Noname manuscript No.
(will be inserted by the editor)

The Palindromic Cyclic Reduction and related algorithms

Bruno Iannazzo · Beatrice Meini

the date of receipt and acceptance should be inserted later

Abstract The cyclic reduction algorithm is specialized to palindromic matrix polynomials and a complete analysis of applicability and convergence is provided. The resulting iteration is then related to other algorithms as the evaluation/interpolation at the roots of unity of a certain Laurent matrix polynomial, the trapezoidal rule for a certain integral and an algorithm based on the finite sections of a tridiagonal block Toeplitz matrix.

Keywords Cyclic Reduction, trapezoidal rule, Gauss-Chebyshev quadrature, matrix geometric mean, matrix sign, matrix square root, polar decomposition, Laurent matrix polynomial.

1 Introduction

The cyclic reduction (CR) algorithm, originally introduced for solving block tridiagonal system, has been successfully used to solve quadratic matrix equations and to compute the constant coefficient L_0 of the inverse $\mathcal{L}(z) = \sum_{i=-\infty}^{+\infty} L_i z^i$ of the Laurent matrix polynomial $\mathcal{H}(z) = Pz^{-1} + Q + Rz$, where P, Q and R are square matrices [8].

The matrix L_0 has some interesting properties and plays an important role in the computation of certain matrix functions. In fact, it has been proved in [15] that L_0 coincides with the square root, the matrix sign or the polar factor of a matrix A for suitable choices of Q and $P = R$ as functions of A. Also the geometric mean of two matrices A and B can be viewed as the matrix

Bruno Iannazzo
Dipartimento di Matematica e Informatica, Università di Perugia, Via Vanvitelli 1, 06123 Perugia, Italy. E-mail: bruno.iannazzo@dmi.unipg.it

Beatrice Meini
Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56100 Pisa, Italy. E-mail: meini@dm.unipi.it

coefficient L_0 for suitable choices of Q and P as functions of A and B. These properties hold for $P = Q$, that is, when the polynomial $\mathcal{H}(z)$ is *palindromic*.

Motivated by this interplay bewtween CR and the inverse of a palindromic Laurent polynomial, we specialize the cyclic reduction algorithm to the case of palindromic matrix polynomials. We refer to the algorithm obtained in this way as *palindromic cyclic reduction* (PCR).

We investigate the theoretical and computational properties of PCR and relate this algorithm to other apparently different algorithms like Newton's iteration, quadrature formulae, evaluation/interpolation at the roots of the unity, the finite sections algorithm, and averaging methods. This analysis provides a nice unification of several numerical methods, in the style of what has been done for the geometric mean of matrices in [12].

As a consequence, we find a scaling technique that much improves the speed of convergence of PCR, moreover, we prove necessary and sufficient conditions for applicability and convergence. This is an important result, since so far only sufficient conditions for convergence were known (see [8]) and no scaling technique was available. Moreover, a precise analysis of the convergence speed is performed, based on the location of the eigenvalues of a suitable matrix.

The paper is organized as follows. In Section 2 we recall CR and specialize it to the palindromic case, and we recall the interplay between matrix functions and PCR. In Section 3 we provide the different formulations of PCR and we show the connection with Newton's method for computing the matrix square root. In Section 4 the scaled version of PCR is presented. In Section 5 we give a complete convergence analysis for PCR, namely, we describe the set of initial values for which the sequences converge and provide the order and the rate of convergence in all cases. In Sections 6 and 7 we relate PCR with some other algorithms, namely quadrature formulae, evaluation/interpolation at the roots of unity, averaging techniques, and the finite sections algorithm for the solution of large linear systems. In Section 8 we show the numerical behavior of PCR and highlight the benefit of the scaling technique. Conclusions are drawn in Section 9.

2 Preliminaries

In this section we recall properties of CR and PCR, palindromic quadratic matrix equations, and the interplay between PCR and certain matrix functions.

2.1 Cyclic reduction and palindromic cyclic reduction

The cyclic reduction algorithm, in the form described in [4,7], generates the set of rational matrix iterations

$$\begin{cases} P_0 = P, \quad Q_0 = Q, \quad R_0 = R, \\ P_{k+1} = -P_k Q_k^{-1} P_k, \\ Q_{k+1} = Q_k - P_k Q_k^{-1} R_k - R_k Q_k^{-1} P_k, \qquad k = 0, 1, 2, \dots \\ R_{k+1} = -R_k Q_k^{-1} R_k, \end{cases} \tag{1}$$

where P, Q and R are given square matrices of the same size. If we define also the sequence $\widehat{Q}_{k+1} = \widehat{Q}_k - P_k Q_k^{-1} R_k$, with $\widehat{Q}_0 = Q$, and if the sequences in (1) are well defined and $\widehat{Q}_k$ is invertible for k sufficiently large, then a solution of the quadratic matrix equation

$$PX^2 + QX + R = 0, \tag{2}$$

can be obtained as

$$X = -\lim_{k\to\infty} \widehat{Q}_k^{-1} R. \tag{3}$$

Under some assumptions, verified in several applications, the matrix sequences generated by CR converge quadratically, but it is well known that for certain initial values CR can converge linearly or even diverge (see [8]). Nevertheless, a complete description of the set of all initial values for which the sequences (1) converge, together with the rate of convergence, is still an open problem.

Here we study CR (1) in the special case where $R = P$, i.e., equation (2) is the palindromic matrix equation $PX^2 + QX + P = 0$. For this reason, we will use the term Palindromic Cyclic Reduction (PCR) to denote the iteration (1) in the special case where $R = P$.

In PCR the sequence $\{R_k\}_k$ coincides with the sequence $\{P_k\}_k$ and equations (1) take the simpler form

$$\begin{cases} P_0 = P, \quad Q_0 = Q, \\ P_{k+1} = -P_k Q_k^{-1} P_k, \\ Q_{k+1} = Q_k - 2P_k Q_k^{-1} P_k, \qquad k = 0, 1, 2, \dots \end{cases} \tag{4}$$

2.2 Matrix functions

We recall the definition of some matrix functions; for a thorough treatise on this topic we refer the reader to the book [10].

Given a matrix $A \in \mathbb{C}^{n\times n}$ and a (possibly complex valued) function f, which is sufficiently regular, a definition of $f(A)$ can e given in terms of the Jordan canonical form of A, say $M^{-1}AM = J_1 \oplus \cdots \oplus J_s$, where J_i, for $i = 1, \dots, s$, is a Jordan block of size k_i corresponding to an eigenvalue λ_i (the λ_i's need not be distinct and $k_1 + \cdots + k_s = n$).

We set $f(A) = M(f(J_1) \oplus \cdots \oplus f(J_s))M^{-1}$, where

$$f(J_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{f^{(k_i-1)}(\lambda_i)}{(k_i-1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ 0 & & & f(\lambda_i) \end{bmatrix},$$

for $i = 1, \dots, s$, where k_i is the size of J_i. The definition makes sense if f is defined on λ_i for $i = 1, \dots, s$, and differentiable at λ_i when $k_i > 1$.

The *principal matrix square root* of A, denoted by the symbol $A^{1/2}$ is defined as $f(A)$, where, for $z \in \mathbb{C}\setminus(-\infty, 0)$, $f(z)$ is the principal square root of z, that is the square root of z with nonnegative real part. The definition makes sense if A has no negative real eigenvalues and semisimple null eigenvalues. In this case, the matrix $A^{1/2}$ is the unique solution of the matrix equation $X^2 = A$ whose eigenvalues belong to the right half complex plane.

Analogously, the *sign of a matrix* A having no purely imaginary eigenvalues, denoted by $\mathrm{sign}(A)$, is defined as $f(A)$, where, for no purely imaginary complex numbers,

$$f(z) = \begin{cases} 1 & \text{if } \mathrm{Re}(z) > 0, \\ -1 & \text{if } \mathrm{Re}(z) < 0. \end{cases}$$

The *unitary polar factor* of a nonsingular matrix A, denoted by $\mathrm{polar}(A)$, is the unique matrix U such that

$$A = UH, \quad U^*U = I,$$

where H is Hermitian positive definite. It holds that $\mathrm{polar}(A) = A(A^*A)^{-1/2}$ (see [10, Chapter 8]).

Given A and B Hermitian positive definite matrices, the matrix geometric mean of A and B is defined as

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2},$$

or, equivalently, $A\#B = A(A^{-1}B)^{1/2}$ (see [3, Chapter 4]).

We recall a useful lemma which is part of a general result on the relationship between the Jordan canonical form of A and the one of $f(A)$, whose complete statement and proof can be found for instance in [11] or [17, Thm. 9.4.7].

Lemma 1 *Let $A \in \mathbb{C}^{n\times n}$ and let $f(z)$ be such that $f(A)$ is well defined. Let λ be an eigenvalue of A such that $f(z)$ differentiable at λ. If $f'(\lambda) \neq 0$, then λ is a semisimple eigenvalue of A if and only if $f(\lambda)$ is a semisimple eigenvalue of $f(A)$.*

2.3 Palindromic matrix equations

Consider the palindromic quadratic matrix equation

$$PX^2 + QX + P = 0, \tag{5}$$

where P and Q are square matrices of the same size. When Q is nonsingular, it may be convenient to consider the equivalent matrix equation

$$MX^2 + X + M = 0, \tag{6}$$

where $M = Q^{-1}P$. The next theorem of [15] shows that just the location of the eigenvalues of $M = Q^{-1}P$ gives necessary and sufficient conditions for the existence of a unique solution X_* of (5) which is a function of M and with spectral radius at most one.

Theorem 1 *Let P, Q be two square matrices, with Q nonsingular, and set $M = Q^{-1}P$. The following conditions are equivalent:*

1. *The matrix $I - 4M^2$ admits a principal square root;*
2. *The matrix M has no real eigenvalues of modulus greater than $1/2$ and the real eigenvalues of modulus $1/2$ (if any) are semisimple;*
3. *The matrix equation (5) has a unique solution X_* which is a function of M and whose eigenvalues lie in the closed unit disk; moreover, its explicit expression is*
$$X_* = -2M(I + (I - 4M^2)^{1/2})^{-1}. \tag{7}$$

We recall a useful corollary of Theorem 1, proved in [15].

Corollary 2 *Let P, Q be two square matrices, with Q nonsingular, and set $M = Q^{-1}P$. The Laurent matrix polynomial*
$$\mathcal{H}(z) = Pz^{-1} + Q + Pz \tag{8}$$
is invertible in an open annulus containing the unit circle if and only if the matrix M has no real eigenvalues of modulus greater than or equal to $1/2$. In that case the invertibility domain of $\mathcal{H}(z)$ contains the annulus $\mathcal{A}_R = \{R < |z| < 1/R\}$, where $R = \rho(X_*)$, and X_* is defined in (7). Moreover, by setting $\mathcal{H}(z)^{-1} = \mathcal{L}(z) = \sum_{i=-\infty}^{+\infty} L_i z^i$ one has $L_i = L_{-i}$ for $i > 0$ and
$$L_0 = (I - 4M^2)^{-1/2}Q^{-1}. \tag{9}$$

2.4 PCR and matrix functions

As shown in [15], a nice property of PCR is that the limit of the sequence $\{Q_k\}_k$ of (4) can be put in correspondence with the matrix functions defined in Section 2.2. In fact, according to the type of the matrices P and Q, the sequence $\{Q_k\}_k$ is convergent and its limit Q_∞ can be the square root of a matrix A having no nonpositive real eigenvalues, the sign of a matrix A having no purely imaginary eigenvalues, the polar factor of an invertible matrix A and the geometric mean of two positive definite matrices A and B. More specifically, by setting $P = \frac{1}{4}(S - T)$ and $Q = \frac{1}{2}(S + T)$ in (4), Table 1 reports the limit Q_∞ corresponding to suitable choices of S and T, as shown in [15, Theorem 8].

Q_∞	S	T
$A^{1/2}$	I	A
$\mathrm{sign}(A)$	A	A^{-1}
$\mathrm{polar}(A)$	A	A^{-*}
$A\#B$	A	B

Table 1 The special values of Q_∞ for suitable choices of S and T, with $P = \frac{1}{4}(S - T)$ and $Q = \frac{1}{2}(S + T)$

3 Equivalent formulations of PCR

We show that PCR can be reformulated in different ways. This fact leads to a deeper understanding of PCR and to different algorithms for computing the same sequences.

Observe that, by setting $X_k = Q_k$ and $H_k = 2P_{k+1}$ in (4), PCR can be written as

$$\begin{cases} X_0 = Q, \quad H_0 = -2PQ^{-1}P, \\ X_{k+1} = X_k + H_k, \\ H_{k+1} = -\dfrac{1}{2}(H_k X_{k+1}^{-1} H_k), \qquad k = 0, 1, 2, \dots. \end{cases} \tag{10}$$

Recurrences (10) can be expressed in terms of the matrices X_k, $k \geqslant 0$, only. Indeed, we find that

$$\begin{aligned} X_{k+1} &= X_k + H_k = X_k - \frac{1}{2} H_{k-1} X_k^{-1} H_{k-1} \\ &= X_k - \frac{1}{2}(X_k - X_{k-1}) X_k^{-1} (X_k - X_{k-1}) = \frac{1}{2}(X_k + 2X_{k-1} - X_{k-1} X_k^{-1} X_{k-1}). \end{aligned}$$

The resulting iteration is the three-term recurrence

$$X_{k+1} = \frac{1}{2}(X_k + 2X_{k-1} - X_{k-1} X_k^{-1} X_{k-1}), \quad k = 1, 2, \dots, \tag{11}$$

starting with $X_0 = Q$, $X_1 = Q - 2PQ^{-1}P$.

The number of arithmetic operations to compute the sequence $\{X_k\}_k$ is the same as the number of operation needed in formulas (4); however, in this formulation, only a sequence of matrices is generated.

By setting $Y_k = X_{k-1} X_k^{-1}$, and by multiplying on the left by X_k^{-1} in (11), we find that

$$Y_{k+1} = \frac{1}{2}\left(I + 2Y_k^{-1} - (Y_k^{-1})^2\right), \quad k = 0, 1, 2, \dots, \tag{12}$$

starting with $Y_0 = I - 2(PQ^{-1})^2$. If we introduce also the matrix sequence $\{Z_k\}_k$ defined as

$$Z_{k+1} = Y_k Z_k, \quad k = 0, 1, 2, \dots, \tag{13}$$

with $Z_0 = Q$, we find that $Z_k = X_k$ for any $k \geqslant 0$. Therefore the sequences $\{Y_k\}_k$ and $\{Z_k\}_k$ defined in (12) and in (13), respectively, represent another formulation of PCR. Observe that the matrices Y_k, $k \geqslant 0$, commute each other and, in the case of convergence of PCR, converge to the identity matrix.

We point out that formulations (10) and (11) are similar to the ones given for the Newton method for the square root in [13] and for the averaging iteration for the geometric mean in [12], respectively; while formulation (12) is apparently new.

The expression (10) of PCR enables us to clarify the deep interplay between PCR and Newton's method. Indeed simplified Newton's method applied to

compute the principal square root of a matrix A consists in generating the sequence

$$Y_0 = I, \quad Y_{k+1} = \frac{1}{2}(Y_k + AY_k^{-1}) = \frac{1}{2}(Y_k + Y_k^{-1}A), \quad k \geqslant 0, \tag{14}$$

(see [10, Chap. 6] and the references therein). Now, set $M = Q^{-1}P$ and $A = I - 4M^2$. By using the "incremental" formulation given in [13] for the rational iteration (14), and by using the property that the matrices Y_k commute each other and with the matrix A, one can easily prove by induction on k that X_k is well defined if and only if Y_k is, and that $X_k = QY_k$ and $H_k = Q(Y_{k+1} - Y_k)$. Therefore, up to the matrix multiplication by Q, the sequences $\{X_k\}_k$ and $\{H_k\}_k$ in (10) coincide with the sequence generated by simplified Newton's method and with the Newton increment at each step, respectively.

This fact allows one to derive a scaling technique that accelerates the convergence and necessary and sufficient conditions for convergence of PCR, as shown in the following Sections 4 and 5.

4 Scaling technique

The equivalence between PCR (4) and Newton's iteration (14) allows us to introduce a scaling technique to accelerate the convergence of PCR .

In fact, the scaling technique has been introduced in [9] to accelerate the convergence of the iteration $X_{k+1} = \frac{1}{2}(X_k + X_k^{-1})$, which provides the matrix sign function of A, when $X_0 = A$. The scaling technique has been adapted to the incremental Newton method [13] when the matrix whose square root is required ($I - 4M^2$ in our case) is nonsingular. By applying this scaling technique to (14) and by using the equivalence with (4), we arrive at the following scaled version of PCR:

$$\begin{cases} X_0 = Q, \quad H_0 = -2PQ^{-1}P, \\ \gamma_k = \left| \dfrac{\det(X_k)^2}{\det(I - 4M^2)\det(Q)^2} \right|^{-1/(2n)}, \\ \widehat{X}_k = \gamma_k X_k, \qquad \widehat{H}_k = \gamma_k^{-1}(H_k + X_k/2) - \gamma_k X_k/2, \\ X_{k+1} = \widehat{X}_k + \widehat{H}_k, \qquad H_{k+1} = -\dfrac{1}{2}(\widehat{H}_k X_{k+1}^{-1} \widehat{H}_k), \qquad k = 0, 1, 2, \ldots. \end{cases}$$

The parameter γ_k is obtained at no additional cost during the inversion of X_k. Therefore the scaling technique does not increase the computational cost of each step of the algorithm, but in certain cases the convergence is dramatically accelerated.

5 Applicability and convergence properties of PCR

Two important issues of cyclic reduction are conditions for applicability and properties of convergence. We show that for PCR we can give specific and strong properties both for applicability and for convergence.

Concerning applicability conditions, by Theorem 13 of [5], PCR has no breakdown if and only if $T_{2^{k+1}-1}[P,Q,P]$ is nonsingular, for $k = 0,1,2,\ldots$, where $T_m[P,Q,P]$ is the leading $m \times m$ block submatrix of the infinite block tridiagonal block Toeplitz matrix

$$T[P,Q,P] = \begin{bmatrix} Q & P & & \\ P & Q & P & \\ & \ddots & \ddots & \ddots \end{bmatrix}.$$

For simplicity, we may consider the matrix $T_m[M,I,M]$, where $M = Q^{-1}P$. Using the orthogonal similarity $H_m \otimes I$, where $(H_m)_{ij} = \sqrt{\frac{2}{m+1}} \sin \frac{ij\pi}{m+1}$, $i,j = 1,\ldots,m$, is the matrix defining the sine transform, we find that the matrix $T_m[M,I,M]$ is similar to $I + V_m \otimes M$, where V_m is the $m \times m$ matrix

$$V_m = \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{bmatrix}, \tag{15}$$

having eigenvalues $2\cos\frac{j\pi}{m+1}$, for $j = 1,\ldots,m$.

Thus, the eigenvalues of $T_m[M,I,M]$ are $1 + 2\lambda_i \cos\frac{j\pi}{m+1}$, for $i = 1,\ldots,r$, $j = 1,\ldots,m$, where $\lambda_1,\ldots,\lambda_r$ are the eigenvalues of M. The matrix $T_{2^{k+1}-1}[P,Q,P]$ is singular for some $k > 0$ if and only if an eigenvalue of M belongs to the set $S := \bigcup_{k=1}^{\infty} S_{2^{k+1}-1}$, where

$$S_m = \left\{ z \in \mathbb{C} \,:\, z = -\frac{1}{2\cos(k\pi/(m+1))},\ k = 1,\ldots,m,\ k \neq \frac{m+1}{2} \right\}.$$

Observe that the set S is dense in the set $(-\infty,-1/2) \cup (1/2,+\infty)$. Therefore, if M has no real eigenvalues of modulus greater than $1/2$, then PCR has no breakdown.

The equivalence between PCR and Newton's method allows us to prove a stronger result than Theorem 13 of [5]. Indeed, it gives necessary and sufficient conditions for both applicability and convergence of PCR. This result is expressed by the following

Theorem 3 *Let P, Q be two square matrices, with Q nonsingular, and set $M = Q^{-1}P$. Then PCR is well defined and convergent if and only if the matrix M has no real eigenvalues of modulus greater than $1/2$ and the real eigenvalues of modulus $1/2$ (if any) are semisimple. In particular, if M has no real eigenvalues of modulus $1/2$ then PCR converges quadratically and $P_k \to 0$, $Q_k \to Q\psi(M)$, where $\psi(z) = (1-4z^2)^{1/2}$; if M has at least one semisimple eigenvalue of modulus $1/2$, then PCR converges linearly to the same limit.*

Proof By Theorem 1, $A = I - 4M^2$ has a principal square root if and only if the matrix M has no real eigenvalues of modulus greater than or equal to $1/2$ and the real eigenvalues of modulus $1/2$ (if any) are semisimple.

The proof is achieved by applying iteration (14) to $A = I - 4M^2$ and by observing that the sequence Y_k converges to $A^{1/2} = \psi(M)$, if and only if $Q_k = QY_k \to Q\psi(M)$ and $P_k = Q(Y_k - Y_{k-1}) \to 0$.

Iteration (14) is well defined and converges quadratically to $A^{1/2}$ if M has no real eigenvalues of modulus greater than $1/2$, while it converges linearly to $A^{1/2}$ if in addition M has some real eigenvalue of modulus $1/2$ (all of which are semisimple). This result is now customary and can be found in [10, Thm. 6.9–6.10].

Conversely, we use the fact that sequence (14) converges if and only if the sequence $Y_{k+1} = \frac{1}{2}(Y_k + J_i Y_k^{-1})$, $Y_0 = I$ converges for each Jordan block J_i of A. In particular a necessary condition for the convergence of sequence (14) is that the scalar sequence $y_{k+1} = \frac{1}{2}(y_k + \lambda_i y_k^{-1})$, $y_0 = 1$ converges for each eigenvalue λ_i of A (see [10,14]).

If M has a real eigenvalue of modulus $1/2$ corresponding to a nontrivial Jordan block, then by Theorem 1, A has a zero eigenvalue corresponding to a nontrivial Jordan block as well. Consider the sequence $Y_k = (y_{ij}^{(k)})$ obtained by iteration (14), where A is a nontrivial Jordan block relative to the zero eigenvalue; observe that $y_{22}^{(k)} = 1/2^k$, $y_{12}^{(k)} > 0$ and thus $y_{12}^{(k+1)} = \frac{1}{2}(y_{12}^{(k)} + (y_{22}^{(k)})^{-1}) \geqslant 2^{k-1}$. Therefore the matrix sequence $\{Y_k\}_k$ diverges.

If M has real eigenvalues outside the interval $[-1/2, 1/2]$ then the matrix A has a negative eigenvalue λ. The scalar iteration corresponding to λ does not converge, in fact the sequence obtained for $y_0 = 1$ is real or it is not well defined and cannot converge to the fixed points of the iteration which are $\pm\sqrt{-\lambda}\mathbf{i}$. Thus, the matrix iteration cannot converge.

Observe that the conditions of convergence of PCR stated by the above theorem are equivalent to any of the conditions of Theorem 1.

We wish to point out that Theorem 3 is an important result, since for general CR, no necessary and sufficient condition for the applicability is known, but only sufficient conditions for convergence have been found so far (see [5, 8]).

The matrix X_* of (7) can be directly computed by PCR, as stated by the following corollary of Theorem 3:

Corollary 4 *If PCR is convergent, then* $X_* = -2(Q+Q_\infty)^{-1}P$, *where* $Q_\infty = \lim_k Q_k$.

Proof It holds that

$$-2(Q+Q_\infty)^{-1}P = -2(Q+Q(I-4M^2)^{1/2})^{-1}P = -2(I+(I-4M^2)^{1/2})^{-1}M = X_*,$$

where the last equality follows from the fact that M commutes with any matrix function of itself [10, Thm. 1.13].

Observe also that $Q_\infty := \lim_k Q_k = Q_0 + 2\sum_{k=0}^{\infty} P_k$, therefore $\sum_{k=0}^{\infty} P_k = \frac{1}{2}(Q_\infty - Q)$. In particular, the sequence $\{\widehat{Q}_k\}_k$ is not needed in PCR in order to compute X_* by means of formula (3), since $\lim_k \widehat{Q}_k = Q_0 + \sum_{k=0}^{\infty} P_k = \frac{1}{2}(Q_\infty + Q)$. Indeed, the latter equality and formula (3) yield Corollary 4.

To conclude we give a precise estimate of the rate of convergence of PCR.

Theorem 5 *Assume that M has no real eigenvalues of modulus greater than $1/2$, and that the real eigenvalues of modulus $1/2$ (if any) are semisimple. The following properties hold:*

- *if M has no real eigenvalues of modulus $1/2$, then $\|P_k\| = O(\xi^{2^k})$, $\|Q_k - Q\psi(M)\| = O(\xi^{2^k})$, where ξ is any real number such that $\rho(X_*) < \xi < 1$, that is, convergence is at least quadratic;*
- *if M has at least one semisimple eigenvalue of modulus $1/2$, then $\|P_k\| = O(2^{-k})$ and $\|Q_k - Q\psi(M)\| = O(2^{-k})$, that is, convergence is at least linear.*

Proof The rate of convergence in case of quadratic convergence follows from Theorem 16 of [5]. The linear convergence of rate $1/2$ follows from Theorem 6.10 of [10].

6 Quadrature formulae, evaluation/interpolation and PCR

Assume that the matrix $M = Q^{-1}P$ has no real eigenvalues of modulus greater than or equal to $1/2$. According to Theorem 3, PCR is convergent and the inverse of the limit of the sequence $\{Q_k\}_k$ is $L_0 = (I - 4M^2)^{-1/2}Q^{-1}$, that is the constant coefficient of the inverse of the Laurent matrix polynomial $\mathcal{H}(z) = Pz^{-1} + Q + Pz$ given in (9). We show that the matrix L_0 can be obtained also by different methods, relying on an integral representation, and on the evaluation/interpolation of $\mathcal{L}(z) = \mathcal{H}(z)^{-1}$ at the roots of one. Surprisingly, we prove that the latter two approaches are strictly related to PCR.

We assume that the hypotheses of Corollary 2 hold, therefore the Laurent power series $\mathcal{L}(z)$ is convergent in the annulus $\mathcal{A}_R$ containing the unit circle.

In the sequel, we set $\omega_N = e^{\frac{2\pi \mathbf{i}}{N}}$, thus ω_N is a primitive N-th root of the unity.

6.1 Quadrature formulae

Since the function $\mathcal{L}(z) = \mathcal{H}(z)^{-1} = \sum_{i=-\infty}^{+\infty} L_i z^i$ is convergent in an annulus containing the unit circle, from a classic result on block Toeplitz matrices (see for instance [6, Theorem 3.1]), one has

$$L_0 = \frac{1}{2\pi}\int_0^{2\pi} \mathcal{L}(e^{\mathbf{i}\vartheta})d\vartheta.$$

Observe that $\mathcal{L}(e^{\mathbf{i}\vartheta}) = \mathcal{H}(e^{\mathbf{i}\vartheta})^{-1} = (Q + 2P\cos\vartheta)^{-1}$, therefore the above formula can be rewritten as

$$L_0 = \frac{1}{2\pi}\int_0^{2\pi} (Q + 2P\cos\vartheta)^{-1}\, d\vartheta, \tag{16}$$

and, by using the symmetry of the cosine we find that

$$L_0 = \frac{1}{\pi}\int_0^{\pi} (Q + 2P\cos\vartheta)^{-1} d\vartheta. \tag{17}$$

Setting $t = \cos\vartheta$ in (17), one obtains the formula

$$L_0 = \frac{1}{\pi}\int_{-1}^{1} \frac{(Q + 2Pt)^{-1}}{\sqrt{1-t^2}} dt. \tag{18}$$

The trapezoidal rule with $k+1$ nodes applied to (16) provides the approximation

$$W_k = \frac{1}{k}\sum_{j=0}^{k-1}\left(Q + 2P\cos\frac{2\pi j}{k}\right)^{-1}, \tag{19}$$

where we have used the periodicity of the integrand function. The integrand function $(Q + 2P\cos\vartheta)^{-1}$ is infinitely differentiable in $[0, 2\pi]$ and the odd ordered derivatives are periodic of period 2π. With these hypotheses a classical result in quadrature theory states that the order of convergence of the trapezoidal rule with $k+1$ nodes is greater than any power of h, where $h = 2\pi/k$ (see for instance Corollary 1 of Section 5.4 in [2]) and thus it is well suited for approximating the integral (16).

The trapezoidal rule on $k+1$ nodes applied to (17) leads to the formula

$$T_k = \frac{1}{k}\left[\frac{(Q+2P)^{-1}}{2} + \sum_{j=1}^{k-1}\left(Q + 2P\cos\frac{\pi j}{k}\right)^{-1} + \frac{(Q-2P)^{-1}}{2}\right]. \tag{20}$$

Again, the symmetry of the cosine function gives $T_k = W_{2k}$ where W_k is defined in (19), hence the quadrature formula (20) is a convenient way to get (19) for an even number of nodes.

Formula (18) is well suited for the Gauss-Chebyshev quadrature, whose formula for k nodes is

$$C_k = \frac{1}{k}\sum_{j=0}^{k-1}\left(Q + 2P\cos\frac{(2j+1)\pi}{2k}\right)^{-1}. \tag{21}$$

Observe that the nodes of the Gauss-Chebyshev quadrature are midpoints of the intervals chosen for the trapezoidal rule, so we obtain $T_{2k} = \frac{1}{2}(T_k + C_k)$. Therefore there is a nice explicit relationship between the trapezoidal rule and the Gauss-Chebyshev quadrature.

6.2 Evaluation/interpolation at the roots of one

We show how the block coefficient L_0 can be approximated by means of evaluation/interpolation at the roots of one.

Let N be an even positive integer, and let

$$P^{(N)}(z) = \sum_{j=-N/2}^{N/2-1} P_j^{(N)} z^j$$

be the Laurent matrix polynomial that verifies the interpolation conditions $P^{(N)}(\omega_N^i) = \mathcal{L}(\omega_N^i)$, for $i = 0, \dots, N-1$. Then one has (see for instance [6, Thm. 3.8])

$$L_0 = P_0^{(N)} - 2\sum_{j=1}^{\infty} L_{Nj}.$$

Therefore, the error obtained approximating L_0 by $P_0^{(N)}$ is

$$\left\|L_0 - P_0^{(N)}\right\| = 2\left\|\sum_{j=1}^{\infty} L_{Nj}\right\|,$$

where $\|\cdot\|$ is any matrix norm. Since $\mathcal{L}(z)$ is convergent in the annulus $\mathcal{A}_R = \{R < |z| < 1/R\}$, from the Cauchy integral formula (see for instance [6, Thm. 3.6]), the norm $\|L_i\|$ of the block coefficients converges to zero faster than ρ^i, for any $R < \rho < 1$, therefore there exists a constant c, independent of N, such that

$$\left\|L_0 - P_0^{(N)}\right\| \leqslant c\rho^N.$$

In other words, the sequence $\left\{P_0^{(N)}\right\}_N$ converges at least linearly to L_0, and the rate of convergence is bounded by the internal radius R of the invertibility annulus of $\mathcal{H}(z)$.

The coefficients of the Laurent matrix polynomial $P^{(N)}(z)$ in the order

$$P_0^{(N)}, P_1^{(N)}, \dots, P_{N/2-1}^{(N)}, P_{-N/2}^{(N)}, \dots, P_{-1}^{(N)},$$

can be obtained as Fu, where $F = \Omega \otimes I$, with Ω the Fourier matrix of size N, and $u_i = \mathcal{L}(\omega_N^i)$ for $i = 0, \dots, N-1$ (see [6, Sec. 3.1.2]). Thus, we may give an explicit expression to $P_0^{(N)}$ using the fact that the first row of the Fourier matrix is made of ones:

$$P_0^{(N)} = \frac{1}{N}\sum_{i=0}^{N-1} \mathcal{L}(\omega_N^i).$$

On the other hand

$$\mathcal{L}(\omega_N^i) = \mathcal{H}(\omega_N^i)^{-1} = \left(Q + 2P\cos\frac{2\pi i}{N}\right)^{-1},$$

therefore

$$P_0^{(N)} = \frac{1}{N}\sum_{i=0}^{N-1}\left(Q + 2P\cos\frac{2\pi i}{N}\right)^{-1}. \tag{22}$$

Observe that the latter formula coincides with (19) when $k = N$; therefore the trapezoidal rule on the integral (16) is equivalent to the evaluation/interpolation of $\mathcal{L}(z)$ at the roots of one.

6.3 Connections with PCR and the generalized averaging iteration

We show that the quadrature formulas derived in the previous section are strictly related with PCR and with the generalized averaging iteration [1,16].

Let us define the Laurent matrix polynomial

$$\mathcal{H}^{(k)}(z) = P_k z^{-1} + Q_k + P_k z, \quad k \geqslant 0, \tag{23}$$

where P_k and Q_k are defined in (4), and set

$$\mathcal{L}^{(k)}(z) = \mathcal{H}^{(k)}(z)^{-1}, \quad k \geqslant 0. \tag{24}$$

Since $\mathcal{H}^{(k)}(z)$ is palindromic, one has $Q_k = \mathcal{H}^{(k)}(\mathbf{i}) = \mathcal{H}^{(k)}(-\mathbf{i})$; therefore from (24) one obtains

$$Q_k = \mathcal{L}^{(k)}(\pm\mathbf{i})^{-1}. \tag{25}$$

By using the functional formulation of CR (see [8]), one has

$$\mathcal{L}^{(k+1)}(z^2) = \frac{1}{2}\left(\mathcal{L}^{(k)}(z) + \mathcal{L}^{(k)}(-z)\right). \tag{26}$$

One may easily verify by induction on k that the latter formula can be rewritten as

$$\mathcal{L}^{(k)}(z^{2^k}) = \frac{1}{2^k}\sum_{j=0}^{2^k-1}\mathcal{L}(\omega_{2^k}^j z).$$

Therefore, from the latter equation and from (25), it follows that

$$Q_k^{-1} = \frac{1}{2^k}\sum_{j=0}^{2^k-1}\mathcal{L}(\omega_{2^k}^j \zeta), \tag{27}$$

where ζ is such that $\zeta^{2^k} = \pm\mathbf{i}$, for instance, $\zeta = \omega_{2^{k+2}}$.

On the other hand, with $\zeta = \omega_{2^{k+2}}$

$$\mathcal{L}(\omega_{2^k}^j \zeta) = \left(Q + 2P\cos\frac{(4j+1)\pi}{2^{k+1}}\right)^{-1},$$

therefore

$$Q_k^{-1} = \frac{1}{2^k}\sum_{j=0}^{2^k-1}\left(Q + 2P\cos\frac{(4j+1)\pi}{2^{k+1}}\right)^{-1} = \frac{1}{2^k}\sum_{j=0}^{2^k-1}\left(Q + 2P\cos\frac{(2j+1)\pi}{2^{k+1}}\right)^{-1}, \tag{28}$$

where the latter equality is obtained by using the symmetry of the cosine function. Observe that the latter formula coincides with C_{2^k} given in (21); therefore the Gauss-Chebyshev quadrature on 2^k nodes of the integral (18) coincides with the inverse of the matrix Q_k of PCR. This connection between PCR and Gaussian quadrature has been already observed in the context of matrix geometric means in [12].

We have already observed that the Gaussian quadrature of (18) with k nodes and the trapezoidal rule applied to (17) with $k+1$ nodes lead to sequences C_k and T_k related through the relation $T_{2^{k+1}} = \frac{1}{2}(T_{2^k} + C_{2^k})$. Now we give another interesting relationship.

Proposition 6 *Let P and Q be square matrices such that Q is invertible and $M = Q^{-1}P$ has no real eigenvalues of modulus greater than or equal to $1/2$. Let T_k and C_k be the sequences defined in* (20) *and* (21)*, then, C_{2^k}, T_{2^k} and $C_{2^k}^{-1} + T_{2^k}^{-1}$ are invertible matrices and we have $C_{2^{k+1}} = 2(C_{2^k}^{-1} + T_{2^k}^{-1})^{-1}$. In other words, $C_{2^{k+1}}$ is the harmonic mean of T_{2^k} and C_{2^k}.*

Proof We know that $C_{2^k} = Q_k^{-1}$, $C_1 = Q^{-1}$, and

$$T_1 = \frac{1}{2}\big((Q+2P)^{-1} + (Q-2P)^{-1}\big) = A^{-1}Q^{-1},$$

where $A = I - 4M^2$ with $M = Q^{-1}P$, so C_{2^k} and T_1 are invertible.

We show by induction that $T_{2^k} = C_1 C_{2^k}^{-1} T_1 = T_1 C_{2^k}^{-1} C_1$, proving at the same time that T_{2^k} is invertible. The statement is true for $k = 0$, while for $k \geqslant 0$, the equivalence between Newton's method for the square root and PCR, proved in Section 3, where a recurrence for $Y_k = Q^{-1}Q_k$ is given, gives

$$\begin{aligned} C_1 C_{2^{k+1}}^{-1} T_1 = Q^{-1}Q_{k+1}T_1 = Y_{k+1}T_1 = \frac{1}{2}(Y_k + Y_k^{-1}A)T_1 = \frac{1}{2}(Q^{-1}Q_kT_1 + Q_k^{-1}QAT_1) \\ = \frac{1}{2}(C_1C_{2^k}^{-1}T_1 + C_{2^k}QAA^{-1}Q^{-1}) = \frac{1}{2}(T_{2^k} + C_{2^k}) = T_{2^{k+1}}, \end{aligned}$$

and thus $T_{2^{k+1}}$ is invertible. In a similar manner we can prove that $T_{2^k} = T_1 C_{2^k}^{-1} C_1$.

We use the property $T_{2^k} = C_1 C_{2^k}^{-1} T_1 = T_1 C_{2^k}^{-1} C_1$ together with the relation $T_{2^k} + C_{2^k} = 2T_{2^{k+1}}$ to complete the proof, in fact

$$C_{2^k}^{-1} + T_{2^k}^{-1} = C_1^{-1}T_{2^k}T_1^{-1} + C_1^{-1}C_{2^k}T_1^{-1} = 2C_1^{-1}T_{2^{k+1}}T_1^{-1} = 2C_{2^{k+1}}^{-1},$$

and we have $C_{2^{k+1}} = 2(C_{2^k}^{-1} + T_{2^k}^{-1})^{-1}$.

In summary, $T_{2^{k+1}}$ is the arithmetic mean of C_{2^k} and T_{2^k}, while $C_{2^{k+1}}$ is the harmonic mean of the two matrices. Both sequences converge quadratically to the same limit, that is L_0 (assuming that no real eigenvalue of $Q^{-1}P$ has modulus greater than or equal to $1/2$). The resulting iteration is

$$\begin{cases} A_1 = \frac{1}{2}\big((Q+2P)^{-1} + (Q-2P)^{-1}\big), \quad B_1 = Q^{-1}, \\ A_{k+1} = \frac{1}{2}(A_k + B_k), \\ B_{k+1} = 2A_k(A_k + B_k)^{-1}B_k, \quad k = 1, 2, \ldots, \end{cases}$$

which is a generalization of the averaging technique of Anderson and Trapp [1] (see also [16]), who observed that, starting from two definite positive matrices (or operators) A and B, and repeatedly computing the arithmetic and harmonic means, one generates a coupled iteration which converges to the geometric mean $A\#B$ (generalizing a classic result for scalars). Namely, for A and B posive definite, the sequence

$$\begin{cases} A_1 = A, \quad B_1 = B, \\ A_{k+1} = \dfrac{1}{2}(A_k + B_k), \\ B_{k+1} = 2A_k(A_k + B_k)^{-1}B_k, \quad k = 1, 2, \dots, \end{cases}$$

is such that A_k and B_k converge to $A\#B$.

7 A finite sections algorithm and PCR

We present an algorithm for computing the coefficient L_0 of the inverse of the Laurent matrix polynomial $\mathcal{H}(z) = Pz^{-1} + Q + Pz$ based on finite sections. Perhaps surprising, we will show that the algorithm produces the inverse of one of the sequences obtained by PCR.

Consider the $k \times k$ block tridiagonal matrix

$$T_k = \begin{bmatrix} Q & P & & \\ P & Q & \ddots & \\ & \ddots & \ddots & P \\ & & P & Q \end{bmatrix},$$

and consider the block linear system

$$T_{2k-1} \begin{bmatrix} X_1^{(k)} \\ X_2^{(k)} \\ \vdots \\ X_{2k-1}^{(k)} \end{bmatrix} = e_k \otimes I_n$$

where $X_j^{(k)}$, $j = 1, \dots, 2k-1$, are square matrices of the same size n as P and Q, and e_k is the k-th column of I_{2k-1}.

We aim to show that for the limit of the central term of the block vector $(X_j^{(k)})_j$ it holds that $L_0 = \lim_{k\to\infty} X_k^{(k)}$.

The matrix T_k can be written as

$$T_k = I_k \otimes Q + V_k \otimes P,$$

where V_k is the $k \times k$ matrix defined in (15). The eigenvalues and eigenvectors of V_k are explicitly known. In fact $D_k = H_k V_k H_k$, where H_k is the discrete sine transform matrix, i.e., the symmetric and orthogonal matrix whose j-th

column has entries $\sqrt{\frac{2}{k+1}}\sin\frac{ij\pi}{k+1}$, for $i = 1,\dots,k$, and D_k is the diagonal matrix whose diagonal elements are $2\cos\frac{j\pi}{k+1}$, for $j = 1,\dots,k$.

Therefore, T_k is similar to the block diagonal matrix

$$\mathcal{D}_k = (H_k^T \otimes I_n)T_k(H_k \otimes I_n) = I_k \otimes Q + D_k \otimes P.$$

Hence

$$\begin{aligned} X_k^{(k)} = (e_k^T \otimes I_n)T_{2k-1}^{-1}(e_k \otimes I_n) &= (e_k^T \otimes I_n)(H_{2k-1}^T \otimes I_n)\mathcal{D}_{2k-1}^{-1}(H_{2k-1} \otimes I_n)(e_k \otimes I_n) \\ &= (e_k^T H_{2k-1}^T \otimes I_n)\mathcal{D}_{2k-1}^{-1}(H_{2k-1}e_k \otimes I_n). \end{aligned}$$

Observe that $e_k^T H_{2k-1}^T = \frac{1}{\sqrt{k}}\left(\sin\frac{j\pi}{2}\right)_{j=1,2,\dots,2k-1}$, therefore

$$X_k^{(k)} = \frac{1}{k}\sum_{j=1}^{2k-1}\left(\sin\frac{j\pi}{2}\right)^2\left(Q + 2P\cos\frac{j\pi}{2k}\right)^{-1} = \frac{1}{k}\sum_{j=0}^{k-1}\left(Q + 2P\cos\frac{(2j+1)\pi}{2k}\right)^{-1}. \tag{29}$$

Formula (29) coincides with the Gauss-Chebyshev quadrature formula C_k in (21), and we have already observed in Section 6.3 that $C_{2^k} = Q_k^{-1}$. The convergence of $X_k^{(k)}$ can be proved using the convergence properties of the Gauss-Chebyshev quadrature.

8 Numerical experiments

We have performed some numerical experiments using MATLAB to compare different formulations of PCR and to give numerical evidence of the effectiveness of the scaling technique.

We have chosen A_ε, B_ε as $n \times n$ symmetric positive definite matrices, depending on $0 < \varepsilon < \frac{1}{2}$, and such that $M = Q^{-1}P$, where $P = \frac{1}{4}(B_\varepsilon - A_\varepsilon)$, $Q = \frac{1}{2}(A_\varepsilon + B_\varepsilon)$, has eigenvalues $\frac{1}{2} - \varepsilon, \frac{1}{3}, \dots, \frac{1}{n+1}$. In this way, the eigenvalues of M belong to the interval $(-\frac{1}{2}, \frac{1}{2})$ and, as ε tends to zero, an eigenvalue of M tends to $\frac{1}{2}$. With these hypotheses, the sequence Q_k converges to $A_\varepsilon \# B_\varepsilon$, the geometric mean of A_ε and B_ε. We have computed $A_\varepsilon \# B_\varepsilon$ with high precision by using the Symbolic Toolbox and we have compared the values obtained using the different algorithms: the plain PCR (4) (PCR), the incremental version (10), the three-term recurrence (3-term) implemented in the more stable form

$$X_{k+1} = \frac{1}{2}(X_k + X_{k-1} + X_{k-1}X_k^{-1}(X_{k-1} - X_{k-2})),$$

formula (12) and the scaled version of the PCR described in Section 4 (Scaled PCR). Tables 2 and 3 report, with $n = 10$ and , $\varepsilon = 10^{-5}$ and $\varepsilon = 10^{-10}$, respectively, the maximum of the component-wise relative error and the relative error in the spectral norm

$$\frac{\|\widetilde{Q} - A_\varepsilon \# B_\varepsilon\|}{\|A_\varepsilon \# B_\varepsilon\|}$$

	error (component-wise)	error (norm)
PCR	$5.7 \cdot 10^{-15}$	$4.7 \cdot 10^{-15}$
3-term	$6.9 \cdot 10^{-15}$	$5.5 \cdot 10^{-15}$
Scaled PCR	$3.0 \cdot 10^{-15}$	$2.6 \cdot 10^{-15}$

Table 2 $n = 10$ and $\varepsilon = 10^{-5}$

	error (component-wise)	error (norm)
PCR	$1.9 \cdot 10^{-12}$	$1.5 \cdot 10^{-12}$
3-term	$2.9 \cdot 10^{-12}$	$2.3 \cdot 10^{-12}$
Scaled PCR	$7.3 \cdot 10^{-13}$	$6.3 \cdot 10^{-13}$

Table 3 $n = 10$ and $\varepsilon = 10^{-10}$

where $\widetilde{Q}$ is the computed geometric mean when the iteration has reached numerical convergence. We have reported just the plain PCR, the three-term recurrence, and the scaled PCR, since the two other versions gave essentially the same results as plain PCR. The worst errors in the case $\varepsilon = 10^{-10}$ are due to a worst conditioning of the problem.

We observe that the algorithms are comparable in terms of accuracy. In some examples, not reported here, the plain PCR performs better than the three-term recurrence, while in other examples, the behavior is reversed. In any test we made, the scaled version gave the better accuracy (and faster convergence), so it should be preferred in practical computation.

Figure 1 reports the relative error of PCR and Scaled PCR at each iteration. We observe that the scaling does not influence too much the convergence in the case $\varepsilon = 10^{-5}$, while it improves a lot the convergence in the case $\varepsilon = 10^{-10}$.

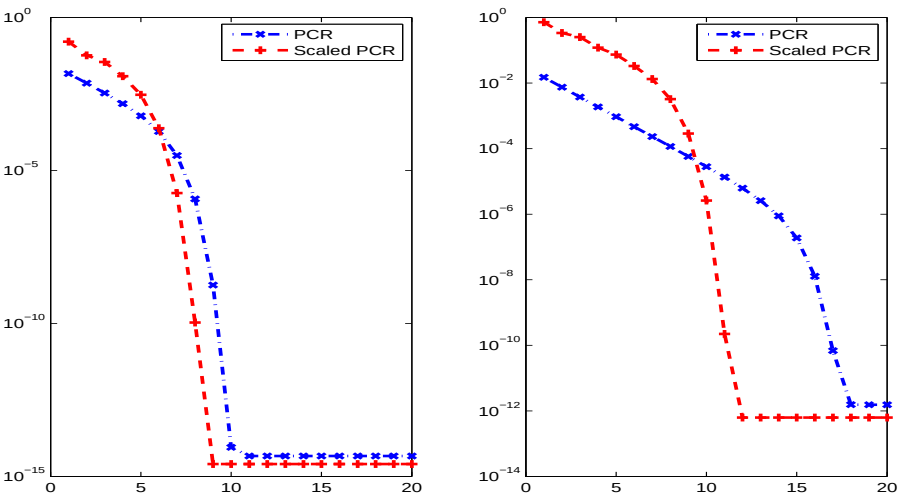

Fig. 1 Convergence behavior of PRC and Scaled PCR for the data matrices of Section 8 with $n = 10$ and $\varepsilon = 10^{-5}$ (left) and $\varepsilon = 10^{-10}$ (right)

9 Conclusions

We have related PCR to several algorithms and quadrature formulae, getting a nice unification of numerical methods, in the style of what has been done for the geometric mean of matrices in [12].

As a side result we have introduced a scaling technique and we have given necessary and sufficient conditions for applicability and convergence of PCR. Unfortunately, these results cannot be easily generalized to get a complete study of the convergence and a scaling for general CR, which remain yet important open problems. We hope that the results of this paper could help to give insights into the general problem.

References

1. W. N. Anderson and G. Trapp. Operator means and electrical networks. In *Proc. 1980 IEEE International Symposium on Circuits and Systems*, pages 523–527, 1980.
2. K. E. Atkinson. *An introduction to numerical analysis.* John Wiley & Sons Inc., New York, second edition, 1989.
3. R. Bhatia. *Positive definite matrices.* Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.
4. D. Bini and B. Meini. On the solution of a nonlinear matrix equation arising in queueing problems. *SIAM J. Matrix Anal. Appl.*, 17(4):906–926, 1996.
5. D. A. Bini, L. Gemignani, and B. Meini. Computations with infinite Toeplitz matrices and polynomials. *Linear Algebra Appl.*, 343/344:21–61, 2002.
6. D. A. Bini, G. Latouche, and B. Meini. *Numerical Methods for Structured Markov Chains.* Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005. Oxford Science Publications.
7. D. A. Bini and B. Meini. Effective methods for solving banded Toeplitz systems. *SIAM J. Matrix Anal. Appl.*, 20(3):700–719 (electronic), 1999.
8. D. A. Bini and B. Meini. The cyclic reduction algorithm: from Poisson equation to stochastic processes and beyond. In memoriam of Gene H. Golub. *Numer. Algorithms*, 51(1):23–60, 2009.
9. R. Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85:267–279, 1987.
10. N. J. Higham. *Functions of Matrices: Theory and Computation.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.
11. R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.
12. B. Iannazzo. The geometric mean of two matrices from a computational viewpoint. Available on arXiv - http://arxiv.org/abs/1201.0101.
13. B. Iannazzo. A note on computing the matrix square root. *Calcolo*, 40(4):273–283, 2003.
14. B. Iannazzo. A family of rational iterations and its application to the computation of the matrix pth root. *SIAM J. Matrix Anal. Appl.*, 30(4):1445–1462, 2008/09.
15. B. Iannazzo and B. Meini. Palindromic matrix polyomials, matrix functions and integral representations. *Linear Algebra Appl.*, 434(1):174–184, 2011.
16. F. Kubo and T. Ando. Means of positive linear operators. *Math. Ann.*, 246(3):205–224, 1979/80.
17. P. Lancaster and M. Tismenetsky. *The theory of matrices.* Computer Science and Applied Mathematics. Academic Press Inc., Orlando, FL, second edition, 1985.

Appendix. Technical results related to the symmetries of functions of cosine

We give the proofs of some equalities used in the paper, based on the symmetry of the cosine function.

Lemma 2 *Let f be a function continuous in the interval $[-1,1]$, then*

$$\frac{1}{2\pi}\int_0^{2\pi} f(\cos\vartheta)d\vartheta = \frac{1}{\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta.$$

Proof

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} f(\cos\vartheta)d\vartheta &= \frac{1}{2\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta + \frac{1}{2\pi}\int_{\pi}^{2\pi} f(\cos\vartheta)d\vartheta \\ &= \frac{1}{2\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta - \frac{1}{2\pi}\int_{\pi}^{0} f(2\pi-\cos\vartheta)d\vartheta \\ &= \frac{1}{2\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta + \frac{1}{2\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta = \frac{1}{\pi}\int_0^{\pi} f(\cos\vartheta)d\vartheta\end{aligned}$$

Lemma 3 *For $k \geqslant 1$ and W_k and T_k as in* (19) *and* (20)*, respectively it holds that $W_{2k} = T_k$.*

Proof The equality follows from

$$\begin{aligned}kW_{2k} &= \frac{1}{2}\sum_{j=0}^{2k-1}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1} \\ &= \frac{1}{2}(Q+2P)^{-1}+\frac{1}{2}\sum_{j=1}^{k}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1}+\frac{1}{2}(Q-2P)^{-1}+\frac{1}{2}\sum_{j=k+1}^{2k-1}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1} \\ &= \frac{1}{2}(Q+2P)^{-1}+2\frac{1}{2}\sum_{j=1}^{k}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1}+\frac{1}{2}(Q-2P)^{-1} = kT_k,\end{aligned}$$

where we have used

$$\sum_{j=k+1}^{2k-1}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1} = \sum_{j=1}^{k-1}\left(Q+2P\cos\frac{\pi(2k-j)}{k}\right)^{-1} = \sum_{j=1}^{k-1}\left(Q+2P\cos\frac{\pi j}{k}\right)^{-1}.$$

Lemma 4 *Let f be a function continuous in $[-1,1]$, then for $k \geqslant 0$ it holds that*

$$\sum_{j=0}^{2^k-1} f\left(\cos\frac{(2j+1)\pi}{2^{k+1}}\right) = \sum_{j=0}^{2^k-1} f\left(\cos\frac{(4j+1)\pi}{2^{k+1}}\right).$$

Proof

$$\begin{aligned}
\sum_{j=0}^{2^k-1} f\left(\cos\frac{(2j+1)\pi}{2^{k+1}}\right) &= \sum_{j=0}^{2^{k-1}-1} f\left(\cos\frac{(4j+1)\pi}{2^{k+1}}\right) + \sum_{j=0}^{2^{k-1}-1} f\left(\cos\frac{(4j+3)\pi}{2^{k+1}}\right)\\
&= \sum_{j=0}^{2^{k-1}-1} f\left(\cos\frac{(4j+1)\pi}{2^{k+1}}\right) + \sum_{j=2^k-1}^{2^{k-1}} f\left(\cos\frac{(2^{k+2}-4j-4+3)\pi}{2^{k+1}}\right)\\
&= \sum_{j=0}^{2^{k-1}-1} f\left(\cos\frac{(4j+1)\pi}{2^{k+1}}\right) + \sum_{j=2^{k-1}}^{2^k-1} f\left(\cos\left(2\pi-\frac{(4j+1)\pi}{2^{k+1}}\right)\right)\\
&= \sum_{j=0}^{2^k-1} f\left(\cos\frac{(4j+1)\pi}{2^{k+1}}\right).
\end{aligned}$$