SUPERLINEAR PCG METHODS
FOR SYMMETRIC TOEPLITZ SYSTEMS

STEFANO SERRA

ABSTRACT. In this paper we deal with the solution, by means of preconditioned conjugate gradient (PCG) methods, of $n \times n$ symmetric Toeplitz systems $A_n(f)\mathbf{x} = \mathbf{b}$ with nonnegative generating function f. Here the function f is assumed to be continuous and strictly positive, or is assumed to have isolated zeros of even order. In the first case we use as preconditioner the *natural* and the *optimal* τ approximation of $A_n(f)$ proposed by Bini and Di Benedetto, and we prove that the related PCG method has a superlinear rate of convergence and a total arithmetic cost of $O(n \log n)$ ops. Under the second hypothesis we cannot guarantee that the *natural* τ matrix is positive definite, while for the *optimal* we show that, in the ill-conditioned case, this can be really a bad choice. Consequently, we define a new τ matrix for preconditioning the given system; then, by applying the Sherman–Morrison–Woodbury inversion formula to the preconditioned system, we introduce a small, constant number of subsidiary systems which can be solved again by means of the previous PCG method. Finally, we perform some numerical experiments that show the effectiveness of the devised technique and the adherence with the theoretical analysis.

1. INTRODUCTION

In this paper, we discuss the solution of Toeplitz systems $A_n\mathbf{x} = \mathbf{b}$ by using the preconditioned conjugate gradient (PCG) method. Here we suppose that the Toeplitz matrix A_n is symmetric and is *generated* by a continuous 2π-periodic function f, defined in the fundamental interval $I = [-\pi, \pi]$ and periodically extended to the whole real axis, in the sense that the entries along the diagonals are given by the Fourier coefficients of the function f. Since we are interested in the symmetric case, we have to suppose that f is also an even function; more precisely we have

$$(1) \qquad [A_n]_{j,k} = a_{|j-k|} = \frac{1}{\pi} \int_0^\pi f(x) \cos((j-k)x)dx, \quad 0 \le j,k \le n-1.$$

We point out that the generating function f is given in some applications of Toeplitz systems. Classical examples are the kernels of the Wiener–Hopf equations [21], the spectral density functions in stationary stochastic processes [25, 24], and the point–spread functions in image processing [29].

When the generating function is continuous and positive there are many types of preconditioners [10, 14, 4, 18, 19, 5] chosen in the circulant algebra [16], in the τ

class [3], or in the Hartley one [5], where each matrix operation costs $O(n \log n)$ ops and $O(\log n)$ parallel steps [39, 6, 5]. These preconditioners lead to superlinearly convergent PCG methods [10, 14, 4, 5, 9] in the case where f is assumed to belong to the Wiener class, i.e., the previously defined Fourier coefficients give rise to an absolutely summable sequence. Under more restrictive hypotheses about the *regularity properties* of f, it is also shown that the related PCG methods have a quadratic rate of convergence [10, 14, 4, 5]. More recently, in the circulant and the Hartley case, the proof of superlinear convergence has been extended under the weakened assumption of continuity and positivity of f (see [12, 27]).

The first result we present here is this extension for the *optimal τ* preconditioner and a similar result for the *natural* circulant and τ preconditioners. The main tool is the Jackson characterization [26] of the Weierstrass approximation theorem like in the Chan, Yeung paper [12]. However, by counting the number of *outliers* [2] of the preconditioned matrices [12, 27], we observe that the spectral clustering around 1, achieved by the τ preconditioner, is slightly better.

When f has zeros, we know that the related matrices are asymptotically ill-conditioned [25, 32, 36]. Unfortunately, in this case, circulant and Hartley preconditioners fail; in [18], Di Benedetto has shown that the *natural* circulant preconditioner [10] is singular when, for instance, f vanishes in $x = 0$. Moreover, this author has proved that the *optimal* circulant preconditioner [14] and the *optimal* Hartley preconditioner [14] are positive definite, but the preconditioned matrices have a spectral cluster around 1 and also $O(n^\beta)$ eigenvalues in any open interval $(0, \epsilon)$ with β depending on the order of the zeros of f (for a richer analysis see [38, 20]). By virtue of the powerful convergence theory for PCG methods due, mainly, to Axelsson and Lindskog [2], it follows that the PCG methods related to these Hartley and circulant preconditioners are not competitive with the superfast methods [1, 17] and with the τ based or band-Toeplitz based PCG methods (see [19, 18, 7, 11, 37, 32, 36, 33]).

In this paper, under the nonnegative assumption, we propose a new τ preconditioner which joins the good computational features of the τ algebra and a strict approximation property with respect to the matrix $A_n(f)$. We obtain an algorithm which is based on the use of the Sherman–Morrison–Woodbury formula and on superlinear PCG methods, and which has an arithmetic cost of $O(n \log n)$ ops and $O(\log n)$ parallel steps. We observe that, for the case of f having zeros, the only superlinear PCG method previously devised was the one obtained by using band-Toeplitz preconditioners (see [37]).

The paper is organized as follows. In Section 2 we prove that the *optimal τ* preconditioner, introduced and discussed in [4], assures a superlinear convergence speed in the continuous, positive case. A similar analysis is carried out for the *natural* circulant and τ preconditioners. In Section 3, we introduce a new symmetric positive definite (SPD) τ preconditioner for which we deduce good approximation properties with respect to $A_n(f)$. In Section 4, we devise an algorithm associated to the former preconditioner and we discuss the related arithmetic and parallel costs. In Section 5, we perform some numerical experiments which plainly confirm the effectiveness of the proposed ideas, and we make a practical and theoretical comparison among the best strategies. Finally, in Section 6, we discuss some "peculiar" and apparently strange features of the PCG methods related to the *optimal* preconditioners.

2. THE τ PRECONDITIONERS

In this section we define the τ matrix algebra, and we introduce the *natural* and *optimal* τ preconditioners [4]. For any integer n, the $n \times n$ τ algebra is generated by the symmetric Toeplitz matrix whose first row is the second vector of the canonical base. Given a symmetric Toeplitz matrix A_n whose first row is $(a_0, a_1, \ldots, a_{n-1})$, the *natural* τ matrix $\tau(A_n)$ of A_n can be obtained by means of the Hankel correction:

$$(2) \qquad \tau(A_n) = A_n - \mathrm{HC}(A_n),$$

where $\mathrm{HC}(A_n)$ is the Hankel matrix whose first row is $(a_2, a_3, \ldots, a_{n-1}, 0, 0)$. On the other hand, the *optimal* approximation τ_{opt} of A_n is defined by minimizing the Frobenius distance from A_n over τ. The first row of τ_{opt} is defined by $(\phi_1, \phi_2, \ldots, \phi_n)$ with $\phi_j = \frac{n-j+3}{n+1}a_{j-1} - \frac{n-j-1}{n+1}a_{j+1}$, $j = 2, \ldots, n-2$ and $\phi_1 = a_0 - \frac{n-2}{n+1}a_2$, $\phi_{n-1} = \frac{4}{n+1}a_{n-2}$, $\phi_n = \frac{3}{n+1}a_{n-1}$ [4, 23].

From a spectral point of view it is possible to prove the following lemma.

Lemma 2.1. Let $A_n(f)$ be the symmetric Toeplitz matrix as defined in (1). By supposing that the function f is continuous and $m_f < M_f$ are the minimum and the maximum of f, respectively, the following relations hold:

1. The eigenvalues $\lambda_i(A_n(f))$ belong to (m_f, M_f) [25].
2. The eigenvalues of the *natural* τ approximation are generated by $s_n(f)\left(\frac{j\pi}{n+1}\right)$, where $s_n(f)$ is the n-th degree Fourier expansion of f [3, 4, 6] and $j = 1, \ldots, n$.
3. The eigenvalues of the *optimal* τ approximation are related to the Cesaro sums of f and are contained in the interval (m_f, M_f) [20].

As a simple consequence, we may obtain information about the extreme spectral properties of the matrix $\tau(A_n(f))$.

Definition 2.1. The class of continuous functions f for which the modulus of continuity $\omega(f, \delta)$ is $o(|\log \delta|^{-1})$ is the Dini–Lipschitz class and is denoted by $\mathcal{C}_*$. The complementary set is called the W class.

Corollary 2.1. If $f \in C_*$, then for any $\epsilon > 0$ there exists an integer value $\bar{n}$ such that $\forall n \geq \bar{n}$ the eigenvalues $\lambda_i(\tau(A_n(f)))$ belong to $(m_f - \epsilon, M_f + \epsilon)$. The same is true for the *optimal* τ matrix under the weaker assumption that $f \in C$ with $\epsilon = 0$ and $\bar{n} = 1$.

Proof. It follows plainly from the first part of Lemma 2.1 and from the convergence of the Fourier series of $f \in C_*$ to f (see [30, 40]). $\qquad \square$

Therefore, when the generating function f is strictly positive, by virtue of Corollary 2.1, we deduce that the *natural* τ approximation of $A_n(f)$ is asymptotically positive definite and the spectra of $A_n(f)$ and $\tau(A_n(f))$ are asymptotically the same. Under the nonnegative assumption, the situation is more complicated. The presence of points x_s in which the function f vanishes implies that we cannot guarantee the positive definiteness of $\tau(A_n(f))$ since we cannot guarantee the positiveness of the related Fourier expansion. In Section 5, we will present cases where the *natural* τ preconditioner is positive definite and cases where it is indefinite. Therefore, this observation justifies the search for a new τ preconditioner (see Section 3) when we consider the asymptotically ill-conditioned case, namely the case where f has zeros [25, 28, 34, 35].

Now, let us come back to the Toeplitz linear system with a positive continuous generating function. The aim of the following statements is the proof of a super-linear rate of convergence for the PCG method whose preconditioner is the *natural* τ correction or the *optimal* τ approximation of $A_n(f)$.

First we give a simple result about the global localization of the preconditioned matrices $P_\tau^1(f) = (\tau(A_n(f)))^{-1} A_n(f)$ and $P_\tau^{\text{opt}}(f) = (\tau_{\text{opt}})^{-1} A_n(f)$.

Lemma 2.2. If $f \in C_*$, then, for any $\epsilon > 0$, there exists an integer value $\bar{n}$ such that, $\forall n \geq \bar{n}$, we have that all the eigenvalues of $P_\tau^1(f)$ belong to $\left(\frac{m_f}{M_f} - \epsilon, \frac{M_f}{m_f} + \epsilon \right)$. The same is true for the $P_\tau^{\text{opt}}(f)$ under the weaker assumption that $f \in C$ with $\epsilon = 0$ and $\bar{n} = 1$.

Proof. It is a direct consequence of the first part of Lemma 2.1 and Corollary 2.1. □

The subsequent step is to show the *clustering property* around 1 (see [10, 4]) for the spectrum of $P_\tau^1(f)$. Let us denote the space of all n-th degree real trigonometric polynomials by $\mathcal{P}_n$, and all n-th degree cosine polynomials by $\mathcal{P}_n^c$. We have

$$\mathcal{P}_n = \left\{ \sum_{k=-n}^{n} b_k e^{\mathbf{i}kx} : b_k = \bar{b}_{-k}, \forall |k| \leq n \right\}, \quad \mathbf{i}^2 = -1,$$

and $\mathcal{P}_n^c$ is the real linear space described by $\mathcal{P}_n$, where the coefficients b_k are assumed to be real. Since $\mathcal{P}_n$ is a convex closed set of the Banach space of continuous periodic functions, it follows that the infimum $\inf_{p \in \mathcal{P}_n} \|f - p\|_\infty$ is really attained by a polynomial p^* of $\mathcal{P}_n$. In the proof of the main result (Theorem 2.1) we would like to associate p^* with a τ matrix. Unfortunately, the τ algebra is a real symmetric class of matrices related to even trigonometric sums (see part 2 of Lemma 2.1), that is, to cosine trigonometric polynomials. Therefore, we need to prove that p^* belongs to the linear real subspace $\mathcal{P}_n^c$ and this fact, actually, holds true in the light of the following elegant characterization due to Jackson [26].

Lemma 2.3. Let f be a continuous and 2π-periodic function. Suppose that, in addition, f is also even. Then for any positive n, there exists p^* belonging to $\mathcal{P}_n^c \subset \mathcal{P}_n$ such that $\|f - p^*\|_\infty \leq \omega\left(f; \frac{\pi}{n+1}\right)$.

Theorem 2.1. Let f be a C_* and 2π-periodic even function. Then, for any positive ϵ, there exist N and $M > 0$ such that for all $n > N$, the difference matrix $A_n(f) - \tau(A_n(f))$ has, at most, M eigenvalues outside $(-\epsilon, \epsilon)$. The same is true for the *optimal* τ matrix under the weaker assumption that $f \in C$.

Proof. For any $\epsilon > 0$, there exists $\delta > 0$ such that, when $|x_1 - x_2| < \delta$, one finds $|f(x_1) - f(x_2)| < \epsilon$. Let $K = \left\lceil \frac{\pi}{\delta} \right\rceil$; then $\omega\left(f; \frac{\pi}{K+1}\right) \leq \epsilon$. Hence, by Lemma 2.3, there is a cosine trigonometric polynomial p^* of degree K which verifies $\|f - p^*\|_\infty \leq \epsilon$. Therefore, by exploiting the linearity of the Fourier operator and of the τ transformation, the difference matrix $A_n(f) - \tau(A_n(f))$ may be viewed as

$$(3) \quad A_n(f) - \tau(A_n(f)) = A_n(f - p^*) - \tau(A_n(f - p^*)) + A_n(p^*) - \tau(A_n(p^*)).$$

Finally, by using Lemma 2.1, Corollary 2.1, and Lemma 2.3, we find that there exists N such that if $n > N$, the matrix $A_n(f - p^*) - \tau(A_n(f - p^*))$ has a Euclidean norm bounded by 3ϵ. Since the matrix $A_n(p^*) - \tau(A_n(p^*))$ has rank $2(K - 1)$, by virtue of equation (2) and of the Courant–Fischer minimax theorem [22], the thesis

follows. Concerning the matrix τ_{opt} the proof is identical and basically exploits the convergence of the Cesaro sum to f for any continuous function f (see also Theorem 1 of [12]). □

Observe that the proof of Theorem 2.1 is slightly simpler than the one given by Chan and Yeung [12] for the *optimal* circulant preconditioner C_{opt} [14]. Basically, the reason is that the preconditioner $\tau(A_n(f))$ has a simpler definition. Moreover, the *optimal* τ preconditioner (as the *optimal* circulant preconditioner) guarantees an approximation of $A_n(f)$ not only in the C_* class, but in the general class of continuous functions. However, it should be stressed that this is only an *academic* difference, since, in practice, the functions belonging to $W = C\backslash C_*$ are pathological examples to be defined *ad hoc*, like the Lebesgue one presented in [40]. By joining the previous remarks and results, we easily find Corollary 2.2.

Corollary 2.2. Let f_1 be a C_* positive and 2π-periodic even function and f_2 be a continuous positive and 2π-periodic even function. Then, for any positive ϵ, there exist N and $M > 0$ such that for all $n > N$, we have that, at most, M eigenvalues of the matrices $P_\tau^1(f_1) - I$ and $P_\tau^{opt}(f_2) - I$ have absolute value larger than ϵ.

Finally, for completeness, it is interesting to point out that the *natural* circulant preconditioner [10] assures clustering of the spectra of the preconditioned matrices under the same hypotheses used for the *natural* τ preconditioner. The proof is exactly the same as in Theorem 2.1.

Now, in the light of the analysis given in [10], we deduce that the related PCG methods, when applied to the preconditioned system

$$T^{-1}A_n(f)\mathbf{x} = T^{-1}\mathbf{b}, \qquad T = \tau(A_n(f)) \ \text{ or } \ T = \tau_{\mathrm{opt}},$$

converge superlinearly. In particular, the number of required iterations, to achieve a preassigned accuracy δ, is uniformly bounded by an absolute constant and, therefore, the total cost is dominated by $c(\delta, f)n\log n$ with a $c(\delta, f)$ positive constant value independent of the dimension n. This is the same result obtained by Chan and Yeung using circulant preconditioners and Jin using Hartley preconditioners. We observe that our choice is more restrictive, because we consider the *symmetric case* and not the *Hermitian case*. However, in the symmetric case considered we take advantage of using τ preconditioners, because they produce a better clustering (see also [4]) and are associated with the sine transform that has a cost practically equal to the Fourier or Hartley transforms [6].

3. The new τ preconditioner

When the generating function f is nonnegative and possesses zeros, the class of symmetric Toeplitz matrices $\{A_n(f)\}_n$ is asymptotically ill-conditioned. In this case, the classic iterative solvers, generally, fail in the sense that they do not converge or converge very slowly. More precisely, the rate of reduction of the error tends to 1 (no reduction of the error) as the dimension n tends to infinity. For this kind of problem, a good preconditioner is not only welcome but necessary. In the light of the second part of Lemma 2.1, it is trivial to conclude that we are not guaranteed the positive definiteness of the natural τ preconditioner. An example of this is given by $A_n(x^4)$: actually $\tau(A_n(f))$ is not positive definite, since, for instance, its first 8 minimal eigenvalues, for $n = 512$, have a magnitude of -10^{-4} (see also Section 6).

The alternative classical choice, i.e., the *optimal* τ preconditioner is guaranteed to be positive definite but it is not able to "match" the zeros of f and then is not suitable in order to obtain a proper cluster around 1 of the preconditioned spectrum (see Section 6).

The new SPD τ preconditioner is constructed as follows. Let g be the minimal polynomial matching all the zeros of f with their order (which is supposed to be even). It is worth noting that this minimal polynomial g is different from the one proposed and analysed in [7, 11, 37] owing to the requirement that g has to be an even function like f. Then we may view f as $g \cdot h$ where h is a strictly positive function, and we obtain the new preconditioner as $\tau(A_n(g))\tau(A_n(h))$. This matrix, at the present moment, is defined only theoretically while we are interested in computational algorithms. In the following scheme we propose how to calculate this preconditioner.

1. We suppose that we know the zeros of the even function f, namely $x_1, \ldots, x_s \in [0, \pi]$ and the related orders $2k_1, \ldots, 2k_s$. Observe that, if $\hat{x} \neq 0$, then the polynomial $g_{\hat{x}} = (2\cos x - 2\cos(\hat{x}))^2$ is nonnegative and is clearly the even polynomial of minimal degree which has a zero in $\hat{x}$ of order 2. If $\hat{x} = 0$ the minimal nonnegative even polynomial is trivially $g_0 = 2 - 2\cos x$. Consequently g is easily constructed as

$$g = \prod_{i=1}^{s} g_{\hat{x}_i}^{k_i}.$$

By using FFTs of constant order (depending on the number and the order of the zeros), it is not expensive ($O(l \log l)$ ops with l independent of n) to calculate the Fourier coefficients of g and, therefore, we have economically computed the matrix $A_n(g)$ and its τ correction $\tau(A_n(f))$.

2. Since we have to calculate the function h, in the sense of computing its Fourier coefficients, we employ the relation $f = g \cdot h$. We face the *inverse* problem of recovering these coefficients of h, by using the Fourier representation of the functions f and g. This problem can be solved in $O(n)$ ops by using a deconvolution algorithm [8], involving the vectors $\mathbf{x}_g$ and $\mathbf{x}_f$, where we have stored the Fourier coefficients of the considered functions. Notice that the matrix $\tau(A_n(h))$ is easily obtained from the entries of $A_n(h)$ by exploiting the simple equation (2).

Finally, we point out that $\tau(A_n(g))\tau(A_n(h))$ is symmetric since it is, by construction, a τ matrix and the τ algebra is structurally symmetric. In addition, as $\tau(A_n(g))$ and $\tau(A_n(h))$ are positive definite, the resulting product has positive eigenvalues and therefore it is an SPD matrix.

4. A NEW ALGORITHM IN THE ILL-CONDITIONED CASE

In this section we first perform an analysis of the "approximation properties" of the matrix $\tau(A_n(g))\tau(A_n(h))$ with regard to the problem matrix $A_n(f)$. Then, by exploiting some information from this analysis, we propose a new procedure to solve the given linear system.

Theorem 4.1. Let $\tau(2)$ be the preconditioner introduced in Section 3 and let $P_\tau^2(f) = (\tau(2))^{-1}A_n(f)$ be the second preconditioned matrix. Then $P_\tau^2(f)$ can be viewed as the sum of two contributions $P_\tau^1(h)$ and LR. Here, as h is strictly positive and continuous, $P_\tau^1(h)$ has a spectrum clustered around 1 (see Corollary 2.2), while LR is a matrix having a rank independent of n.

Proof. Preliminarily, we observe that, because of the relation $f = g \cdot h$ where g is polynomial of fixed degree, we find $A_n(f) = A_n(g)A_n(h) + LR_1$ with the rank of LR_1 depending only on the degree of g. Moreover, by (2), we have $\tau(A_n(g)) = A_n(g) + LR_2$, rank$LR_2 = $ constant and therefore by applying the Sherman–Morrison–Woodbury (SMW) formula (see equation (5)) we deduce

$$\tau(A_n(g))^{-1} = (I + LR_3)A_n^{-1}(g),$$

with rank$LR_3 = $ constant. Then, by indicating with LR_j "low" and constant rank matrices, the following relations hold true:

$$
\begin{aligned}
P_\tau^2(f) &= (\tau(2))^{-1}(A_n(g)A_n(h) + LR_1) \\
&= \tau(A_n(h))^{-1}(I + LR_3)A_n^{-1}(g)(A_n(g)A_n(h) + LR_1) \\
&= \tau(A_n(h))^{-1}(I + LR_3)(A_n(h) + LR_4) \\
&= \tau(A_n(h))^{-1}(A_n(h) + LR_5) \\
&= \tau(A_n(h))^{-1}A_n(h) + LR_6.
\end{aligned}
$$

The theorem is proved by putting $LR = LR_6$ and by observing that $P_\tau^1(h) = \tau(A_n(h))^{-1}A_n(h)$. $\qquad\square$

Unfortunately, the matrix LR is not symmetric and cannot be symmetrized by the same transformation that symmetrizes $P_\tau^1(h)$. Therefore, we cannot apply the Courant–Fischer minimax characterization [22] but we can deduce a proper clustering around 1 of the singular values and a superlinear rate of convergence for the related PCG method when applied to the normal equations.

An alternative possibility is given by the representation of $P_\tau^2(f)$ derived in Theorem 4.1. The original system $A_n(f)\mathbf{x} = \mathbf{b}$ is equivalent to

$$(4) \qquad (A_n(h) + LR_5)\mathbf{x} = \hat{\mathbf{b}}, \qquad \hat{\mathbf{b}} = (\tau(A_n(g)))^{-1}\mathbf{b},$$

where $\hat{\mathbf{b}}$ is computable sequentially in $O(n)$ ops by using a band-solver or in parallel in $O(\log n)$ steps by means of fast sine transforms. By following the relations proved in Theorem 4.1, it is also easy to verify that the matrices LR_i, $i = 1, \ldots, 6$, may be formally and numerically computed in $O(n \log n)$ ops and $O(\log n)$ parallel steps. Finally, the solution of the original system is obtained by applying the SMW inversion formula to the system (4). We recall that the SMW formula can be applied in the following way: given $B = A + LR$ and

$$
\begin{aligned}
LR &= X \cdot Y, \quad X, Y^T \in \mathbf{R}^{c \times n}, \\
c &= \operatorname{rank}(LR) = \text{constant}
\end{aligned}
$$

we obtain $\mathbf{x} = B^{-1}\mathbf{b}$ as

$$(5) \qquad B^{-1}\mathbf{b} = A^{-1}\mathbf{b} - A^{-1}X(I_c + YA^{-1}X)^{-1}YA^{-1}\mathbf{b}.$$

The cost of the calculation $\mathbf{x}$ is "reduced" to the solution of some systems with coefficient matrix A (which are assumed easy to solve) and to the inversion of the $c \times c$ matrix $I_c + YA^{-1}X$. In our case, we observe that the subsidiary systems having as coefficient matrix $A_n(h)$, h positive even continuous function, can be solved with the superlinear PCG method devised in [4] and extended to the continuous case in Section 2. Therefore the total cost is bounded by $c_1 n \log n$ ops and $c_2 \log n$ parallel steps where the coefficients c_j are independent of the dimension n and depend only on the functions f, g and on the preassigned accuracy δ.

TABLE 1. $f(x) = H^*(x)$, superlinear PCG

n	16	32	64	128	256	512
Iter (Prec=I)	8	15	20	31	45	74
Iter (Prec=$\tau(A_n(f))$)	6	7	8	9	10	9

TABLE 2. $f(x) = 1 - e^{-x^2}$

n	I	C_{opt}	$B_n^{\min}$	$B_{n,5}$	$\tau(1)$	$\tau(2)$
128	42	10	17	3	4	4
512	143	17	17	3	4	4

5. NUMERICAL EXPERIMENTS AND COMPARISONS

In order to show experimentally the validity of the analysis of Section 2, we need an even function f which is positive and continuous but not in the Wiener class. As pointed out in [12], a possible candidate is the Hardy–Littlewood series [40] given by

$$H(x) = \sum_{k=1}^{\infty} \left[\frac{e^{ik \log k}}{k} e^{ikx} + \frac{e^{-ik \log k}}{k} e^{-ikx} \right], \quad x \in [-\pi, \pi].$$

Unfortunately, this function is not positive and is not even. Consequently, we define $H^*(x)$ as $\frac{H(x)+H(-x)}{2} + 3.02$. This function is continuous, even, positive and does not belong to the Wiener class. For $n = 512$, the Euclidean condition number of $A_n(H^*)$ is equal to $2.37 \cdot 10^3$ and it increases as n grows. Table 1 shows the number of steps, in order to reach an accuracy of $\delta < 10^{-7}$, when the identity and our τ preconditioner is applied to the original system with $b_i = 1$ for any i.

Now, we consider the ill-conditioned case by considering nonnegative generating functions having zeros. We compare the convergence rate of the minimal band Toeplitz preconditioner [7], with the near optimal band Toeplitz preconditioner [37], with the optimal circulant preconditioner [14], with the *natural* τ preconditioner [4], and with our τ preconditioner for two different generating functions. They are $1 - e^{-x^2}$ and x^4 and are associated to ill-conditioned matrices A_n having Euclidean condition numbers equal to $O(n^2)$ and $O(n^4)$, respectively (see for instance [34, 35]). The matrices A_n are formed by evaluating the Fourier coefficients of the generating functions by using FFTs (see [11]). In the considered tests, the vector of all ones is the right-hand side vector, the zero vector is the initial guess, and the stopping criterion is $\|r_q\|_2 / \|r_0\|_2 \leq 10^{-7}$, where r_q is the residual vector after q iterations. All computations are done by using Matlab.

In Tables 2 and 3, I denotes that no preconditioning is used, C_{opt} is the T. Chan *optimal* circulant preconditioner [14], $B_{n,l}$ is the near optimal band-Toeplitz preconditioner [37] (l indicates the semibandwidth), $B_n^{\min}$ is the minimal band-Toeplitz preconditioner [7], $\tau(1)$ is the Bini–Di Benedetto preconditioner, and $\tau(2)$ is our τ preconditioner.

Notice the "strange" behavior of the *natural* τ preconditioner $\tau(1)$. In the case of $f(x) = 1 - e^{-x^2}$, it leads to a very fast PCG method, while, in the case of $f(x) = x^4$, the convergence is not so good and depends strongly on the size n of the problem. The reason is very simple: as proved in Corollary 2.1, the minimal eigenvalues can be less than m_f, which is equal to zero for the two considered functions. Fortunately,

TABLE 3. $f(x) = x^4$

n	I	C_{opt}	$B_n^{\min}$	$B_{n,5}$	$\tau(1)$	$\tau(2)$
128	587	77	24	11	22	8
512	7457	406	29	13	88	10

TABLE 4. Operation count for the i-th iteration

	$P = B_{n,5}$	$P = \tau(2)$
Multiplying $A_n(f)$ by a vector	$18n \log n + qn$	$18n \log n + qn$
Solution of $P\mathbf{y} = \mathbf{c}_i$	$2k_i n$	$4n \log n + n$
Other operations	$2n + k$	$2n + k$

for the first generating function, $\tau(1)$ is an SPD matrix, while in the second case, for $n = 128$, $\tau(1)$ has 3 eigenvalues in $[-2.4 \cdot 10^{-3}, -1.2 \cdot 10^{-3}]$ and, for $n = 512$, has 8 eigenvalues in $[-3.1 \cdot 10^{-5}, -1.0 \cdot 10^{-4}]$. Evidently, the presence of a "few" negative eigenvalues causes a dramatic slowdown in the convergence rate of the considered PCG method.

5.1. **Operation count.** First we observe a qualitative difference between the behavior of the PCG methods presented. The *optimal* circulant preconditioner leads to a sublinear PCG method, while $B_n^{\min}$ and $B_{n,5}$ are optimal preconditioners, and $\tau(2)$ is superoptimal. On the other hand, the behaviour of the PCG method associated with $\tau(1)$ strongly depends on the specific generating function. However, the operation count for each iteration is asymptotically the same ($O(n \log n)$), but it is better for the PCG methods related to the band Toeplitz preconditioners. For instance, the precise count for $\tau(2)$ and $B_{n,5}$ are shown in Table 4, where $k_i = 25$ for the first iteration (the *LU* factorization), and $k_i = 10$ otherwise [22].

Therefore, by taking into account the number of iterations, the PCG related to $B_{n,5}$ is slightly less expensive than the one related to $\tau(2)$. However, it should be stressed that the construction of $\tau(2)$ requires the Fourier coefficients of f and g, while the construction of $B_{n,5}$ needs also a precise evaluation of f and g in the Chebyshev points. Whence, in some cases it is easy to practically define $\tau(2)$ but we may have difficulties in defining $B_{n,5}$.

6. WHY ARE OPTIMAL PRECONDITIONERS SO BAD?

We did not make an explicit comparison with the *optimal* τ preconditioner, because, in the ill-conditioned case, even if it is always SPD, it does not have good approximation properties with respect to the extremes of the spectrum of $A_n(f)$. Let us consider the following elementary example: the finite difference discretization of the fourth derivative on the unit interval $[0, 1]$. The resulting matrix is the band Toeplitz operator generated by $p(x) = 16 \sin^4\left(\frac{x}{2}\right)$. By a direct calculation we obtain that the eigenvalues of the *optimal* circulant and τ approximations are

$$\lambda_i = p(x_i) + \frac{2}{n} p(x_i) + \frac{12 - 4\cos(x_i)}{n}, \qquad x_i = \frac{2(i-1)\pi}{n}, \quad i = 1, \dots, n$$

and

$$\lambda_i = p(z_i) + \frac{2}{n+1} \sin^2(z_i), \qquad z_i = \frac{i\pi}{n+1}, \quad i = 1, \dots, n,$$

TABLE 5. $f(x) = (2 - 2\cos(x))^2$

n	I	C_{opt}	τ_{opt}	H_{opt}	A_{opt}	$\tau(1) = \tau(2)$
32	18	14	10	22	14	2
128	162	31	16	55	31	2

respectively. It follows that $\forall g(n)$ increasing function of n with $\lim_{n \to \infty} \frac{g(n)}{n} = 0$, $\forall i = 1, \ldots, g(n)$, $\lambda_i(A_n(p))$ and $\lambda_i(G)$ (G the *optimal* τ or circulant or anticirculant or Hartley preconditioner) behave in a very different way when n tends to infinity. In particular, for i constant wrt n, $\lambda_i(A_n(p)) \sim 1/n^4$ and $\lambda_i(G) \sim 1/n^s$ with $s = 1$ in the circulant case and $s = 3$ in the τ case. Then the condition number of the preconditioned matrices grows, at least, as n^{4-s}. So while the *natural* τ preconditioner can be good or bad (see Tables 2 and 3), the *optimal* preconditioners, in the presence of zeros of high orders, are generally **bad**. For evidence of this, see [23] and Tables 2, 3, and 5, where A_{opt} and H_{opt} indicate the optimal anticirculant and Hartley preconditioners.

By generalizing the argument used for Table 5, we can say that the "averaging schemes", which are inherent to minimization in the Frobenius norm, destroy the information about the infinitesimal order of the small eigenvalues and this is the main reason for the observed slowdown of the related PCG methods. However, for symmetric Toeplitz problems, the optimal τ approach is the best, among the optimals, owing also to the structural symmetry of the τ algebra.

REFERENCES

[1] G. Ammar and W. Gragg, "Superfast solution of real positive definite Toeplitz systems", *SIAM J. Matrix Anal. Appl.*, 9, pp. 61–76 (1988). MR **89b:**65065

[2] O. Axelsson and G. Lindskog, "The rate of convergence of the preconditioned conjugate gradient method", *Numer. Math.*, 48, pp. 499–523 (1986). MR **88a:**65037b

[3] D. Bini and M. Capovani, "Spectral and computational properties of band symmetric Toeplitz matrices", *Linear Algebra Appl.*, 52/53, pp. 99–126 (1983). MR **85k:**15008

[4] D. Bini and F. Di Benedetto, "A new preconditioner for the parallel solution of positive definite Toeplitz systems", *Proc. 2nd ACM SPAA*, Crete, Greece, July 1990, pp. 220–223.

[5] D. Bini and P. Favati, "On a matrix algebra related to the discrete Hartley transform", *SIAM J. Matrix Anal. Appl.*, 14, pp. 500–507 (1993). MR **94h:**65026

[6] E. Bozzo, *Matrix Algebras and Discrete Transforms*. PhD. thesis in Comp. Sci., Dept. Comp. Sci., University of Pisa, Italy, 1994.

[7] R.H. Chan, "Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions", *IMA J. Numer. Anal.*, 11, pp. 333–345 (1991).

[8] R.H. Chan and W.-K. Ching, "Toeplitz-circulant preconditioners for Toeplitz systems and their applications to queueing networks with batch arrivals", *SIAM J. Sci. Comput.*, 17, pp. 762–772 (1996). MR **97b:**65054

[9] R.H. Chan and M.K. Ng, "Conjugate gradient method for Toeplitz systems", *SIAM Rev.*, 38, pp. 427–482, (1996). MR **97i:**65048

[10] R.H. Chan and G. Strang, "Toeplitz equations by conjugate gradients with circulant preconditioner", *SIAM J. Sci. Statist. Comput.*, 10, pp. 104–119 (1989). MR **90d:**65069

[11] R.H. Chan and P. Tang, "Fast band-Toeplitz preconditioners for Hermitian Toeplitz systems", *SIAM J. Sci. Comput.*, 15, pp. 164–171 (1994). MR **94j:**65043

[12] R.H. Chan and M.-C. Yeung, "Circulant preconditioners for Toeplitz matrices with positive continuous generating functions", *Math. Comp.*, 58, pp. 233–240 (1992). MR **92e:**65040

[13] R.H. Chan and M.-C. Yeung, "Jackson's theorem and circulant preconditioned Toeplitz systems", *J. Approx. Theory*, 70, pp. 191–205 (1992). MR **93c:**65046

[14] T.F. Chan, "An optimal circulant preconditioner for Toeplitz systems", *SIAM J. Sci. Statist. Comput.*, 9, pp. 766–771 (1988). MR **89e:**65046

[15] E. Cheney, *Introduction to Approximation Theory.* McGraw–Hill, New York, 1966. MR **36:**5568

[16] P. Davis, *Circulant Matrices.* John Wiley and Sons, New York, 1979. MR **81a:**15003

[17] F. de Hoog, "A new algorithm for solving Toeplitz systems of equations", *Linear Algebra Appl.*, 88/89, pp. 123–138 (1987). MR **88d:**65053

[18] F. Di Benedetto, "Analysis of preconditioning techniques for ill-conditioned Toeplitz matrices", *SIAM J. Sci. Comput.*, 16, pp. 682–697 (1995). MR **95m:**65082

[19] F. Di Benedetto, G. Fiorentino and S. Serra, "C.G. preconditioning for Toeplitz matrices", *Comput. Math. Appl.*, 25, pp. 35–45 (1993). MR **93h:**65063

[20] F. Di Benedetto and S. Serra Capizzano, "A unifying approach to matrix algebra preconditioning", Numer. Math., in press.

[21] I. Gohberg and I. Fel'dman, *Convolution Equations and Projection Methods for Their Solution*, Transl. Math. Monogr., **41**, Amer. Math. Soc., Providence, RI, 1974. MR **50:**8149

[22] G.H. Golub and C.F. Van Loan, *Matrix Computations.* The Johns Hopkins University Press, Baltimore, MD, 1983. MR **85h:**65063

[23] B. Grasso, *Tecniche di Precondizionamento per la Risoluzione Numerica di Sistemi Lineari Tramite Proiezioni su Particolari Algebre Matriciali.* Undergraduate thesis, Dept. Math., University of Genova, Italy, 1995.

[24] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series.* Second edition, Chelsea, New York, 1984. MR **88f:**62003

[25] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications.* Second Edition, Chelsea, New York, 1984. MR **88b:**42031

[26] D. Jackson, *The Theory of Approximation.* Amer. Math. Soc., New York, 1930.

[27] X.-Q. Jin, "Hartley preconditioners for Toeplitz systems generated by positive continuous functions", *BIT*, 34, pp. 367–371 (1994). MR **97i:**65050

[28] M. Kac, W. Murdock, and G. Szegö, "On the eigenvalues of certain Hermitian forms", *J. Rational Mech. Anal.*, 2, pp. 767–800 (1953). MR **15:**538b

[29] A. Oppenheim, *Applications of Digital Signal Processing.* Prentice–Hall, Englewood Cliffs, NJ, 1978.

[30] W. Rudin, *Principles of Mathematical Analysis.* McGraw–Hill, New York, 1985. MR **52:**5893

[31] S. Serra, "Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems", *BIT*, 34, pp. 579–594 (1994). MR **98a:**65052

[32] S. Serra, "Conditioning and solution of Hermitian (block) Toeplitz systems by means of preconditioned conjugate gradient methods", *Proc. in Advanced Signal Processing Algorithms, Architectures, and Implementations - SPIE Conference*, F. Luk ed., San Diego, CA, July 1995, pp. 326–337.

[33] S. Serra, "Preconditioning strategies for Hermitian Toeplitz systems with nondefinite generating functions", *SIAM J. Matrix Anal. Appl.*, 17, pp. 1007–1019 (1996). MR **98a:**65042

[34] S. Serra, "On the extreme eigenvalues of Hermitian (block) Toeplitz matrices", *Linear Algebra Appl.*, 270, pp. 109–129 (1998). CMP 98:05

[35] S. Serra, "On the extreme spectral properties of Toeplitz matrices generated by L^1 functions with several minima/maxima", *BIT*, 36, pp. 135–142 (1996). MR **98a:**65051

[36] S. Serra, "Asymptotic results on the spectra of preconditioned Toeplitz matrices and some applications", *TR nr. 9, LAN - Dept. Math., University of Calabria,* (1995).

[37] S. Serra, "Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems", *Math. Comp.*, 66, pp. 651–665 (1997). MR **97h:**65056

[38] E.E. Tyrtyshnikov, "Circulant preconditioners with unbounded inverses", *Linear Algebra Appl.*, 216, pp. 1–23 (1995). MR **95m:**15019

[39] C.F. Van Loan, *Computational Frameworks for the Fast Fourier Transform.* SIAM, Philadelphia, PA, 1992. MR **93a:**65186

[40] A. Zygmund, *Trigonometric Series.* Cambridge University Press, London, 1959. MR **21:**6498

Dipartimento di Informatica, Corso Italia 40, 56100 Pisa (ITALY)

E-mail address: serra@mail.dm.unipi.it