Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Approximation Theory 139 (2006) 327–345
Journal of Approximation Theory
www.elsevier.com/locate/jat

Factorization of multivariate positive Laurent polynomials

Jeffrey S. Geronimo[a,*,1], Ming-Jun Lai[b,2]

[a]School of Mathematics, The Georgia Institute of Technology, Atlanta, GA 30332, USA
[b]Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

Received 16 February 2005; accepted 21 September 2005

Communicated by Zlatos
Available online 15 November 2005

Dedicated to Barry Simon on the occasion of his 60th birthday

---

Abstract

Recently Dritschel proved that any positive multivariate Laurent polynomial can be factorized into a sum of square magnitudes of polynomials. We first give another proof of the Dritschel theorem. Our proof is based on the univariate matrix Fejér–Riesz theorem. Then we discuss a computational method to find approximates of polynomial matrix factorization. Some numerical examples will be shown. Finally we discuss how to compute nonnegative Laurent polynomial factorizations in the multivariate setting.

---

1. Introduction

We are interested in computing factorizations of nonnegative Laurent polynomials into sum of squares of polynomials. That is, let

$$P(z) = \sum_{k=-n}^{n} p_k z^k$$

---

[*] Corresponding author.

E-mail addresses: geronimo@math.gatech.edu (J.S. Geronimo), mjlai@math.uga.edu (M.-J. Lai).

[1] This author is partly supported by an NSF grant.

[2] This author is partly supported by the National Science Foundation under grant EAR-0327577 and the School of Mathematics, at the Georgia Institute of Technology where he visited during the fall, 2004.

doi:10.1016/j.jat.2005.09.010

be a Laurent polynomial, where $z = e^{i\theta}$. Suppose that $P(z) \geqslant 0$ for $|z| = 1$. One would ask if there exists a polynomial $Q(z) = \sum_{k=0}^{n} q_k z^k$ such that

$$P(z) = Q(z)^* Q(z), \tag{1}$$

where $Q(z)^*$ denotes the complex conjugate of $Q(z)$. This is the well-known Fejér–Riesz factorization problem and it was resolved by Fejér [7] and by Riesz [21]. A natural question is whether the results of Fejér and Riesz can be extended to the multivariate setting. More generally, given a nonnegative multivariate trigonometric polynomial $P(z) := P(z_1, z_2, \ldots, z_d)$ with coordinate degrees $\leqslant n$, does there exist a finite number of polynomials $Q_k(z)$ such that

$$P(z) = \sum_k Q_k^*(z) Q_k(z), \tag{2}$$

i.e., can $P(z)$ be written as a sum of square magnitudes (sosm) of polynomials. There is a vast amount of literature related to the study of this problem and the results relevant to this paper may be summarized as follows:

1. When $P(z)$ is nonnegative on the multi-torus $|z_1| = |z_2| = \cdots = |z_d| = 1$ and the coordinate degrees of Q_k are less than or equal to n, then the answer to the question is negative. (See [4,24].)
2. When $P(z)$ is strictly positive on the multi-torus and the coordinate degrees of Q_k are not specified, Dritschel has shown that the answer to the question is positive [6]. However, the nonnegative case remains unresolved.
3. In the bivariate setting, Geronimo and Woerdeman gave a necessary and sufficient condition in order for $P(z) = |Q(z)|^2$, where $Q(z)$ is a stable polynomial, i.e., $Q(z) \neq 0$ inside and on the bi-torus [8].
4. In the bivariate setting, there exist rational Laurent polynomials $Q_k(z)$ such that (2) holds. Furthermore, Q_k can be so chosen that the determinants of Q_k are only one variable Laurent polynomials (cf. [1]).
5. In [17], an algorithm was proposed to find polynomials P_k such that $P = \sum_k |P_k|^2$. The algorithm uses semi-definite programming.

Although the mathematical problem appears to be theoretical, it has many applications in engineering, e.g., the design of autoregressive filters, construction of orthonormal wavelets (cf. [5]), construction of tight wavelet framelets (cf. [16]), spectral estimation in control theory (cf. [25]) and many other engineering applications mentioned in [17]. Thus, how to compute such factorization polynomials $Q_1, Q_2, \ldots,$ is interesting and useful for applications.

In this paper, we iteratively reduce the problem of factorization of multivariate positive Laurent polynomials to a problem of factorization of univariate positive definite polynomial matrices and thus present a new elementary proof of Dritschel's Theorem. The proof suggests a computational method (a Bauer type method [2,3]) for computing the above factorization. The Bauer method has been studied and generalized to the multivariate and operator settings by many researchers, e.g., [26,10,19,25]. It was argued by Bauer [3] that his method converges exponentially fast. See [15,9] for different proofs of the exponential convergence of their Bauer type methods. The Bauer method was extended to the multivariate case in [11,20]. For the factorization of univariate positive definite polynomial matrix, a linear convergence of the Bauer type method was proved in [26]. Later van der Mee et al. [18] used Banach algebra techniques to show that the method converges exponentially fast for real matrices. For the convenience of the reader we present an elementary proof based on an extension of the method in [15].

The paper is organized as follows. In Section 2, we first give a different proof of Dritschel's Theorem. As mentioned above the key to the proof is to iteratively reduce the factorization of a multivariate strictly positive Laurent polynomial to a problem of factorizing a positive definite univariate matrix of Laurent polynomials. In Section 3, we explain a Bauer type method to compute the factorization of positive definite Laurent polynomial matrices. The convergence of the method is shown to be exponentially fast. Then in Section 4, some numerical examples are computed following the procedure in Sections 2 and 3. In Section 5, the factorization of nonnegative Laurent polynomials is considered and the paper is concluded with some remarks in Section 6.

2. Dritschel's theorem

We begin with reviewing the concept of the symbol of a bi-infinite Toeplitz matrix and discussing its properties [12, p. 16]. For a given univariate Laurent polynomial $P(z) = \sum_{k=-n}^{n} p_k z^k$, we may view $P(z)$ as the symbol of a bi-infinite Toeplitz matrix $\mathcal{P} := (p_{i-j})_{i,j\in\mathbf{Z}}$. Indeed, for any absolutely summable sequence $\mathbf{x} = (x_i)_{i\in\mathbf{Z}}$, i.e., $\sum_{i\in\mathbf{Z}} |x_i| < \infty$, let $F(\mathbf{x}) = \sum_{j\in\mathbf{Z}} x_j z^j$ be the discrete Fourier transform (or z-transform) of $\mathbf{x}$. Let $\mathbf{y} = \mathcal{P}\mathbf{x}$, then it is easy to see that

$$F(\mathbf{y}) = P(z)F(\mathbf{x}).$$

If the matrix $\mathcal{P}$ has a factorization $\mathcal{Q}$ which is a banded upper triangular Toeplitz matrix such that

$$\mathcal{P} = \mathcal{Q}^\dagger \mathcal{Q},$$

the discrete Fourier transform of $\mathbf{y} = \mathcal{Q}^\dagger\mathcal{Q}\mathbf{x}$ is $F(\mathbf{y}) = Q(z)^* Q(z) F(\mathbf{x})$, where $\mathcal{Q}^\dagger$ denotes the complex conjugate transpose of matrix $\mathcal{Q}$ and $\mathcal{Q}(z)^*$ the complex conjugate of the Laurent polynomial $\mathcal{Q}(z)$. Thus, finding $P(z) = Q(z)^* Q(z)$ is equivalent to finding a banded upper triangular Toeplitz matrix $\mathcal{Q}$ such that $\mathcal{P} = \mathcal{Q}^\dagger\mathcal{Q}$.

It is easy to show that if $P(z)\geqslant 0$ for all $|z| = 1$, then $\mathcal{P}$ is Hermitian and nonnegative definite. Clearly, $\mathcal{P}$ is Hermitian since $P(z)$ is real. Furthermore, for any absolutely summable sequence $\mathbf{x}$, we need to show that $\mathbf{x}^\dagger\mathcal{P}\mathbf{x}\geqslant 0$. Again writing $\mathbf{y} = \mathcal{P}\mathbf{x}$, we know that

$$\mathbf{x}^\dagger\mathbf{y} = \frac{1}{2\pi}\int_0^{2\pi} \overline{F(\mathbf{x})} F(\mathbf{y})\,\mathrm{d}\theta,$$

where $z = e^{i\theta}$ and it follows that

$$\mathbf{x}^\dagger\mathcal{P}\mathbf{x} = \frac{1}{2\pi}\int_0^{2\pi} |F(\mathbf{x})|^2 P(z)\,\mathrm{d}\theta \geqslant 0,$$

for any nonzero sequence $\mathbf{x}$. In particular, for

$$\mathbf{x} = (\ldots, 0, x_{-N}, \ldots, x_0, \ldots, x_N, 0, \ldots)^T,$$

the left-hand side in the above inequality gives $\mathbf{x}^\dagger P_N \mathbf{x}$, where P_N is a central section of $\mathcal{P}$. The above argument shows that P_N is nonnegative definite.

In the following we will assume that $P(z)$ is strictly positive, in the sense that there exists a positive number $\varepsilon > 0$ such that $P(z)\geqslant\varepsilon$. When $P(z)$ is a matrix, we mean that $P(z)\geqslant\varepsilon I$, where I is the identity matrix of the same size as that of $P(z)$. When $P(z)$ is strictly positive, we have

$$\mathbf{x}^\dagger \mathcal{P}\mathbf{x} = \frac{1}{2\pi}\int_0^{2\pi} |F(\mathbf{x})|^2 P(z)\,\mathrm{d}\theta \geqslant \varepsilon\|\mathbf{x}\|^2.$$

It follows that if $P(z)\geqslant\varepsilon > 0$, then $P_N\geqslant\varepsilon > 0$.

We now consider the factorization of multivariate Laurent polynomials. Let us begin with a bivariate Laurent polynomial $P(z_1, z_2)$. That is, let

$$P(z_1, z_2) = \sum_{j=-n}^{n}\sum_{k=-n}^{n} p_{jk} z_1^j z_2^k \geqslant 0$$

be a Laurent polynomial of coordinate degrees $\leqslant n$. We would like to find a finite number of polynomials Q_k such that

$$P(z_1, z_2) = \sum_k |Q_k(z_1, z_2)|^2.$$

Denote by $\mathbf{z_1} = [1, z_1, z_1^2, \ldots, z_1^n]^T$ and write

$$P(z_1, z_2) = \mathbf{z_1}^\dagger \widetilde{P}(z_2)\mathbf{z_1}$$

for a Hermitian matrix $\widetilde{P}(z_2) = \sum_{k=-n}^{n} \tilde{p}_k z_2^k$, where each p_k is an $(n + 1) \times (n + 1)$ Toeplitz matrix. With a slight modification of an observation of [17, Theorem 2.1], we note that there are many ways to write $\widetilde{P}(z_2)$. If there is one $\widetilde{P}(z_2)$ which is nonnegative definite then we can use the matrix Fejér–Riesz factorization theorem (cf. e.g., in [14,22,23,17], see also Section 3) to find $\widetilde{Q}(z_2)$ such that

$$\widetilde{P}(z_2) = \widetilde{Q}^\dagger(z_2)\widetilde{Q}(z_2).$$

That is, we have

$$P(z_1, z_2) = (\widetilde{Q}(z_2)\mathbf{z_1})^\dagger \widetilde{Q}(z_2)\mathbf{z_1},$$

which is clearly a sum of magnitude squares of polynomials.

The above discussion can be generalized to the multivariate setting and using an observation of [6] to the case that the size of $\widetilde{P}(z_2)$ is larger than $(n+1) \times (n+1)$. For simplicity, let us consider a trivariate Laurent polynomial $P(z_1, z_2, z_3)$ in $z_1 = e^{i\theta_1}, z_2 = e^{i\theta_2}, z_3 = e^{i\theta_3}$ of coordinate degrees $\leqslant n$. We first write $P(z_1, z_2, z_3)$ in a matrix format

$$P(z_1, z_2, z_3) = \sum_{k=-n}^{n} p_k(z_2, z_3) z_1^k = \mathbf{z_1}^\dagger \widehat{P}(z_2, z_3)\mathbf{z_1},$$

with

$$\mathbf{z_1} = [1, z_1, \ldots, z_1^{m_1}]^T \tag{3}$$

and $m_1 \geqslant n$. There are many ways to write $\widehat{P}(z_2, z_3)$. To capture this define the set of matrices

$$\mathcal{F} = \left\{ (p_{i,j}(z_2, z_3))\ 0\leqslant i, j\leqslant m_1 : \sum_{\substack{i-j=k\\ |k|\leqslant m_1}} p_{i,j}(z_2) = p_k(z_2), \right\}.$$

Note that the matrices in $\mathcal{F}$ are banded since $p_k = 0,\ |k| > n$. We look for a matrix $\widehat{P}(z_2, z_3)$ in $\mathcal{F}$ that is positive definite for $|z_2| = 1 = |z_3|$. The polynomial matrix $\widehat{P}(z_2, z_3)$ can be written as

$$\widehat{P}(z_2, z_3) = \sum_{k=-n}^{n} \tilde{P}_k(z_3) z_2^k,$$

where each $\tilde{P}_k(z_3)$ is an $(m_1 + 1) \times (m_1 + 1)$ Toeplitz matrix. Thus we can write

$$\widehat{P}(z_2, z_3) = \mathbf{z_2}^{\dagger} \bar{P}(z_3) \mathbf{z_2},$$

where

$$\mathbf{z_2} = [I_{m_1}, z_2 I_{m_1}, \ldots, z_2^{m_2} I_{m_1}]^T,$$

with I_{m_1} being the $(m_1 + 1) \times (m_1 + 1)$ identity matrix and $m_2 \geqslant n$. The polynomial $\bar{P}(z_3)$ is a matrix polynomial of size $(m_1+1)(m_2+1) \times (m_1+1)(m_2+1)$. If it is nonnegative definite we can factor it into a polynomial matrix $Q(z_3)$, i.e., $\bar{P}(z_3) = Q(z_3)Q(z_3)^{\dagger}$ by the matrix Fejér–Riesz theorem and we have

$$P(z_1, z_2, z_3) = (Q(z_3)\mathbf{z_2 z_1})^{\dagger}(Q(z_3)\mathbf{z_2 z_1}),$$

which is a sum of square magnitudes of polynomials in z_1, z_2, z_3.

Our task then is to produce a positive definite polynomial matrix for any given positive multivariate Laurent polynomial. We resume our discussion on the two variable case again and rewrite $P(z_1, z_2)$ as follows:

$$P(z_1, z_2) = \sum_{k=-n_1}^{n_1} p_k(z_2) z_1^k = \mathbf{z_{m_1}}^{\dagger} P_{m_1}(z_2) \mathbf{z_{m_1}},$$

where $m_1 \geqslant n_1$, $\mathbf{z_{m_1}} = [1, z_1, z_1^2, \ldots, z_1^{m_1}]^T$, and

$$P_{m_1}(z_2) = [p_{jk}(z_2)]_{0 \leqslant j,k \leqslant m_1}$$

with polynomial entries $p_{j,k}(z_2)$ given by

$$p_{jk}(z_2) = \frac{1}{m_1 + 1 - |j - k|} p_{k-j}(z_2), \quad \forall j, k = 0, \ldots, m_1.$$

Note that $p_{jk}(z_2) = 0$ for $|j - k| > n_1$. Under this decomposition we can show that for some m_1 large enough, the matrix $P_1(z_2)$ will be positive definite when $P(z_1, z_2)$ is positive definite. To see this note that $P(z_1, z_2)$ is the symbol of the following bi-infinite Toeplitz matrix:

$$\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & p_0(z_2) & p_{-1}(z_2) & \cdots & p_{-n}(z_2) & 0 & \cdots \\ \ddots & p_1(z_2) & p_0(z_2) & \ddots & \ddots & \ddots & \ddots \\ \ddots & p_2(z_2) & p_1(z_2) & \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & p_n(z_2) & p_{n-1}(z_2) & \cdots & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{4}$$

The positivity of $P(z_1, z_2)$ implies that any central section of this matrix, i.e., any square block with the diagonal consistent with the main diagonal

$$\mathrm{diag}(\ldots, p_0(z_2), p_0(z_2), p_0(w_2), \ldots)$$

is positive as explained at the beginning of this section. Typically, we have

$$p_0(z_2) > 0, \quad \begin{bmatrix} p_0(z_2) & p_{-1}(z_2) \\ p_1(z_2) & p_0(z_2) \end{bmatrix} > 0, \quad \begin{bmatrix} p_0(z_2) & p_{-1}(z_2) & p_{-2}(z_2) \\ p_1(z_2) & p_0(z_2) & p_{-1}(z_2) \\ p_2(z_2) & p_1(z_2) & p_0(z_2) \end{bmatrix} > 0, \ldots .$$

For convenience, we denote by $\mathcal{P}_2$ and $\mathcal{P}_3$ to be the 2×2 and 3×3 matrices above and in general $\mathcal{P}_k$ to denote the $k \times k$ central block matrix from the bi-infinite Toeplitz matrix (4) above.

Now look at the matrix $P_{m_1}(z_2)$ given by

$$\begin{bmatrix} \frac{1}{m_1+1} p_0(z_2) & \frac{1}{m_1} p_{-1}(z_2) & \cdots & \frac{1}{m_1+1-n_1} p_{-n_1}(z_2) & 0 & \cdots \\ \frac{1}{m_1} p_1(z_2) & \frac{1}{m_1+1} p_0(z_2) & \frac{1}{m_1} p_{-1}(z_2) & \ddots & \ddots & \ddots \\ \frac{1}{m_1-1} p_2(z_2) & \frac{1}{m_1} p_1(z_2) & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \frac{1}{m_1+1-n_1} p_{n_1}(z_2) & \ddots & \cdots & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \frac{1}{m_1+1} p_0(z_2) \end{bmatrix}.$$

With $\mathbf{x} = [x_0, x_1, \ldots, x_{m_1}]^T$, we need to prove that $\mathbf{x}^* P_{m_1}(z_2)\mathbf{x} > 0$. First write

$$\mathbf{x}^\dagger P_{m_1} \mathbf{x} = \frac{1}{m_1+1} \mathbf{x}^\dagger \mathcal{P}_{m_1} \mathbf{x} + \frac{1}{m_1+1} \mathbf{x}^\dagger R_{m_1} \mathbf{x}$$

with a remainder matrix R_{m_1}. The ℓ_2 norm of R_{m_1} can be estimated directly to give

$$\|R_{m_1}\|_2 \leqslant \frac{2n_1(n_1+1)C_1}{\sqrt{3}(m_1-n_1)},$$

where we have used the fact that R_{m_1} is a banded matrix and

$$C_1 = \sup_{\substack{i=1,\ldots,n_1, \\ |z_2|=1}} |p_i(z_2)|. \tag{5}$$

If $P(z_1, z_2) \geqslant \varepsilon$ then $\mathbf{x}^\dagger \mathcal{P}_{m_1} \mathbf{x} \geqslant \varepsilon \|\mathbf{x}\|_2$, so that if $\dfrac{2n_1(n_1+1)C}{\sqrt{3}(m_1-n_1)} < \varepsilon$, then $\mathbf{x}^\dagger P_{m_1}(z_2)\mathbf{x} > 0$. Thus an application of the matrix Fejér–Riesz Theorem yields:

Theorem 2.1. *Let* $P(z_1, z_2) = \sum_{k=-n_1}^{n_1} p_k(z_2) z_1^k \geqslant \varepsilon > 0$ *be strictly positive on bi-torus* $|z_1| = 1 = |z_2|$. *Then* $P(z_1, z_2)$ *can be factored into a sum of square magnitudes of polynomials in* z_1 *and* z_2. *The total number of terms in the sum is less than or equal to* $m_1 + 1$ *with* m_1 *being an integer such that*

$$\frac{2n_1(n_1+1)C_1}{\sqrt{3}(m_1-n_1)} < \varepsilon,$$

where C_1 is the positive constant given in (5). *The degrees of each of the polynomials is bounded by m_1 in z_1 and n_2 in z_2.*

We remark that when $P(z_1, z_2)$ has different coordinate degrees n_1, n_2, it may be worthwhile depending upon C_1 to choose the smaller among n_1 and n_2 in order to have a fewer terms in the sum of square magnitudes of polynomials for $P(z_1, z_2)$.

Next we generalize the result in Theorem 2.1 to the multivariate setting which is known from [6].

Theorem 2.2 (*Dritschel [6]*). *Let $P(z_1, \ldots, z_d)$ be a multivariate Laurent polynomial which is strictly positive on the multivariate torus $|z_1| = |z_2| = \cdots = |z_d| = 1$, where $d \geqslant 2$ is an integer. Then $P(z_1, \ldots, z_d)$ can be expressed as a sum of square magnitudes of polynomials in $z_1, \ldots, z_d$.*

Proof. We shall use the arguments in the proof of the previous theorem. Write $P(z_1, z_2, \ldots, z_d) = P(z_1, z) = \sum_{j=-n_1}^{n_1} p_j(z) z_1^j > 0$, where z is the usual multi-variable notation beginning with z_2. We know that $P(z_1, z)$ is the symbol of the bi-infinite Toeplitz matrix given by (4) with z_2 replaced by the multivariable z.

It follows that any central section along the main diagonal is strictly positive definite as explained before. Write

$$P(z_1, z) = \mathbf{z_1}^\dagger P_{m_1}(z) \mathbf{z_1}, \tag{6}$$

where $\mathbf{z_1}$ given by Eq. (3) and $P_{m_1}(z) = \left[p_{j,k}\right]_{0 \leqslant j,k \leqslant m}$ is a matrix of size $(m_1 + 1) \times (m_1 + 1)$ with entries

$$p_{jk} = \frac{1}{m_1 + 1 - |j - k|} p_{j-k}(z), \quad \forall j, k = 0, 1, \ldots, m_1.$$

If $P > \varepsilon$ the argument in Theorem 2.1 shows that for m_1 large enough there is an $\varepsilon_1 > 0$ such that $\mathbf{x}^\dagger P_{m_1}(z)\mathbf{x} > \varepsilon_1 \|\mathbf{x}\|^2$ on the $d - 1$ torus if $\dfrac{2n_1(n_1+1)\widehat{C}_1}{\sqrt{3}(m_1 - n_1)} < \varepsilon$, where in this case $\widehat{C}_1 = \sup_{i, |z_j|=1, j=2,\ldots,d} |p_i(z)|$. Write $P_{m_1}(z_2, z') = \sum_{k=-n_2}^{n_2} \widetilde{p}_k(z') z_2^k$, where $\widetilde{p}_k$ are $(m_1 + 1) \times (m_1 + 1)$ Toeplitz matrices and $z' = (z_3, \ldots, z_d)$. Now set

$$\widehat{p}_{jk} = \frac{1}{m_2 + 1 - |j - k|} \widetilde{p}_{j-k}(z'), \quad \forall j, k = 0, \ldots, m_2$$

with $m_2 \geqslant m_1$ and $P_{m_2}(z') = \left[\widehat{p}_{j,k}\right]_{0 \leqslant j,k \leqslant m_2}$. As above we have that

$$\mathbf{x}^\dagger P_{m_2} \mathbf{x} = \frac{1}{m_2 + 1} \mathbf{x}^\dagger \mathcal{P}_{m_2} \mathbf{x} + \frac{1}{m_2 + 1} \mathbf{x}^\dagger \mathcal{R}_{m_2} \mathbf{x}.$$

As above the norm of $\mathcal{R}_{m_2}$ can be bounded by

$$\|\mathcal{R}_{m_2}\|_2 \leqslant \frac{2n_2(n_2 + 1)C_2}{\sqrt{3}(m_2 - n_2)},$$

where $C_2 = \sup_{i, |z_2| = \cdots |z_d| = 1} \|\tilde{p}_i(z')\|_2$, we find for m_2 sufficiently large, P_{m_2} is a positive matrix polynomial. We continue the process until we arrive at the positive trigonometric matrix polynomial $P_{m_{d-1}}(z_d)$ which can be factored by the matrix Fejér–Riesz theorem. We have thus established the proof. $\square$

Note that the number of factors will be $(m_1+1)(m_2+1)\cdots(m_{d-1}+1)$ and the degrees of the polynomials at most m_1 for $z_1 \ldots m_{d-1}$ for z_{d-1} and n_d for z_d. We note that we could have avoided the use of the matrix Fejér–Riesz theorem by eliminating all variables then using a square root of a positive matrix (see [17]). We will consider an alternative computationally attractive method for computing factorizations in the next section.

3. Computing approximate factorizations

As shown in the previous section, an important step in the factorization of multivariate Laurent polynomials is to compute the factorization of univariate polynomial matrices. Recall a computational algorithm for factorization of one variable Laurent trigonometric polynomials was developed in [15]. (This is a Bauer type method. See Remark 6.1 for differences.) This method can be extended to factorize positive definite polynomial matrices in the univariate setting. Let us first introduce some necessary notation and definitions in order to explain the method in more detail.

Let ℓ_2 stand for the space of all bi-infinite square summable sequences. Let $\|\mathbf{x}\|_2$ denote the standard norm on ℓ_2. We note that any bounded operator A from $\ell^2 \mapsto \ell^2$ can be expressed by a bi-infinite matrix.

Definition 3.1. A bi-infinite matrix $A = (a_{ik})_{i,k\in\mathbf{Z}}$ is said to be of exponential decay off its diagonal if

$$\|a_{ik}\|_2 \leqslant K r^{|i-k|}$$

for some constant K and $r \in (0,1)$, where $\mathbf{Z}$ is the collection of all integers. A is banded with band width b if $a_{ik} = 0$ for all $i, k \in \mathbf{Z}$ with $|i-k| > b$.

We suppose that A is a bounded operator throughout this section. If A is a positive operator, then there exists the unique positive bounded bi-infinite square root matrix Q of A such that $Q^2 = A$. If $A = B^\dagger B$ for bi-infinite Cholesky factorization B of A with positive entries on its diagonal, then there exists a unitary matrix U such that $B = UQ$.

Recall from the previous section that given any Laurent polynomial $P(z)$, we can view $P(z)$ to be the symbol of a bi-infinite Toeplitz matrix $\mathcal{P}$. The computational scheme introduced in [15] roughly speaking is to choose a central section

$$P_N = (p_{j-k})_{-N\leqslant j,k\leqslant N}$$

of matrix $\mathcal{P}$ and compute a Cholesky factorization, i.e. $P_N = C_N^\dagger C_N$ with C_N being an upper triangular matrix with positive diagonal entries if P_N is positive definite. If P_N is nonnegative definite use the singular value decomposition (SVD) to first find Q_N such that $P_N = Q_N^2$ and then find a Householder matrix H_N such that $C_N = H_N Q_N$ is upper triangular. The nonzero entries in the middle row of C_N approximate those in the middle row (in fact any row) of $\mathcal{C}$ whose symbol $C(z)$ is a factorization of $P(z)$, i.e., $P(z) = C(z)^* C(z)$.

For the extension of this method to polynomial matrices, let

$$\ell_k^m = \{\mathbf{x} = \{x_i\}_{i\in\mathbf{Z}}, x_i \in \mathbf{R}^m, \|\mathbf{x}\|_k < \infty\}, \quad k = 1, 2$$

and $B(\ell_2^m)$ be the set of bounded linear operators on ℓ_2^m. Let $\Pi_N \in B(\ell_2^m)$ be the projection given by

$$\Pi_N \mathbf{x} = \mathbf{y}, \quad \mathbf{y} = \{y_i : y_i = 0, |i| > N, y_i = x_i, \ |i| \leqslant N\}.$$

If $P \in B(\ell_2^m)$ is positive definite we will be interested in considering the $(2N+1)m \times (2N+1)m$ submatrix of P centered at the index zero which will be called the Nth central section. Note that the Nth central section is positive definite. We will also be interested in extensions of various finite matrices A_N to $B(\ell_2^m)$ given by

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & A_N & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

which with a slight abuse of notation will also be called A_N.

Consider the matrix polynomial $P(z) = \sum_{j=-n}^{n} p_j z^j$ with matrix coefficients p_k's of size $m \times m$, then $\mathcal{P} = (p_{i-j})_{i,j \in \mathbf{Z}} \in B(\ell_2^m)$ defined by $m \times m$ matrix blocks p_k, $-n \leqslant k \leqslant n$ is a bi-infinite block Toeplitz matrix whose symbol is $P(z)$. As shown earlier if $P(z)$ is Hermitian nonnegative definite, so is $\mathcal{P}$. Let $C(z)$ be a factorization of $P(z)$ i.e., $P(z) = C(z)^\dagger$ C(z), then $\mathcal{P} = \mathcal{C}^\dagger \mathcal{C}$, where $\mathcal{C}$ is a bi-infinite upper triangular banded block Toeplitz matrix associated with $C(z)$. On the other hand, if $\mathcal{P} = \mathcal{C}^\dagger \mathcal{C}$ for a upper triangular banded block Toeplitz matrix, then the symbol $C(z)$ of $\mathcal{C}$ satisfies $P(z) = C(z)^\dagger C(z)$. If $P(z)$ is positive definite then it follows from the matrix Fejér–Riesz theorem (cf. [14,23,22,17]) that it is possible to choose $\mathcal{C}$ so that it has positive diagonal entries. We shall prove the following (see also [18]):

Theorem 3.1. *Let $P(z) = \sum_{-n}^{n} p_k z^k$ be an $m \times m$ matrix polynomial that is positive definite for $|z| = 1$. Let $\mathcal{P} = (p_{i-j})_{i,j,\in \mathbf{Z}} = \mathcal{C}^\dagger \mathcal{C}$ where $\mathcal{C}$ is an upper triangular banded block Toeplitz with positive diagonal entries, P_N be the Nth central section of $\mathcal{P}$, and C_N the Cholesky factor of P_N (which we extend as described above). Then*

$$\|(C_N - \mathcal{C}_N)\delta\|_2 \leqslant K\rho^N,$$

for some $\rho \in (0, 1)$, where $\delta \in \ell_2^m$ is any vector with a finite number of nonzero entries.

For the numerical computation in the next section we will choose δ with zero components except for $\delta_0 = I_m$, the $m \times m$ identity matrix.

The proof of Theorem 3.1 is based upon the following:

Theorem 3.2. *Suppose that $A \in B(\ell_2)$ is a positive banded operator such that $\|A - I\|_2 < 1$. Let Q be the unique positive square root of A, A_N be a central section of A, and $\widehat{Q}_N$ be the positive matrix such that $\widehat{Q}_N^2 = A_N$.* Then

$$\|(Q - \widehat{Q}_N)\delta\|_2 \leqslant K\lambda^N \tag{7}$$

for some $\lambda \in (0, 1)$ and a positive constant K. In Eq. (7) δ is any vector with a fixed number of nonzero entries.

To prove the above Theorem 3.2, we begin with the following lemmas:

Lemma 3.3. *Suppose that A is banded with bandwidth b and $\|A-I\|_2 \leqslant r < 1$. If $Q^2 = A$ with $Q = (q_{ik})_{i,k\in\mathbf{Z}}$, then $|q_{l,k}| \leqslant K r^{\frac{|l-k|}{b}}$. If A is invertible, then the entries of Q^{-1} satisfy a similar bound.*

Proof. We only prove the exponential decay property of Q. The proof of that of Q^{-1} is similar. The uniqueness of Q and the convergence of the following series:

$$\sum_{i=0}^{\infty}(-1)^i \frac{(2i-3)!!}{(2i)!!}(A-I)^i$$

implies that

$$Q = \sqrt{A} = \sqrt{I + (A-I)} = \sum_{i=0}^{\infty}(-1)^i \frac{(2i-3)!!}{(2i)!!}(A-I)^i.$$

A is banded and so is $A-I$. If $A-I$ has bandwidth b, then $(A-I)^i$ is also banded with bandwidth ib. Thus,

$$q_{jk} = \sum_{i\geqslant |j-k|/b}^{\infty} (-1)^i \frac{(2i-3)!!}{(2i)!!}(A-I)^i_{jk},$$

where $(A-I)_{jk}$ denotes the (j,k)th entry of $A-I$ and similar for $(A-I)^i_{jk}$. It follows that

$$|q_{jk}| \leqslant K r^{|j-k|/b}$$

for some constant K. This finishes the proof. □

Let us write

$$Q = \begin{bmatrix} \alpha_1 & B & \alpha_2 \\ B^\dagger & Q_N & C^\dagger \\ \alpha_2^\dagger & C & \alpha_4 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} \beta_1 & a & \beta_2 \\ a^\dagger & A_N & c^\dagger \\ \beta_2^\dagger & c & \beta_4 \end{bmatrix}.$$

Note that $Q^2 = A$ implies $A_N = Q_N^2 + B^\dagger B + C^\dagger C$ or $\widehat{Q}_N^2 - Q_N^2 = B^\dagger B + C^\dagger C$, where $\widehat{Q}_N^2 = A_N$. Thus, we have

$$(Q_N + \widehat{Q}_N)(\widehat{Q}_N - Q_N) = \widehat{Q}_N^2 - Q_N^2 + Q_N\widehat{Q}_N - \widehat{Q}_N Q_N = B^\dagger B + C^\dagger C + R, \qquad (8)$$

where R is defined in the following:

Lemma 3.4 (*cf. Lai [15]*). *Let $R = (r_{jk})_{-N\leqslant j,k\leqslant N} := Q_N\hat{Q}_N - \hat{Q}_N Q_N$. Then $r_{jk} = O(r^{N/(4b)})$ for $k = -N/4+1, \ldots, N/4-1$ and $j = -N, \ldots, N$.*

Proof of Theorem 3.2. From Eq. (8) we find that, $(\widehat{Q}_N - Q_N) = (Q_N + \widehat{Q}_N)^{-1}(B^\dagger B + C^\dagger C + R)$. By Lemma 3.3., we can prove that the entries of $B^\dagger B + C^\dagger C$ have the exponential decay property: $(B^\dagger B + C^\dagger C)_{jk} = O(r^{N-|k|})$, $-N \leqslant k \leqslant N$.

The positivity of A implies that Q is positive and so is Q_N. It follows that $\|Q_N^{-1}\|_2$ is uniformly bounded. Thus, we have

$$\|(Q_N + \widehat{Q}_N)^{-1}\|_2 \leqslant \|Q_N^{-1}\|_2 \leqslant K_1 < \infty$$

for a positive constant K_1 independent of N, where we have used the fact that $\widehat{Q}_N$ is nonnegative. Therefore, we conclude that

$$\begin{aligned}\|(\widehat{Q}_N - Q_N)\delta_N\|_2 &\leqslant \|(Q_N + \widehat{Q}_N)^{-1}\| \|(B^\dagger B + C^\dagger C + R)\delta_N\|_2 \\ &\leqslant K_1 \|(B^\dagger B + C^\dagger C + R)\delta_N\|_2,\end{aligned}$$

where δ_N is the finite vector whose entries match those of δ. The proof is completed by extending Q_N, $\widehat{Q}_N$, replacing δ_N by δ, and noticing that by Lemma 3.3, $\|(Q_N - Q)\delta\|_2 < K_1\lambda^N$, $\lambda < 1$. □

Proof of Theorem 3.1. Suppose that

$$\sup_{|z|=1} \|P(z)\|_2 < 1. \tag{9}$$

Otherwise divide P by a sufficiently large constant so that (9) holds. Let Q be the unique positive square root of $\mathcal{P}$, and $\widehat{Q}_N$ the positive square root of P_N. From Theorem 3.2 we know that $\|(\widehat{Q}_N - Q)\delta\|_2 \leqslant K\rho^N$ with $\rho < 1$. Let U be the unitary matrix such that $\mathcal{C} = UQ$ which is upper triangular. Then

$$\|(\widehat{Q}_N - Q)\delta\|_2 = \|(U\widehat{Q}_N - \mathcal{C})\delta\|_2.$$

The above equation implies that the diagonal elements of $U\widehat{Q}_N$ tend exponentially fast to the positive diagonal entries of $\mathcal{C}$. Moreover, let $(\tilde{q}_{i,0})_{i\in\mathbb{Z}}$ be the central column of $U\widehat{Q}_N$. Since $\mathcal{C}$ is upper triangular and banded with bandwidth b, we have

$$\sum_{i<0} |\tilde{q}_{i,0}|^2 + \sum_{i>b} |\tilde{q}_{i,0}|^2 \leqslant \|(U\widehat{Q}_N - \mathcal{C})\delta\|_2 \leqslant K^2\rho^{2N}, \tag{10}$$

by Theorem 3.2.

Write $U\widehat{Q}_N = \widetilde{Q}_N + L_N^1$, where $\widetilde{Q}_N$ is upper triangular and L_N^1 is strictly lower triangular. Then $U\widehat{Q}_N = q_N + l_N$, where $q_N = \Pi_N \widetilde{Q}_N \Pi_N^\dagger$ and $l_N = L_N^1 + \widetilde{Q}_N - q_N$. Since $\mathcal{C}$ is upper triangular and banded, Eq. (10) shows that $\|l_N\delta\|_2$ tends to zero exponentially fast because $\|l_N\delta\|_2^2 = \sum_{i<0} |\tilde{q}_{i,0}|^2 + \sum_{i>N} |\tilde{q}_{i,0}|^2$. The fact that $\widehat{Q}_N$ is symmetric implies

$$\begin{aligned}P_N &= \widehat{Q}_N^2 = \widehat{Q}_N^\dagger \widehat{Q}_N = (U\widehat{Q}_N)^\dagger (U\widehat{Q}_N) \\ &= (q_N + l_N)^\dagger (q_N + l_N) \\ &= q_N^\dagger q_N + l_N^\dagger q_N + q_N^\dagger l_N + l_N^\dagger l_N,\end{aligned}$$

so that

$$C_N^\dagger C_N - q_N^\dagger q_N = l_N^\dagger q_N + q_N^\dagger l_N + l_N^\dagger l_N.$$

Since Q_N is uniformly bounded so is q_N and we find that $\|l_N^\dagger l_N \delta\|_2$ and $\|q_N^\dagger l_N \delta\|_2$ go to zero exponentially fast. Also, we claim that $\|l_N^\dagger q_N \delta\|_2 < K_3\lambda^N$. Indeed, as we know $\|l_N\delta\|_2 \leqslant K\rho^N$ which implies $\|l_N\delta_i\|_2 \leqslant K\rho^N$ for δ_i which is a zero vector except for the ith component which is 1, $i = 1, 2, \ldots, b$. Write $l_N = (\ell_{ij})_{-N\leqslant i,j\leqslant N}$ with $\ell_{ij} = 0$ for $i > j$. (Note that we arrange the indices so that $\ell_{N,N}$ is on the top left corner of matrix l_N while $\ell_{-N,-N}$ is the low right corner of l_N.) We know that $\sum_{i<j} |\ell_{ij}|^2 < K\rho^N$ for $j = 0, 1, \ldots, b$. Also, let $(q_{N,i})$ be the central column of q_N. Note that $q_{N,i} = \tilde{q}_{i,0}$ for $i = -N, \dots, N$ and $\tilde{q}_{i,0} = 0$ for $i < 0$. It follows that the only nonzero entries of $l_N^\dagger q_N \delta$ are those with $j \geqslant 0$ thus,

$$\sum_{i=-N}^{N} \ell_{ij} \tilde{q}_{i,0} = \sum_{i=0}^{N} \ell_{ij} \tilde{q}_{i,0} = \sum_{i=0}^{b} \ell_{ij} \tilde{q}_{i,0} + \sum_{i=b+1}^{j} \ell_{ij} \tilde{q}_{i,0}.$$

Hence, by Eq. (10) we have

$$\begin{aligned}
\|l_N^\dagger q_N \delta\|_2^2 &\leqslant 2 \sum_{j=0}^{N} \left(\left| \sum_{i=0}^{b} \tilde{q}_{i,0} \ell_{ij} \right|^2 + \left| \sum_{i=b+1}^{j} \ell_{ij} \tilde{q}_{i,0} \right|^2 \right) \\
&\leqslant \sum_{i=0}^{b} |\tilde{q}_{i,0}|^2 \sum_{i=0}^{b} \|l_N \delta_i\|_2^2 + \sum_{i=0}^{N} |\tilde{q}_{i,0}|^2 \sum_{j=0}^{N} \sum_{i=b+1}^{j} |\ell_{ij}|^2 \\
&\leqslant K_1 \rho^{2N} + (N+1) \|\widehat{Q}_N\|_2^2 \rho^{2N} \\
&\leqslant K_2 \lambda^{2N},
\end{aligned}$$

for another $\lambda \in (0, 1)$ and constant $K_2 > 0$. Therefore,

$$\|(C_N^\dagger C_N - q_N^\dagger q_N)\delta\|_2 < K_3 \lambda^N,$$

where we recall that $C_N = [c_{ij}]_{-N \leqslant i,j \leqslant N}$ is the Cholesky factorization of the central section P_N of $\mathcal{P}$. Restricting the above quantities to their finite matrices we note because of the strict positivity of P, $\|q_N\|_2$ is uniformly bounded from below hence q_N^{-1} is uniformly bounded. Furthermore since C_N has the same size as q_N,

$$\|(I - (q_N^\dagger)^{-1} C_N^\dagger C_N q_N^{-1})\delta_N\|_2 < K_4 \lambda^N,$$

where $\delta_N = q_N \delta$ for any δ with finitely many nonzero entries. The above inequality shows that $\|(C_N q_N^{-1} - I)\delta_N\|_2 \leqslant K_4 \lambda^N$. Indeed, writing $(a_{N,ij})_{-N \leqslant i,j \leqslant N} = C_N q_N^{-1}$, we note that $a_{ij} = 0$ for $i < j$ since both C_N and q_N are upper triangular and each entry $a_{N,i,i}$ on the diagonal is bounded below by the uniform boundedness of C_N^{-1} and q_N. Thus

$$\|(I - (q_N^\dagger)^{-1} C_N^\dagger C_N q_N^{-1})\delta\|_2^2 = \sum_{j=-N}^{N} \left| \sum_{i=j}^{N} a_{N,i,j} a_{N,i,0} - \delta_{j0} \right|^2 \leqslant K_4^2 \lambda^{2N}.$$

From the above inequality, we conclude $|a_{N,N,0}| \leqslant K_4 \lambda^N$ for $j = N$. By induction we can show that $|a_{N,j,0}| \leqslant K_4 \lambda^N$ for $j = 1, \dots, N-1$. For $j = 0$, we have

$$\left| \sum_{i=0}^{N} a_{N,i,0}^2 - 1 \right|^2 \leqslant K_4^2 \lambda^{2N}.$$

It follows that $|a_{N,0,0} - 1| \leqslant K_4 N \lambda^N$. Hence, we have

$$\|(C_N q_N^{-1} - I)\delta\|_2 \leqslant K_5 v^N$$

for another real number $v \in (0, 1)$. Therefore,

$$\|(C_N - U\widehat{Q}_N)\delta\|_2 \leqslant \|(C_N - q_N)\delta\|_2 + \|\ell_N\delta\|_2 \leqslant K_6 v^N$$

so that

$$\|(C_N - \mathcal{C})\delta\|_2 \leqslant \|(C_N - U\widehat{Q}_N)\delta\|_2 + \|(U\widehat{Q}_N - \mathcal{C})\delta\|_2 \leqslant K_7 v^N.$$

This completes the proof. □

4. Numerical examples

In this section we give three examples to illustrate how the computational method works for polynomial matrix factorizations.

Example 4.1. We first consider a univariate polynomial matrix

$$P(z) := \begin{bmatrix} 8+z+1/z & 1+z \\ 1+1/z & 1 \end{bmatrix}.$$

It is clear that the matrix is Hermitian and positive definite. We write

$$P(z) = \begin{bmatrix} 8 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} z + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} /z.$$

We assemble a bi-infinite Toeplitz matrix whose 10×10 block is as shown below

$$\begin{bmatrix} 8 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 8 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 8 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 8 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 8 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

We use the Cholesky factorization method to a 40×40 central block and get a lower triangular matrix F. Let $P0$ be the 2×2 block from the middle rows and columns of F (e.g., $(F_{ij})_{19 \leqslant i,j \leqslant 20}$ which is

$$P0 := \begin{bmatrix} \frac{\sqrt{385}}{7} & 0 \\ \frac{6}{\sqrt{385}} & \frac{\sqrt{2310}}{55} \end{bmatrix}.$$

Choose the 2×2 block next to $P0$ in the same rows as that of $P0$ as $P1$. That is,

$$P1 := \begin{bmatrix} \frac{\sqrt{385}}{55} & \frac{-\sqrt{2310}}{385} \\ \frac{\sqrt{385}}{55} & \frac{-\sqrt{2310}}{385} \end{bmatrix}.$$

Define $Q^\dagger(z) = P0 + P1/z$ and then we have $P(z) = Q(z)^\dagger Q(z)$.

Example 4.2. We next consider a bivariate polynomial

$$\begin{aligned}P(x,y) = {} & 41 + 5x^2 + 5y^2 + 15/x + 20/y + 5/x^2 + 5/y^2 + 15x + 20y + 5xy \\ & + 8y/x + 5/(xy) + 8x/y + 2x/y^2 + 3y/x^2 + 3x^2/y + x^2/y^2 \\ & + 2y^2/x + y^2/x^2.\end{aligned}$$

It is a positive polynomial since $P(x,y) = p(x,y)p(1/x,1/y)$ with $p(x,y) = 5 + 2x + 3y + xy + x^2 + y^2$. Let us write

$$P(x,y) = [1, 1/x, 1/x^2]\widetilde{P}(y)\begin{bmatrix}1\\ x\\ x^2\end{bmatrix},$$

with

$$\widetilde{P}(y) := \begin{bmatrix} \frac{41}{3} + \frac{5y^2}{3} + \frac{20}{3y} + \frac{5}{3y^2} + \frac{20}{3}y & \frac{15}{2} + \frac{5}{2}y + \frac{4}{y} + \frac{1}{y^2} & 5 + \frac{3}{y} + \frac{1}{y^2} \\ \frac{15}{2} + 4y + \frac{5}{2y} + y^2 & \frac{41}{3} + \frac{5}{3}y^2 + \frac{20}{3y} + \frac{5}{3y^2} + \frac{20}{3}y & \frac{15}{2} + \frac{5}{2}y + \frac{4}{y} + \frac{1}{y^2} \\ 5 + 3y + y^2 & \frac{15}{2} + 4y + \frac{5}{2y} + y^2 & \frac{41}{3} + \frac{5}{3}y^2 + \frac{20}{3y} + \frac{5}{3y^2} + \frac{20}{3}y \end{bmatrix}.$$

The above matrix polynomial can be rewritten as $\widetilde{P}(y) = \sum_{j=-2}^{2} p_j y^j$ with $p_{-2}, \ldots, p_2$ being given below

$$p_0 = \begin{bmatrix} \frac{41}{3} & \frac{15}{2} & 5 \\ \frac{15}{2} & \frac{41}{3} & \frac{15}{2} \\ 5 & \frac{15}{2} & \frac{41}{3} \end{bmatrix}, \quad p_1 = \begin{bmatrix} \frac{20}{3} & \frac{5}{2} & 0 \\ 4 & \frac{20}{3} & \frac{5}{2} \\ 3 & 4 & \frac{20}{3} \end{bmatrix}, \quad p_{-1} = p_1^\dagger,$$

$$p_2 = \begin{bmatrix} \frac{5}{3} & 0 & 0 \\ 1 & \frac{5}{3} & 0 \\ 1 & 1 & \frac{5}{3} \end{bmatrix}, \quad p_{-2} = p_2^\dagger.$$

We now assemble a bi-infinite Toeplitz matrix whose 9×9 central block is shown as follows:

$$\begin{bmatrix} \frac{41}{3} & \frac{15}{2} & 5 & \frac{20}{3} & \frac{5}{2} & 0 & \frac{5}{3} & 0 & 0 \\ \frac{15}{2} & \frac{41}{3} & \frac{15}{2} & 4 & \frac{20}{3} & \frac{5}{2} & 1 & \frac{5}{3} & 0 \\ 5 & \frac{15}{2} & \frac{41}{3} & 3 & 4 & \frac{20}{3} & 1 & 1 & \frac{5}{3} \\ \frac{20}{3} & 4 & 3 & \frac{41}{3} & \frac{15}{2} & 5 & \frac{20}{3} & \frac{5}{2} & 0 \\ \frac{5}{2} & \frac{20}{3} & 4 & \frac{15}{2} & \frac{41}{3} & \frac{15}{2} & 4 & \frac{20}{3} & \frac{5}{2} \\ 0 & \frac{5}{2} & \frac{20}{3} & 5 & \frac{15}{2} & \frac{41}{3} & 3 & 4 & \frac{20}{3} \\ \frac{5}{3} & 1 & 1 & \frac{20}{3} & 4 & 3 & \frac{41}{3} & \frac{15}{2} & 5 \\ 0 & \frac{5}{3} & 1 & \frac{5}{2} & \frac{20}{3} & 4 & \frac{15}{2} & \frac{41}{3} & \frac{15}{2} \\ 0 & 0 & \frac{5}{3} & 0 & \frac{5}{2} & \frac{20}{3} & 5 & \frac{15}{2} & \frac{41}{3} \end{bmatrix}.$$

We use the Cholesky factorization of a central block matrix of size 120 × 120. Let F be the lower triangular factorization. Then choose Q_0 to be the 3×3 block at the middle rows and columns of F (e.g., $(F_{ij})_{58 \leqslant i,j \leqslant 60}$), Q_1 the 3×3 block next to Q_1 in the same rows of Q_1 and Q_2 the 3×3 block next to Q_1 in the same rows of Q_1. That is

$$Q_0 = \begin{bmatrix} 3.185602126 & 0 & 0 \\ 1.873651218 & 2.539725049 & 0 \\ 1.524622962 & 1.128505745 & 2.269126602 \end{bmatrix},$$

$$Q_1 = \begin{bmatrix} 1.797364251 & 0.08381502303 & -0.0003518239229 \\ 0.7675275947 & 1.633796832 & 0.06150315980 \\ 0.00008111923034 & 0.9665117592 & 1.856367398 \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} 0.5231873284 & 0.007768330871 & 0.08530594055 \\ 0 & 0.6562390159 & 0.1143305535 \\ 0 & 0 & 0.7344969935 \end{bmatrix}.$$

Let $Q(y)^\dagger = Q_0 + Q_1/y + Q_2/y^2$ and then $Q(y)^\dagger Q(y) \approx \widetilde{P}(y)$. In fact the maximum error of each entry of $Q(y)Q(y)^* - \widetilde{P}(y)$ is less than or equal to 10^{-8}.

Example 4.3. Let us consider a bivariate polynomial which has a zero on the bi-torus

$$P(x, y) = 30 + 14/x + 11/y + 4/x/y + 14x + 6x/y + 11y + 6y/x + 4xy.$$

It is the product of $P(x, y) = (4 + 3x + 2y + 1)(4 + 3/x + 2/y + 1)$ which is zero at $x = -1$, $y = -1$. We write

$$P(x, y) = p_0(y) + p_1(y)x + p_{-1}(y)/x$$

for $p_0(y) = 30 + 11/y + 11y$, $p_1(y) = 14 + 6y + 4/y$, and $p_{-1}(y) = 14 + 4y + 6/y$. It is the symbol of an bi-infinite Toeplitz matrix. One of its central section is as shown below

$$\begin{bmatrix} 11/y + 30 + 11y & 4/y + 14 + 6y & 0 & 0 \\ 6/y + 14 + 4y & 11/y + 30 + 11y & 4/y + 14 + 6y & 0 \\ 0 & 6/y + 14 + 4y & 11/y + 30 + 11y & 4/y + 14 + 6y \\ 0 & 0 & 6/y + 14 + 4y & 11/y + 30 + 11y \end{bmatrix}.$$

Since $P(x, y)$ has no simple factors (see the next section), any central sections of the bi-infinite Toeplitz matrix is positive by Lemma 5.1. We consider several central sections P_m of size $m = 32 \times 32, 64 \times 64, 128 \times 128$ and 256×256. For each of these central sections, P_m is a univariate polynomial in y with matrix coefficients and $P_m(y)$ is positive. Thus, $P_m(y) = Q_m(y)^\dagger Q_m(y)$. To compute $Q_m(y)$, we use the computational method in Section 3 to yield an approximation $\tilde{Q}_m$ of Q_m. As the size of central sections increases, the Q_m converges to the corresponding entries in the bi-infinite Toeplitz matrix. We use the entries in the center of the middle rows of $\tilde{Q}_m$ to construct an approximation of $Q_m(y)$ and hence the factorization of $P(x, y)$ and listed below

$$\left[\begin{array}{cc} \text{size} & \text{factorization} \\ 16 \times 16 & 4.01207952 + 2.984741799x + 2.000226870y + 0.996712925xy \\ 32 \times 32 & 4.004041536 + 2.994924757x + 2.000034879y + 0.998949058xy \\ 64 \times 64 & 4.001381387 + 2.998269650x + 2.000005690y + 0.999648058xy \\ 128 \times 128 & 4.00069369 + 2.999134582x + 1.99999896y + 0.999821915xy \end{array}\right].$$

As we know that the factorization is $4 + 3x + 2y + 1$, the approximations are very good.

5. Nonnegative bivariate trigonometric polynomials

Finally we consider the problem of factorization of nonnegative multivariate polynomials. Let us start with $P(z, w) \geqslant 0$. If for some z_0 with $|z_0| = 1$, $P(z_0, w) = 0$ for all w with $|w| = 1$, we say that $P(z, w)$ has a simple factor at z_0. If $P(z, w)$ has a simple factor at z_0, then $P(z, w)$ has factors $(z - z_0)$ and $(1/z - 1/z_0)$. Let us factor them out. Then $P(z, w)/((z - z_0)(1/z - 1/z_0))$ is still nonnegative. Similarly, if $P(z, w_0) = 0$ for all z with $|z| = 1$, $P(z, w)$ has a simple factor at w_0. In this case, $P(z, w)$ has two factors $(w - w_0)$ and $(1/w - 1/w_0)$ which can be factored out from $P(z, w)$. Without loss of generality, we may assume that $P(z, w) \geqslant 0$ does not have any simple factors. Writing $P(z, w) = \sum_{j=-n}^{n} p_j(w) z^j$, we view that $P(z, w)$ is a polynomial of z and it is the symbol of a bi-infinite Toeplitz matrix in (4) with w in place of z_2. We have the following:

Lemma 5.1. *Suppose that $P(z, w) \geqslant 0$ does not have any simple factors. Then any central section of the bi-infinite Toeplitz matrix in* (4) *is strictly positive definite.*

Proof. Since $P(z, w) \geqslant 0$, we know that any central section of the matrix in (4) is nonnegative definite. Suppose that a central section $T_m(w)$ of the matrix in (4) is not positive definite for $w = w_0$. Then there exists a vector $\mathbf{x}$ such that $T_m(w_0)\mathbf{x} = 0$, i.e., $\mathbf{x}^\dagger T_m(w_0)\mathbf{x} = 0$. Thus, we have, for $z = e^{i\theta}$,

$$0 = \mathbf{x}^\dagger T_m(w_0)\mathbf{x} = \frac{1}{2\pi} \int_0^{2\pi} F(\mathbf{x})^* P(z, w_0) F(\mathbf{x})\, \mathrm{d}\theta.$$

It follows that

$$|F(\mathbf{x})|^2 P(z, w_0) = 0, \quad \text{a.e.}$$

and hence, $P(z, w_0) \equiv 0$ since $|F(\mathbf{x})| \neq 0$, a.e. and $P(z, w_0)$ is a Laurent polynomial. That is, $P(z, w)$ has a simple factor at w_0. This contradicts the assumption on $P(z, w)$. $\square$

Thus, for a central section P_m of size $m \times m$ in the matrix in (4), P_m is positive. Since P_m is a polynomial matrix in w, by the matrix Fejér–Riesz factorization theorem (cf. [14,22,23,17]), P_m can be factorized into Q_m, i.e., $P_m(w) = Q_m(w)^\dagger Q_m(w)$. Intuitively, the polynomial Q_m is a good approximation of the factorization of the bi-infinite Toeplitz matrix $\mathcal{P}$ in (4) for m sufficiently large. In the previous section, we presented an example (Example 4.3.) of $P(z, w)$ which is nonnegative without simple factors. Using our symbol approximation method, we compute an approximation of the factorization of P_m for $m = 16, 32, 64$, and 128. The numerical computation shows the factorizations converge.

Let us now discuss the convergence a little bit more in detail. For simplicity, let $\mathcal{A}$ be a bi-infinite Toeplitz matrix associated with a univariate Laurent polynomial $A(z)$ and $\mathcal{A}_N = (a_{jk})_{-N \leqslant j,k \leqslant N}$ be a central section of size $(2N+1) \times (2N+1)$ for a positive integer N. Suppose that each $\mathcal{A}_N$ is strictly positive. Thus we can obtain a factorization $\mathcal{A}_N = \mathcal{B}_N^* \mathcal{B}_N$ by Cholesky factorization with positive entries on its diagonal of $\mathcal{B}$.

Lemma 5.2. *For any* $\mathbf{x}, \mathbf{y} \in \ell_2$, $\mathbf{x}^\dagger \mathcal{A}_N \mathbf{y} := \mathbf{x}_N^\dagger \mathcal{A}_N \mathbf{y}_N$ *converges to* $\mathbf{x}^\dagger \mathcal{A} \mathbf{y}$ as $N \longrightarrow +\infty$, *where* $\mathbf{x}_N = (x_{-N}, \ldots, x_0, \ldots, x_N)^\dagger$ *is the central section of size* $2N+1$ *of* $\mathbf{x}$ *around the index* 0 *and similar for* $\mathbf{y}_N$.

Proof. For an integer $N > 0$,

$$\begin{aligned}
&\mathbf{x}^\dagger \mathcal{A}_N \mathbf{y} - \mathbf{x}^\dagger \mathcal{A} \mathbf{y} \\
&\quad = \frac{1}{2\pi} \int_0^{2\pi} \left(F(\mathbf{x}_N)^* A(z) F(\mathbf{y}_N) - F(\mathbf{x})^* A(z) F(\mathbf{y}) \right) \mathrm{d}\theta \\
&\quad = \frac{1}{2\pi} \int_0^{2\pi} \left(F(\mathbf{x}_N) - F(\mathbf{x}) \right)^* A(z) F(\mathbf{y}_N) \, \mathrm{d}\theta \\
&\qquad + \frac{1}{2\pi} \int_0^{2\pi} F(\mathbf{x})^* A(z) \left(F(\mathbf{y}_N) - F(\mathbf{y}) \right) \mathrm{d}\theta,
\end{aligned}$$

where $z = e^{i\theta}$. In the first equality we used the fact that $\mathbf{x}^\dagger \mathcal{A}_N \mathbf{x} = (\Pi_N \mathbf{x})^\dagger \mathcal{A} \Pi_N \mathbf{x}$ where Π_N is the projection defined in Section 3. Thus

$$\begin{aligned}
&|\mathbf{x}^\dagger \mathcal{A}_N \mathbf{y} - \mathbf{x}^\dagger \mathcal{A} \mathbf{y}| \\
&\quad \leqslant \|\mathbf{x} - \mathbf{x}_N\|_2 \|A(z)\|_\infty \|\mathbf{y}\|_2 + \|\mathbf{y} - \mathbf{y}_N\|_2 \|A(z)\|_\infty \|\mathbf{x}\|_2 \\
&\quad \to 0,
\end{aligned}$$

as $N \to +\infty$. Here, $\|A(z)\|_\infty$ denotes the maximum norm of $A(z)$ over the circle $|z| = 1$. This completes the proof. □

A consequence of the above Lemma 5.2 is that $\|\mathcal{B}_N \mathbf{x}\|_2^2$ converges to $\mathbf{x}^\dagger \mathcal{A} \mathbf{x}$. If $\mathcal{A}$ can be factored to $\mathcal{A} = \mathcal{B}^\dagger \mathcal{B}$. Then $\|\mathcal{B}_N \mathbf{x}\|_2 \longrightarrow \|\mathcal{B} \mathbf{x}\|_2$. The following is another consequence of Lemma 5.2.

Lemma 5.3. *Let* $\mathcal{B}_N$ *be a factorization of* $\mathcal{A}_N$, *i.e.*, $\mathcal{A}_N = \mathcal{B}_N^\dagger \mathcal{B}_N$. *Then* $\|\mathcal{B}_N\|$ *is bounded independent of* N.

Proof. By Lemma 5.2, there exists a constant C such that for N large enough,

$$\|\mathcal{B}_N \mathbf{x}\|_2^2 = \mathbf{x}^\dagger \mathcal{A}_N \mathbf{x} \leqslant \mathbf{x}^\dagger \mathcal{A} \mathbf{x} + C = \|\mathbf{x}\|_2^2 \|A(z)\|_\infty + C.$$

Hence, $\|\mathcal{B}_N\| := \max\limits_{\substack{\mathbf{x} \in \ell_2 \\ \|\mathbf{x}\|_2 = 1}} \|\mathcal{B}_N \mathbf{x}\|_2$ is bounded. □

Note that for every N, $\mathcal{B}_N$ banded with the same band width as that of $\mathcal{A}$. Thus, each row (or column) of $\mathcal{B}_N$ has finitely many nonzero entries. Lemma 5.3 implies that each row (or column) of $\mathcal{B}_N$ is bounded in ℓ_2 norm and hence each entry in any row is bounded. Therefore, there exists a subsequence of $\mathcal{B}_{N_j}$ such that each entry with indices (j,k) in $\mathcal{B}_{N_i}$ converges as $i \longrightarrow +\infty$. That is, for any vector $\mathbf{x} = (x_i)_{i\in\mathbf{z}} \in \ell_2$ with finitely many nonzero entries, we have

$$\mathcal{B}_{N_i}\mathbf{x} \longrightarrow \mathcal{B}\mathbf{x}$$

for a bi-infinite matrix $\mathcal{B}$. By Lemma 5.3 again, we have $\mathbf{x}^\dagger\mathcal{B}^\dagger\mathcal{B}\mathbf{y} = \mathbf{x}^\dagger\mathcal{A}\mathbf{y}$. Then $\mathcal{B}^\dagger\mathcal{B} = \mathcal{A}$. Note that $\mathcal{B}$ is an upper triangular matrix with the same band width as that of $\mathcal{A}$. If $\mathcal{B}$ is a Toeplitz matrix, we immediately know that $A(z)$ has a factorization such that $A(z) = B(z)^*B(z)$. Therefore, we end with:

Theorem 5.4. *Let $P(z,w)$ be a nonnegative Laurent polynomial with no simple zeros. Let $\mathcal{P}$ be a bi-infinite Toeplitz matrix with Laurent polynomial entries in w. Then $\mathcal{P}$ naturally induces a nonnegative operator $\mathcal{B}$ on ℓ_2 such that $\mathcal{P} = \mathcal{B}^\dagger\mathcal{B}$ and there is a subsequence of $\mathcal{B}_N$ convergent to $\mathcal{B}$ entrywise, where $\mathcal{B}_N$ is a factorization of a central section $\mathcal{P}_N$ of $\mathcal{P}$, i.e., $\mathcal{B}_N^\dagger\mathcal{B}_N = \mathcal{P}_N$. If $\mathcal{B}$ is Toeplitz, then $P(z,w)$ can be factored into a sum of square magnitudes of finitely many polynomials in z and w.*

Theorem 5.4 provides a computational method to check if a nonnegative Laurent polynomial $P(z,w)$ can be factorized. That is, we compute Cholesky factorization of central sections of the bi-infinite Toeplitz matrix $\mathcal{P}$ associated with $P(z,w)$ and observe if the factorization matrices converge to a Toeplitz matrix or not. If they converge, $P(z,w)$ has a factorization.

6. Remarks

1. It is interesting to point out that the symbol approximation method discussed in [15] is very much like the Bauer method in [2]. One slight difference is that the singular value decomposition (SVD) instead of the Cholesky decomposition is used to factorize the matrices associated with Laurent polynomial $P(z)\geqslant 0$. Another slight difference is that the central section $P_N = (p_{ij})_{-N\leqslant i,j\leqslant N}$ in [15] is used instead of $P_N = (p_{ij})_{0\leqslant i,j\leqslant N}$ in [2].
2. When $P(z)$ is a matrix polynomial in the univariate setting [13] have demonstrated a constructive method to factor $P(z) = Q(z)^\dagger Q(z)$ when $P(z)$ has a nonzero monomial determinant.

Acknowledgments

The authors would like to thank the referees for providing references [9–12,14,18–20,22,23] which are directly related to the topic of the research presented in this paper.

References

[1] S. Basu, A constructive algorithm for 2D spectral factorization with rational spectral factors, IEEE Trans. Circuits and Systems 47 (2000) 1309–1318.
[2] F.L. Bauer, Ein direktes Iterationsverfahren zur Hurwitz-zrlegung eines Polynoms, Arch. Elektr. Uebertragung 9 (1955) 285–290.
[3] F.L. Bauer, Beiträge zur Einwicklung numerischer Verfahren für programmmgeteuerte Rechenanlagen II. Direkte Faktorizierung eines Polnoms, Sitz. Ber. Bayer. Akad. Wiss. (1956) 163–203.
[4] A. Calderon, R. Pepinsky, On the phases of Fourier coefficients for positive real periodic functions, in: R. Perpinsky (Ed.), Computing Methods and Phase Problem in X-Ray Crystal Analysis, 1952, pp. 339–346.
[5] I. Daubechies, Ten Lectures on Wavelets, SIAM Publications, Philadelphia, 1992.

[6] M.A. Dritschel, On factorization of trigonometric polynomials, Integral Equations Operator Theory 49 (2004) 11–42.

[7] L. Fejér, Über trigonometrische Polynome, J. Reine Angew. Math. 146 (1915) 53–82.

[8] J. Geronimo, H.J. Woerdeman, Positive extensions, Fejér–Riesz factorization and autoregressive filters in two variables, Ann. Math. 160 (2004) 839–906.

[9] T.N.T. Goodman, C.A. Micchelli, G. Rodrigues, S. Seatzu, On the Cholesky factorization of the Gram matrix of locally supported functions, BIT 35 (1995) 233–257.

[10] T.N.T. Goodman, C.A. Micchelli, G. Rodrigues, S. Seatzu, Spectral factorization of Laurent polynomials, Adv. Comput. Math. 7 (1997) 429–454.

[11] T.N.T. Goodman, C.A. Micchelli, G. Rodrigues, S. Seatzu, On the Cholesky factorization of the Gram matrix of multivariate functions, SIAM J. Matrix Anal. Appl. 22 (2000) 501–526.

[12] U. Grenender, G. Szego, Toeplitz Forms and Their Applications, Chelsea, New York, 1958.

[13] D. Hardin, T. Hogen, Q. Sun, The matrix-valued Riesz Lemma and local orthonormal bases in shift-invariant spaces, Adv. Comput. Math. 20 (2004) 367–384.

[14] H. Helson, Lectures on Invariant Subspaces, Academic Press, New York, 1964.

[15] M.J. Lai, On the computation of Battle–Lemarie's wavelets, Math. Comput. 63 (1994) 689–699.

[16] M.J. Lai, J. Stöckler, Construction of multivariate compactly supported tight wavelet frames, Manuscript, 2004.

[17] J.W. McLean, H.J. Woerdeman, Spectral factorizations and sums of squares representation via semi-definite programming, SIAM J. Matrix Anal. Appl. 23 (2001) 646–655.

[18] C.V.M. van der Mee, S. Seatzu, G. Rodrigues, Block Cholesky factorization of infinite matrices and orthonormalization of vectors of functions, in: Z. Chen, Y. Li, C.A. Micchelli, Y. Xu (Eds.), Advances in Computational Mathematics, Dekker, NY, 1998, pp. 423–455.

[19] C.V.M. van der Mee, S. Seatzu, G. Rodrigues, Spectral factorization of bi-infinite block Toeplitz matrices, in: Recent Advances in Numerical Analysis, Nova Science Publishers, Huntington, NY, 2001, pp. 223–248.

[20] C.V.M. van der Mee, S. Seatzu, G. Rodrigues, Spectral factorization of bi-infinite multi-index block Toeplitz matrices, Linear Algebra Appl. 343/344 (2002) 355–380.

[21] F. Riesz, Über ein Problem des Herr Carathéodory, J. Reine Angew. Math. 146 (1915) 83–87.

[22] M. Rosenblatt, A multidimensional prediction problem, Arch. Mat. 3 (1958) 407–424.

[23] M. Rosenblum, Vectorial Toeplitz operators and Fejér–Riesz theorem, J. Math. Anal. Appl. 23 (1968) 139–147.

[24] W. Rudin, The existence problem for positive definite functions, Illinois J. Math. 7 (1963) 532–539.

[25] A.H. Sayed, T. Kailath, A survey of spectral factorization methods, Numer. Linear Algebra Appl. 8 (2001) 467–496.

[26] D. Youla, N. Kazanjian, Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle, IEEE Trans. Circuits and Systems 25 (1978) 57–69.