c_k > 0, a_k \in A(\mathbb{I})} \left\{ \sum_k c_k : y = \sum_k a_k a_k \right\}.
\]

(64)

The following FS atomic norm minimization (FS-ANM) problem was proposed in [19]:

\[
\min_{y} \|y\|_{A(\mathbb{I})}, \text{subject to } y_{\Omega} = y^\ast_{\Omega}.
\]

(65)

This means that the candidate \( y^\ast \) with the minimum FS atomic norm is used as the signal estimate, and the frequencies composing \( y^\ast \) form the frequency estimates. Note that (63)-(65) degenerate to the existing standard forms in the case of \( \mathbb{I} = \mathbb{T} \).

Since the FS atomic norm defined in (63) is inherently semi-infinite programming (SIP), a finite-dimensional formulation of it is required to practically solve (65), which is dealt with in the ensuing section by applying the FS Vandermonde decomposition.

B. SDP Formulation of FS Atomic Norm

By applying the FS Vandermonde decomposition, the FS atomic norm is cast as SDP in the following theorem.

Theorem 5: It holds that

\[
\|y\|_{A(\mathbb{I})} = \min_{x, t} \frac{1}{2} x + \frac{1}{2} t_0, \quad \text{subject to } \begin{bmatrix} x & y^H \end{bmatrix} \geq 0 \text{ and } T_g \geq 0,
\]

(66)

where \( g \) is as defined previously.

Proof: Let \( F^\ast \) be the optimal objective value of (66). We need to show that \( \|y\|_{A(\mathbb{I})} = F^\ast \).
Moreover, it holds that $x^* \leq \|y\|_{A(I)}$. To do so, let $y = \sum_k c_k a(f_k, \phi_k)$ be a FS atomic decomposition of $y$ on $I$. Then let $t$ be such that $T(t) = \sum_k c_k a(f_k) a^H(f_k)$ and $x = \sum_k c_k$. By Theorem 3 we have that $T_y \geq 0$. Moreover, it holds that

$$\begin{bmatrix} x^* & y^H \\ y & T \end{bmatrix} = \sum_k c_k \begin{bmatrix} \overline{\phi_k} \\ \overline{a(f_k)} \end{bmatrix} \begin{bmatrix} \overline{\phi_k} \\ a(f_k) \end{bmatrix}^H \geq 0. \quad (67)$$

Therefore, $x$ and $t$ constructed above form a feasible solution to the problem in (66), at which the objective function equals

$$\frac{1}{2} x^* + \frac{1}{2} t_0 = \sum_k c_k. \quad (68)$$

It follows that $F^* \leq \sum_k c_k$. Since the inequality holds for any FS atomic decomposition of $y$ on $I$, we have that $F^* \leq \|y\|_{A(I)}$ by the definition of $\|y\|_{A(I)}$.

On the other hand, suppose that $(x^*, t^*)$ is an optimal solution to the problem in (66). By the fact that $T(t^*)$ is nonnegative on $I$ and $T_y(t^*) \geq 0$ and applying Theorem 3 we have that $T(t^*)$ has a FS Vandermonde decomposition on $I$ as in (64), with $(r, p_k, f_k)$ denoted by $(r^*, p_k^*, f_k^*)$. By the fact that $[x^* \ y^H] \geq 0$, we have that $y$ lies in the range space of $T(t^*)$ and thus has the following FS atomic decomposition:

$$y = \sum_{k=1}^{r^*} c_k^* a(f_k^*, \phi_k^*), \quad f_k^* \in I. \quad (69)$$

Moreover, it holds that

$$x^* \geq y^H [T(t^*)] y = \sum_{k=1}^{r^*} \frac{c_k^*}{p_k^*} \quad (70)$$

$$t_0^* = \sum_{k=1}^{r^*} p_k^*. \quad (71)$$

It therefore follows that

$$F^* = \frac{1}{2} x^* + \frac{1}{2} t_0^*$$

$$\geq \frac{1}{2} \sum_k c_k^2 [p_k^*] + \frac{1}{2} \sum_k p_k^*$$

$$\geq \sum_k c_k$$

$$\geq \|y\|_{A(I)}. \quad (72)$$

Combining (72) and the inequality that $F^* \leq \|y\|_{A(I)}$ as shown previously, we conclude that $F^* = \|y\|_{A(I)}$ and complete the proof. At last, it is worth noting that by (72) it must hold that $p_k^* = c_k$ and $\|y\|_{A(I)} = \sum_k c_k^2$. Therefore, the FS atomic decomposition in (69) must achieve the FS atomic norm.

It immediately follows from Theorem 3 that (65) can be written as the following SDP:

$$\begin{array}{ll}
\min \frac{1}{2} x^* + \frac{1}{2} t_0, \\
\text{subject to } [x^* \ y^H] \geq 0, T_y \geq 0 \text{ and } y = y_\Omega.
\end{array} \quad (73)$$

Note that (73) can be solved using off-the-shelf SDP solvers such as SDPT3. Given its solution, the frequencies can be retrieved from the FS Vandermonde decomposition of $T$.

Before proceeding to the next subsection, we note that (65) was solved by studying its dual in (19). In particular, the dual of (65) is given by:

$$\max_{z, Q_0, Q_1} \Re(y_\Omega^* z_\Omega), \text{ subject to } \|z\|_{A(I)}^* \leq 1 \text{ and } z_\Omega^c = 0, \quad (74)$$

where $\Omega^c$ denotes the complement of $\Omega$ and $\|z\|_{A(I)}^*$ is the dual FS atomic norm. By the fact that

$$\|z\|_{A(I)}^* = \sup_{a \in A(I)} \Re(a^H z) = \sup_{f \in I} \Re(a^H (f) z), \quad (75)$$

the constraint that $\|z\|_{A(I)}^* \leq 1$ can be cast as the following:

$$|a^H (f) z| \leq 1 \text{ for any } f \in I, \quad (76)$$

where $q(f) := a^H (f) z \quad (77)$

is referred to as the dual polynomial (7), (19). It follows that $1 - |q(f)|^2$ is a Hermitian trigonometric polynomial nonnegative on $I$ and, by the PRL, admits a Gram matrix parametrization as in (35). With some further derivations that we will omit, it can be shown that (76) holds if and only if the unit polynomial (the right hand side of the inequality in (76)) has the following Gram matrix parametrization:

$$\text{tr}(\Theta_j Q_0) + \text{tr}(\Theta_g Q_1) = \begin{cases} 1, & \text{if } j = 0, \\
0, & \text{otherwise},
\end{cases} \quad (78)$$

and moreover, $Q_0$ and $Q_1$ satisfy that

$$\begin{bmatrix} 1 \\ z^H Q_0 \\ z \end{bmatrix} \geq 0 \text{ and } Q_1 \geq 0. \quad (79)$$

In fact, this is nothing but the bounded real lemma (BRL) for trigonometric polynomials (21, 23) that is a more precise result of PRL when applying to (76). Finally, (74) is cast as the following SDP:

$$\max_{z, Q_0, Q_1} \Re(y_\Omega^* z_\Omega), \text{ subject to } (78), (79) \text{ and } z_\Omega^c = 0. \quad (80)$$

Without surprise, it follows from a standard Lagrangian analysis that (80) is the dual of (73) (note that the analysis uses (33) and (34) and will be left to interested readers). Since strong duality holds (26), the solution to (80) can be obtained for free when solving (73) using a primal-dual algorithm, and vise versa.

In summary, the FS Vandermonde decomposition can be applied to provide a primal SDP formulation of (65), while the trigonometric polynomial based technique in (19) provides a dual SDP formulation. Moreover, the FS Vandermonde decomposition also provides a new method for frequency retrieval. We empirically find that the new method results in higher numerical stability, as compared to the root-finding method in (7), (19), since in the former we can always determine the number of frequencies first by computing rank ($T$), which can effectively reduce the problem dimension and improve stability. In contrast to this, the root-finding method requires to solve all, up to $2N-2$, roots of the polynomial $1-|q(f)|^2$. 
amplitude
R. Amplitude
0
5
10
15
20
25
30
Ground truth
Recovered
Dual polynomial
0.2 0.22 0.24 0.26 0.28 0.3
0
0.5
1
1.5
2
(a) ANM
(b) FS-ANM
Fig. 1. Line spectral estimation results of (a) ANM and (b) FS-ANM.

among which appropriate ones are selected to produce the frequencies.

Example 2: Consider a line spectrum composed of \( K = 3 \) frequencies \( f = [0.22, 0.23, 0.28]^T \) as shown in Fig. 1. To estimate/recover the spectrum, \( M = 16 \) randomly located noiseless samples are acquired among \( N = 64 \) uniform samples. The standard ANM and the FS-ANM methods are implemented using SDPT3 to estimate the line spectrum. In FS-ANM, the prior knowledge that the frequencies lie in \( I = [0.2, 0.3] \) is used. The estimation results are presented in Fig. 1.

It can be seen that FS-ANM exactly recovers the spectrum but ANM does not. For both ANM and FS-ANM, the recovered frequencies retrieved using the Vandermonde decomposition match the locations at which the dual polynomials have unit magnitude. Finally, note that for FS-ANM the frequencies computed using the FS Vandermonde decomposition have recovery errors on the order of \( 10^{-10} \) while those computed using the root-finding method have errors on the order of \( 10^{-6} \).

C. Extension to FS Atomic \( \ell_0 \) Norm

In this subsection, we study the FS atomic \( \ell_0 \) norm defined by:

\[
\|y\|_{\mathcal{A}(I), 0} := \inf_{\mathcal{K} : y = \sum_{k=1}^{K} c_k a_k} \left\{ \| K : y = \sum_{k=1}^{K} a_k f_k s_k \right\}.
\]

\( \|y\|_{\mathcal{A}(I), 0} \) is of interest since it exploits sparsity to the greatest extent possible, while \( \|y\|_{\mathcal{A}(I)} \) is in fact its convex relaxation. It has been vastly demonstrated in the literature on compressed sensing that improved performance can usually be obtained by solving (or approximately solving) \( \ell_0 \) norm based problems (see, e.g., [13, 20, 51]). To possibly compute \( \|y\|_{\mathcal{A}(I), 0} \), a finite-dimensional formulation of it is provided in the following theorem by applying the FS Vandermonde decomposition.

**Theorem 6:** It holds that

\[
\|y\|_{\mathcal{A}(I), 0} = \min_{x, t} \| x \|_T, \quad \text{subject to } \left[ x \ y^T \right] \geq 0 \text{ and } T_g \geq 0, \tag{82}
\]

where \( g \) is as defined previously.

**Proof:** The proof is similar to that of Theorem 5. At the first step, by applying the FS Vandermonde decomposition, we can construct a feasible solution, as in the proof of Theorem 5 to the optimization problem in (82), which concludes that \( \|y\|_{\mathcal{A}(I), 0} \leq r^* \), where \( r^* \) denotes the optimal objective value of (82). At the second step, for any optimal solution that achieves the optimal value \( r^* \), we can similarly obtain an \( r^* \)-atomic FS decomposition of \( y \), which results in that \( \|y\|_{\mathcal{A}(I), 0} \geq r^* \). So we complete the proof.

It follows from Theorem 6 that \( \|y\|_{\mathcal{A}(I), 0} \) is cast as a rank minimization problem, while solving (or approximately solving) the resulting optimization problem is beyond the scope of this paper. It is worth noting that, since \( \|y\|_{\mathcal{A}(I), 0} \) is nonconvex, a trigonometric polynomial based technique, as in [19] for \( \|y\|_{\mathcal{A}(I)} \), cannot be applied in this case to provide a finite-dimensional formulation.

VII. CONCLUSION

In this paper, the FS Vandermonde decomposition of Toeplitz matrices on a given interval was studied. The new result generalizes the classical Vandermonde decomposition result. It was shown by duality to be connected to the theory on trigonometric polynomials. It was also applied to provide a solution to the classical truncated trigonometric \( K \)-moment problem and a primal SDP formulation of the recent FS atomic norm for line spectral estimation with prior knowledge.

ACKNOWLEDGMENT

Z. Yang would like to thank Prof. Petre Stoica of Uppsala University, Sweden for fruitful discussions on an earlier version of this paper.
REFERENCES

[1] C. Carathéodory and L. Fejér, “Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landau’schen Satz,” Rendiconti del Circolo Matematico di Palermo (1884-1940), vol. 32, no. 1, pp. 218–239, 1911.

[2] N. I. Akhiezer, *The classical moment problem: and some related questions in analysis*. Oliver & Boyd Edinburgh, 1965, vol. 5.

[3] U. Grenander and G. Szego, *Toeplitz forms and their applications*. Univ of California Press, 1958.

[4] T. T. Georgiou, “The Carathéodory–Fejér–Pisarenko decomposition and its multivariable counterpart,” IEEE Transactions on Automatic Control, vol. 52, no. 2, pp. 212–228, 2007.

[5] V. F. Pisarenko, “The retrieval of harmonics from a covariance function,” Geophysical Journal International, vol. 33, no. 3, pp. 347–366, 1973.

[6] K. V. Mishra, M. Cho, A. Krüger, and W. Xu, “Spectral super-resolution,” Communications on Pure and Applied Mathematics, vol. 67, no. 6, pp. 906–956, 2014.

[7] 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.

[8] 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.

[9] T. T. Georgiou, “The Carathéodory–Fejer–Pisarenko decomposition and its multivariable counterpart,” IEEE Transactions on Automatic Control, vol. 52, no. 2, pp. 489–509, 2006.

[10] Z. Yang, L. Xie, and P. Stoica, “Vandermonde decomposition of multilevel Toeplitz matrices with application to multidimensional super-resolution,” IEEE Transactions on Information Theory, DOI: 10.1109/TIT.2016.2553041, 2016.

[11] K. Schmidt, “The K-moment problem for compact semi-algebraic sets,” Mathematische Annalen, vol. 289, no. 1, pp. 203–206, 1991.

[12] R. E. Curto and L. A. Fialkow, “Recursiveness, positivity, and truncated Toeplitz forms and their applications,” Transactions of the American mathematical society, vol. 352, no. 6, pp. 2825–2855, 2000.

[13] J.-B. Lasserre, *Moments, positive polynomials and their applications*. SIAM, Philadelphia, PA, 2009.

[14] S. Beygi and U. Mitra, “Multi-scale multi-lag channel estimation using low rank approximation for OFDM,” IEEE Transactions on Signal Processing, vol. 63, no. 18, pp. 4744–4755, 2015.

[15] R. Doviak and D. S. Zrnić, *Doppler radar and weather observations*. Courier Corporation, 1993.

[16] K. V. Mishra, M. Cho, A. Krüger, and W. Xu, “Spectral super-resolution with prior knowledge,” IEEE Transactions on Signal Processing, vol. 63, no. 20, pp. 5342–5357, 2015.

[17] Z. Yang and L. Xie, “A weighted atomic norm approach to spectral super-resolution with probabilistic priors,” in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 4598–4602.

[18] B. Dumitrescu, *Positive trigonometric polynomials and signal processing applications*. Springer, 2007.

[19] L. Gurvits and H. Barnum, “Largest separable balls around the maximally mixed bipartite quantum state,” Physical Review A, vol. 66, no. 6, p. 062311, 2002.

[20] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 2012.

[21] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, “Linear matrix inequality formulation of spectral mask constraints with applications to FIR filter design,” IEEE Transactions on Signal Processing, vol. 50, no. 11, pp. 2702–2715, 2002.

[22] B. Alkire and L. Vandenberghe, “Convex optimization problems involving finite autocorrelation sequences,” Mathematical Programming, vol. 93, no. 3, pp. 331–359, 2002.

[23] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, London, 2004.

[24] E. Candès and L. A. Fialkow, “Recursiveness, positivity, and truncated moment problems,” Houston Journal of Mathematics, vol. 17, no. 4, pp. 603–635, 1991.