NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2001; **08**:83–98 *Prepared using nlaauth.cls [Version: 2000/03/22 v1.0]*

Preconditioners for non-Hermitian Toeplitz systems†

Raymond H. Chan, Daniel Potts, Gabriele Steidl

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong ,
rchan@math.cuhk.edu.hk
Medical University of Lübeck, Institute of Mathematics, Wallstr. 40, D–23560 Lübeck,
potts@math.mu-luebeck.de
University of Mannheim, Institute of Mathematics, D–68131 Mannheim,
steidl@math.uni-mannheim.de

SUMMARY

In this paper, we construct new ω-circulant preconditioners for non-Hermitian Toeplitz systems, where we allow the generating function of the sequence of Toeplitz matrices to have zeros on the unit circle. We prove that the eigenvalues of the preconditioned normal equation are clustered at 1 and that for (N, N)-Toeplitz matrices with spectral condition number $\mathcal{O}(N^\alpha)$ the corresponding PCG method requires at most $\mathcal{O}(N \log^2 N)$ arithmetical operations. If the generating function of the Toeplitz sequence is a rational function then we show that our preconditioned original equation has only a fixed number of eigenvalues which are not equal to 1 such that preconditioned GMRES needs only a constant number of iteration steps independent of the dimension of the problem.
Numerical tests are presented with PCG applied to the normal equation, GMRES, CGS and BICGSTAB. In particular, we apply our preconditioners to compute the stationary probability distribution vector of Markovian queuing models with batch arrival.

KEY WORDS: Non-Hermitian Toeplitz matrices, circulant matrices, Krylov space methods, CG–method, preconditioner,*Mathematics Subject Classification.* 65F10, 65F15, 65T10.

1. INTRODUCTION

Let $C_{2\pi}$ denote the space of 2π-periodic continuous functions and let $\mathcal{W}$ be the Wiener algebra.
We are concerned with the solution of non-Hermitian Toeplitz systems

$$\boldsymbol{A}_N(f)\boldsymbol{x} = \boldsymbol{b} \tag{1.1}$$

arising from a generating function $f \in \mathcal{W}$ with a finite number of zeros, i.e.

$$\boldsymbol{A}_N(f) := (a_{j-k}(f))_{j,k=0}^{N-1} \ ,$$

†Research supported in part by the Hong Kong-German Joint Research Collaboration Grant from the Deutscher Akademischer Austauschdienst and the Hong Kong Research Grants Council.

Received 20 September 1999
Revised 7 August 2000
Accepted 27 August 2000

$$a_k(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{\mathrm{i}kt} \, \mathrm{d}t \,.$$

We are interested in iterative solution methods, more precisely in Krylov space methods. These methods require in each iteration step only multiplications of $\boldsymbol{A}_N(f)$ with vectors and since $\boldsymbol{A}_N(f)$ is Toeplitz these multiplications can be computed in $\mathcal{O}(N \log N)$ arithmetical operations by using fast Fourier transforms. However, in order to keep the number of iterations small, iterative methods must be applied with suitable preconditioning, in general. The construction of "good" preconditioners is the aim of this paper.

Although there exists a rich literature on Hermitian Toeplitz systems (see [4] and the references therein), only few papers consider the non-Hermitian case [3, 5, 6, 9, 12].

The papers [5, 6, 9, 12] construct preconditioners only for generating functions f with strictly positive absolute value, i.e. $|f| \geq c > 0$. In [5], the authors suggested the use of optimal circulant preconditioners $\boldsymbol{M}_N$ of $\boldsymbol{A}_N(f)$ and to solve the normal equation

$$(\boldsymbol{M}_N^{-1} \boldsymbol{A}_N(f))^* \, \boldsymbol{M}_N^{-1} \boldsymbol{A}_N(f) \, \boldsymbol{x} = (\boldsymbol{M}_N^{-1} \boldsymbol{A}_N(f))^* \, \boldsymbol{M}_N^{-1} \, \boldsymbol{b}$$

by the CG method. Here $\boldsymbol{A}^*$ denotes the transposed complex conjugate matrix of $\boldsymbol{A}$. In [6] and [12], τ-preconditioners depending on $|f|^2$ and optimal trigonometric preconditioners of $\boldsymbol{A}_N(f)^* \boldsymbol{A}_N(f)$, respectively, were proposed for preconditioning of the normal equation

$$\boldsymbol{A}_N^*(f) \boldsymbol{A}_N(f) \boldsymbol{x} = \boldsymbol{A}_N^*(f) \boldsymbol{b} \,. \tag{1.2}$$

The paper [9] contains interesting results for Strang preconditioners and for four other special preconditioners for the case that the generating function $f > 0$ is a rational function.

The only paper, where generating functions with zeros were considered, is [3]. Here the authors suggested preconditioners $\boldsymbol{M}_N$ which are products of banded Toeplitz matrices and optimal circulant matrices and examine the distribution of the singular values of the preconditioned matrices $\boldsymbol{M}_N^{-1} \boldsymbol{A}_N(f)$. Unfortunately, it is not clear how the number of iteration steps of the preconditioned CGS method [16] used in their numerical tests depends on this distribution of the singular values.

In this paper, we introduce circulant, respectively ω-circulant preconditioners related to $|f|^2$ for the normal equation (1.2) even if $f \in \mathcal{W}$ has zeros. These preconditioners can be applied with a fewer amount of arithmetical operations than the combined preconditioners in [3]. We show that the singular values of the preconditioned matrix are clustered at 1 and that in case if the spectral condition number of $\boldsymbol{A}_N(f)$ is $\mathcal{O}(N^\alpha)$ the PCG method applied to (1.2) converges in $\mathcal{O}(\log N)$ iteration steps.

We are also interested in Krylov space methods like GMRES or BICGSTAB which do not require the translation of (1.1) to the normal equation. Here we suggest circulant, respectively ω-circulant preconditioners related to f. Unfortunately, the convergence of these methods does no longer depend on the singular values but on the eigenvalues of the preconditioned system. We can not prove clustering results for arbitrary generating functions $f \in \mathcal{W}$. However, for rational functions f, we show that the preconditioned matrices have only a finite (independent of N) number of eigenvalues which are not equal to 1 such that preconditioned GMRES converges in a finite number of steps independent of the dimension of the problem.

This paper is organized as follows: In Section 2, we introduce our circulant, respectively ω-circulant preconditioners and prove corresponding clustering results. Section 3 modifies these results for trigonometric preconditioners. Finally, Section 4 contains numerical examples for various iterative methods and preconditioners. In particular, we apply our preconditioners to the queueing network problem with batch arrivals examined in [3].

2. ω–CIRCULANT PRECONDITIONS

In order to prove our main result in Theorem 2.5 we need a couple of preliminary lemmata. In the following we denote by $\boldsymbol{R}_N(m)$ arbitrary (N,N)–matrices of rank at most m.

Lemma 2.1. *Let $f \in \mathcal{W}$. Then, for any $\varepsilon > 0$ and N sufficiently large, there exists $M = M(\varepsilon)$ independent of N such that*

$$\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f) = \boldsymbol{A}_N(|f|^2) + \boldsymbol{U}_N + \boldsymbol{R}_N(M),$$

where $\|\boldsymbol{U}_N\|_2 \leq \varepsilon$.

To be careful with respect to our setting, we notice the short proof which in general follows the lines of [6].

Proof. Let

$$\begin{aligned}\boldsymbol{a}_+ &:= (a_0, a_1, \dots, a_{N-1})' \quad , \quad \boldsymbol{a}_- := (0, \bar{a}_{-1}, \dots, \bar{a}_{-(N-1)})',\\ \tilde{\boldsymbol{a}}_+ &:= (0, \bar{a}_{N-1}, \dots, \bar{a}_1)' \quad , \quad \tilde{\boldsymbol{a}}_- := (0, a_{-(N-1)}, \dots, a_{-1})',\end{aligned}$$

where $\boldsymbol{y}'$ denotes the transposed vector of $\boldsymbol{y}$ and let $\boldsymbol{L}(\boldsymbol{y})$ be the lower triangular Toeplitz matrix with first column $\boldsymbol{y}$. Then $\boldsymbol{A}_N(f) = \boldsymbol{L}(\boldsymbol{a}_+) + \boldsymbol{L}^*(\boldsymbol{a}_-)$ and we obtain by straightforward computation that

$$\begin{aligned}\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f) &= \boldsymbol{L}^*(\boldsymbol{a}_+)\boldsymbol{L}^*(\boldsymbol{a}_-) + \boldsymbol{L}(\boldsymbol{a}_-)\boldsymbol{L}(\boldsymbol{a}_+) + \boldsymbol{L}^*(\boldsymbol{a}_+)\boldsymbol{L}(\boldsymbol{a}_+) + \boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{L}^*(\tilde{\boldsymbol{a}}_+)\\ &\quad + \boldsymbol{L}(\boldsymbol{a}_-)\boldsymbol{L}^*(\boldsymbol{a}_-) + \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{L}(\tilde{\boldsymbol{a}}_-) - \boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{L}^*(\tilde{\boldsymbol{a}}_+) - \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{L}(\tilde{\boldsymbol{a}}_-)\\ &= \boldsymbol{A}_N(|S_N f|^2) - \boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{L}^*(\tilde{\boldsymbol{a}}_+) - \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{L}(\tilde{\boldsymbol{a}}_-)\\ &= \boldsymbol{A}_N(|f|^2) + \boldsymbol{A}_N(|S_N f|^2 - |f^2|) - \boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{L}^*(\tilde{\boldsymbol{a}}_+) - \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{L}(\tilde{\boldsymbol{a}}_-), \quad (2.3)\end{aligned}$$

where

$$(S_N f)(t) := \sum_{k=-(N-1)}^{N-1} a_k \mathrm{e}^{\mathrm{i}kt}.$$

Since $f \in \mathcal{W}$, the Fourier sums $S_N f$ converge uniformly to f, i.e. for any $\tilde{\varepsilon} > 0$ there exists $\tilde{M}(\tilde{\varepsilon})$ such that

$$|S_N f - f| \leq \tilde{\varepsilon} \quad \text{for all} \quad N \geq \tilde{M}(\tilde{\varepsilon}).$$

This implies that for any $\varepsilon > 0$ there exists $M_1(\varepsilon)$ such that

$$\big||S_N f|^2 - |f|^2\big| \leq \varepsilon/3 \quad \text{for all} \quad N \geq M_1(\varepsilon).$$

Thus

$$\|\boldsymbol{A}_N(|S_N f|^2 - |f|^2)\|_2 \le \varepsilon/3. \tag{2.4}$$

Further, since $f \in \mathcal{W}$, for any $\varepsilon > 0$ there exists $M = M(\varepsilon) \ge M_1(\varepsilon)$ such that

$$\max\left\{ \sum_{k=-M+1}^{N-1} |a_k|, \sum_{k=-M+1}^{N-1} |a_k| \right\} \le \sqrt{\frac{\varepsilon}{3}} \quad \text{for all} \quad N \ge M(\varepsilon). \tag{2.5}$$

Now we split the triangular Toeplitz matrices in (2.3) into banded matrices and matrices of rank $\lfloor M/4 \rfloor$, i.e.

$$\boldsymbol{L}(\tilde{\boldsymbol{a}}_+) = \boldsymbol{B}_+ + \boldsymbol{R}_+ \,,\; \boldsymbol{L}(\tilde{\boldsymbol{a}}_-) = \boldsymbol{B}_- + \boldsymbol{R}_- \,, \tag{2.6}$$

where

$$\begin{aligned}
\boldsymbol{B}_+ &:= \boldsymbol{L}\left((0, \bar{a}_{N-1}, \ldots, \bar{a}_{M+1}, \boldsymbol{o}_M')'\right) \,,\; \boldsymbol{R}_+ := \boldsymbol{L}\left((\boldsymbol{o}_{N-M}', \bar{a}_M, \ldots, \bar{a}_1)'\right) , \\
\boldsymbol{B}_- &:= \boldsymbol{L}\left((0, a_{-(N-1)}, \ldots, a_{-(M+1)}, \boldsymbol{o}_M')'\right) \,,\; \boldsymbol{R}_- := \boldsymbol{L}\left((\boldsymbol{o}_{N-M}', a_{-M}, \ldots, a_{-1})'\right)
\end{aligned}$$

with zero vectors $\boldsymbol{o}_M'$ of length $\lfloor M/4 \rfloor$ and obtain

$$\begin{aligned}
\boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{L}^*(\tilde{\boldsymbol{a}}_+) + \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{L}(\tilde{\boldsymbol{a}}_-) = \boldsymbol{B}_+\boldsymbol{B}_+^* + \boldsymbol{B}_-^*\boldsymbol{B}_- \quad &+ \quad \boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{R}_+^* + \boldsymbol{R}_+\boldsymbol{B}_+^* \\
&+ \quad \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{R}_- + \boldsymbol{R}_-^*\boldsymbol{B}_-.
\end{aligned}$$

Now we have by (2.5) that

$$\|\boldsymbol{B}_+\boldsymbol{B}_+^* + \boldsymbol{B}_-^*\boldsymbol{B}_-\|_2 \le \|\boldsymbol{B}_+\|_1 \|\boldsymbol{B}_+^*\|_1 + \|\boldsymbol{B}_-^*\|_1 \|\boldsymbol{B}_-\|_1 \le \frac{2\varepsilon}{3},$$

while rank $(\boldsymbol{L}(\tilde{\boldsymbol{a}}_+)\boldsymbol{R}_+^* + \boldsymbol{R}_+\boldsymbol{B}_+^* + \boldsymbol{L}^*(\tilde{\boldsymbol{a}}_-)\boldsymbol{R}_- + \boldsymbol{R}_-^*\boldsymbol{B}_-) \le M$. Together with (2.3), (2.4) and Weyl's interlacing theorem [8, p. 184] this yields the assertion. ■

For a function f and equispaced nodes

$$x_{N,l} := w_N + \frac{2\pi l}{N} \quad \left(l = 0, \ldots, N-1; w_N \in \left[0, \frac{2\pi}{N}\right)\right) \tag{2.7}$$

we introduce the ω–*circulant matrices* ($\omega := \mathrm{e}^{\mathrm{i}N w_N}$)

$$\boldsymbol{M}_N(f) := \boldsymbol{W}_N \boldsymbol{F}_N \boldsymbol{D}_N(f) \boldsymbol{F}_N^* \boldsymbol{W}_N^*, \tag{2.8}$$

where

$$\boldsymbol{F}_N := \frac{1}{\sqrt{N}}(\mathrm{e}^{-2\pi\mathrm{i}jk/N})_{j,k=0}^{N-1}, \boldsymbol{W}_N := \mathrm{diag}(\mathrm{e}^{-\mathrm{i}k w_N})_{k=0}^{N-1}, \boldsymbol{D}_N(f) = \mathrm{diag}(f(x_{N,l}))_{l=0}^{N-1}.$$

See [12]. If $w_N = 0$, then $\boldsymbol{M}_N(f)$ is a *circulant matrix*. If q is a trigonometric polynomial, i.e.

$$q(t) = \sum_{k=-s_1}^{s_2} a_k(q) \mathrm{e}^{\mathrm{i}kt},$$

then, by [12], the matrices $\boldsymbol{A}_N(q)$ and $\boldsymbol{M}_N(q)$ are related by

$$\boldsymbol{A}_N(q) = \boldsymbol{M}_N(q) - \boldsymbol{B}_N(q) \tag{2.9}$$

where $\boldsymbol{B}_N(q) := (b_{j-k}(q))_{j,k=0}^{N-1}$ is the Toeplitz matrix of rank $s_1 + s_2$ with

$$\begin{array}{ll} b_{-(N-k)}(q) = \mathrm{e}^{\mathrm{i}Nw_N} a_k(q) & (k = 1, \dots, s_2), \\ b_{N-k}(q) = \mathrm{e}^{-\mathrm{i}Nw_N} a_{-k}(q) & (k = 1, \dots, s_1), \\ b_k(q) = 0 & \text{otherwise.} \end{array}$$

Having Lemma 2.1 in mind, we propose the Hermitian ω-circulant matrix $\boldsymbol{M}_N(|f|^2)$ as preconditioner for $\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f)$. If $|f| > 0$, then $\boldsymbol{M}_N(|f|^2)$ is positive definite. Further, by using Lemma 2.1, it is easy to prove the following lemma.

Lemma 2.2. *Let $f \in \mathcal{W}$ and let $|f| \geq f_{\min} > 0$. Then, for any $\varepsilon > 0$ and N sufficiently large,*

$$\boldsymbol{M}_N(|f|^2)^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f) = \boldsymbol{I}_N + \boldsymbol{U}_N + \boldsymbol{R}_N,$$

where $\|\boldsymbol{U}_N\|_2 \leq \varepsilon$ and $\boldsymbol{R}_N$ is a matrix of low rank independent of N.

Proof. By Lemma 2.1 and since $\|\boldsymbol{M}_N(|f|^2)^{-1}\|_2 \leq 1/f_{\min}^2$, it remains to show that

$$\boldsymbol{M}_N(|f|^2)^{-1}\boldsymbol{A}_N(|f|^2) = \boldsymbol{I}_N + \boldsymbol{U}_N + \boldsymbol{R}_N$$

with low norm and low rank matrices $\boldsymbol{U}_N$ and $\boldsymbol{R}_N$, respectively.

Since $|f|^2$ is continuous and $|f|^2 \geq f_{\min}^2 > 0$, for any $\varepsilon > 0$ there exists a trigonometric polynomial of degree $M = M(\varepsilon)$ such that

$$q - \frac{1}{2}\varepsilon f_{\min}^2 \leq |f|^2 \leq q + \frac{1}{2}\varepsilon f_{\min}^2. \tag{2.10}$$

Using this relation and the fact that $\lambda_{\min}(\boldsymbol{M}_N(|f|^2)) \geq f_{\min}^2$, we conclude that for every $\boldsymbol{o} \neq \boldsymbol{u} \in \boldsymbol{C}^N$

$$\frac{\boldsymbol{u}^*\boldsymbol{A}_N(q)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} - \frac{1}{2}\varepsilon \leq \frac{\boldsymbol{u}^*\boldsymbol{A}_N(|f|^2)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} \leq \frac{\boldsymbol{u}^*\boldsymbol{A}_N(q)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} + \frac{1}{2}\varepsilon. \tag{2.11}$$

By (2.9), the right hand inequality can be written as

$$\frac{\boldsymbol{u}^*\boldsymbol{A}_N(|f|^2)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} \leq \frac{\boldsymbol{u}^*\boldsymbol{M}_N(q)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} + \frac{\boldsymbol{u}^*\boldsymbol{B}_N(q)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} + \frac{1}{2}\varepsilon$$

and further by (2.10) and definition of $\boldsymbol{M}_N$ as

$$\frac{\boldsymbol{u}^*\boldsymbol{A}_N(|f|^2)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}} \leq 1 + \varepsilon + \frac{\boldsymbol{u}^*\boldsymbol{B}_N(q)\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{M}_N(|f|^2)\boldsymbol{u}}.$$

Handling the left-hand inequality of (2.11) in the same way and applying Weyl's interlacing theorem, we obtain the assertion. ■

The more interesting case even for practical purposes appears if we allow f to have zeros. In the following, let $f \in C_{2\pi}$ be given by

$$f = p_s\, h, \tag{2.12}$$

where $h \in \mathcal{W}$ with $|h| \geq h_{\min} > 0$ and

$$p_s(t) := \prod_{j=1}^{m} (\mathrm{e}^{\mathrm{i}t} - \mathrm{e}^{\mathrm{i}t_j})^{s_j}$$

with pairwise distinct zeros $t_j \in [-\pi, \pi)$ and $\sum_{j=1}^{m} s_j = s$. We choose our grid points $x_{N,l}$ $(l = 0, \ldots, N-1)$ for the construction of our ω-circulant preconditioner $\boldsymbol{M}_N(|f|^2)$ such that

$$x_{N,l} \neq t_j \quad (j = 1, \ldots, m; l = 1, \ldots, N-1). \tag{2.13}$$

See also Remark 2.6. Then $\boldsymbol{M}_N(|f|^2)$ is positive definite. Moreover, we will prove that the eigenvalues of $\boldsymbol{M}_N(|f|^2)^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f)$ have a proper cluster at 1. Note that these eigenvalues coincide with the square of singular values of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$.

Theorem 2.3. *Let $f \in \mathcal{W}$ be given by* (2.12). *Let $\boldsymbol{M}_N(f)$ be defined by* (2.8) *and* (2.13). *Then, for any $\varepsilon > 0$ and N sufficiently large,*

$$(\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1})^*(\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}) = \boldsymbol{I}_N + \boldsymbol{U}_N + \boldsymbol{R}_N\,,$$

where $\|\boldsymbol{U}_N\|_2 \leq \varepsilon$ and $\boldsymbol{R}_N$ is a matrix of low rank independent of N.

Proof. By straightforward calculation, we obtain for $0 \leq m \leq N-1, 0 \leq n \leq N-s$

$$(\boldsymbol{A}_N(h)\boldsymbol{A}_N(p_s))_{m,n} = \sum_{j=0}^{s} (p_s)_j h_{k-j} = (\boldsymbol{A}_N(hp_s))_{m,n},$$

i.e.

$$\boldsymbol{A}_N(hp_s) = \boldsymbol{A}_N(h)\boldsymbol{A}_N(p_s) + \boldsymbol{C}_N(s), \tag{2.14}$$

where $\boldsymbol{C}_N(s)$ has only nonzero entries in its last s columns.

Then, by (2.9) and definition of $\boldsymbol{M}_N(f)$

$$\begin{aligned}
\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1} &= (\boldsymbol{A}_N(h)\boldsymbol{A}_N(p_s) + \boldsymbol{C}_N(s))\boldsymbol{M}_N(f)^{-1} \\
&= (\boldsymbol{A}_N(h)(\boldsymbol{M}_N(p_s) - \boldsymbol{B}_N(p_s)) + \boldsymbol{C}_N(s))\boldsymbol{M}_N(f)^{-1} \\
&= \boldsymbol{A}_N(h)\boldsymbol{M}_N(h)^{-1} + \boldsymbol{R}_N(s)\,.
\end{aligned}$$

Thus

$$(\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1})^*(\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}) = (\boldsymbol{A}_N(h)\boldsymbol{M}_N(h)^{-1})^*\boldsymbol{A}_N(h)\boldsymbol{M}_N(h)^{-1} + \boldsymbol{R}_N(2s).$$

The rest of the proof follows immediately by Lemma 2.2 and Weyl's interlacing theorem. ■

The proper clustering of the singular values of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$ leads to a superlinear convergence of the CG-method applied to the normal equation (1.2). In order to estimate the number of iteration steps we have to estimate the smallest singular values of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$. Following the lines of [5], we first prove the following lemma.

Lemma 2.4. *Let $f \in \mathcal{W}$ be given by* (2.12). *Let $\boldsymbol{M}_N(f)$ be defined by* (2.8) *and* (2.13). *Then there exists $c > 0$ independent of N such that*

$$\|\boldsymbol{M}_N(f)\boldsymbol{A}_N(f)^{-1}\|_2 \leq c\,\kappa_2(\boldsymbol{A}_N(f)),$$

where $\kappa_2(\boldsymbol{A}_N(f)) := \|\boldsymbol{A}_N(f)^{-1}\|_2\|\boldsymbol{A}_N(f)\|_2$ denotes the spectral condition number *of $\boldsymbol{A}_N(f)$.*

Proof. Since

$$\|\boldsymbol{M}_N(f)\boldsymbol{A}_N(f)^{-1}\|_2 \leq \|\boldsymbol{A}_N(f)\|_2^{-1}\|\boldsymbol{M}_N(f)\|_2\;\kappa_2(\boldsymbol{A}_N(f))$$

and since $\|\boldsymbol{M}_N(f)\|_2$ is bounded from above, it remains to show that there exists $c > 0$ independent of N such that

$$\|\boldsymbol{A}_N(f)\|_2 \geq c.$$

But this follows immediately from the fact that the singular values of $\boldsymbol{A}_N(f)$ are distributed as $|f|$ (see [11, 19]). ■

Theorem 2.5. *Let $f \in \mathcal{W}$ given by* (2.12) *and let $\kappa_2(\boldsymbol{A}_N(f)) = N^\alpha (\alpha > 0)$. Let $\boldsymbol{M}_N(|f|^2)$ be defined by* (2.8) *and* (2.13). *Then CG applied to*

$$\boldsymbol{M}_N(|f|^2)^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f)\boldsymbol{x} = \boldsymbol{M}_N(|f|^2)^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{b},$$

requires $\mathcal{O}(\log N)$ iteration steps to produce a solution of prescribed precision.

Proof. By a result of Axelsson [1, p. 573] and by Theorem 2.3, the number of iterations of the CG-method to obtain a solution of the above equation up to a prescribed precision τ is given by

$$\left\lceil \left(\ln\frac{2}{\tau} + \sum_{k=1}^{q}\ln\frac{1+\varepsilon}{\sigma_k^2}\right) \Big/ \ln\left(\frac{1+\left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{1/2}}{1-\left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{1/2}}\right)\right\rceil + p + q$$

where $\sigma_1 \leq \ldots \leq \sigma_q$ are the smallest singular values of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$ which are not inside our cluster $[1-\varepsilon, 1+\varepsilon]$ and where $p + q = \text{rank }(\boldsymbol{R}_N)$. Here ε and $\boldsymbol{R}_N$ are taken with respect to Theorem 2.3. By Lemma 2.4 we have that $\sigma_1^2 \geq \|\boldsymbol{M}_N(f)\boldsymbol{A}_N(f)^{-1}\|_2^{-2} \geq cN^{-2\alpha}$ and consequently

$$\sum_{k=1}^{q}\ln\left(\frac{1+\varepsilon}{\sigma_k^2}\right) \leq 2\,\alpha\,q\,c\,\ln N.$$

This yields the assertion. ■

By the following remark it is possible to neglect the grid condition (2.13). In particular this implies that it is always possible to work with circulant instead of ω-circulant preconditioners. To keep proofs short we have introduced the more general ω-circulant preconditioners.

Remark 2.6. Let the grid points $x_{N,l}$ $(l = 0, \ldots, N-1)$ be given by (2.7) where we do not assume (2.13). We define, similar as in [18], $\tilde{\boldsymbol{M}}_N(f)$ by

$$\tilde{\boldsymbol{M}}_N(f) := \boldsymbol{W}_N\boldsymbol{F}_N\text{diag}(f(\tilde{x}_l))_{l=0}^{N-1}\boldsymbol{F}_N^*\boldsymbol{W}_N^*, \tag{2.15}$$

$$\tilde{x}_l := \begin{cases} x_l & \text{if} \quad x_l \neq t_j \quad (j = 1, \ldots, m), \\ x_{\tilde{l}} & \text{otherwise,} \end{cases}$$

where $\tilde{l} \in \{0, \ldots, N-1\}$ is the next higher index to l such that $|f(x_{\tilde{l}})| > 0$. For N large enough we can simply choose $\tilde{l} = l + 1 \bmod N$. Then, since by construction

$$\boldsymbol{M}_N(f) = \tilde{\boldsymbol{M}}_N(f) + \boldsymbol{R}_N(m) \tag{2.16}$$

it is easy to check that the above theorems remain valid with a small fixed number of more outlying eigenvalues. In particular, if we choose $x_{N,l} := 2\pi l/N$ $(l = 0, \ldots, N-1)$ we obtain circulant preconditioners $\tilde{\boldsymbol{M}}_N(f)$. □

Beyond application of PCG to the normal equation (1.2) we can use other iterative methods like GMRES and BICGSTAB for the solution of (1.1). These methods avoid the translation of the original system to the normal equation. However, by Remark 4.1, the arithmetical complexity per iteration step of BICGSTAB is "nearly the same" as the arithmetical complexity of CG applied to the normal equation.

As preconditioner we suggest $\boldsymbol{M}_N(f)$, respectively $\tilde{\boldsymbol{M}}_N(f)$ here. The numerical results concerning the number of iteration steps of preconditioned GMRES and BICGSTAB are very good. Unfortunately, the number of iterations does no longer depend on the distribution of the singular values of the preconditioned matrix but on its eigenvalues. For arbitrary $f \in C_{2\pi}$ of the form (2.12) we were not able to prove properties concerning the distribution of the eigenvalues of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$. But for special generating functions, namely rational functions, we obtain the following result (see also [17]):

Theorem 2.7. *Let f be a rational function of order (s_1, s_2) $(s_1 s_2 \neq 0)$ given by*

$$f(z) = \frac{p(z)}{q(z)} = \frac{p_0 + p_1 z + \ldots + p_{s_1} z^{s_1}}{q_0 + q_1 z + \ldots + q_{s_2} z^{s_2}}.$$

Define $\boldsymbol{M}_N(f)$ by (2.8) with grid points satisfying (2.13) if $f(\mathrm{e}^{\mathrm{i}t_j}) = 0$ $(j = 1, \ldots, m)$. Then

$$\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1} = \boldsymbol{I}_N + \boldsymbol{R}_N(\max\{s_1, s_2\}).$$

Proof. By (2.9), (2.14) and definition of $\boldsymbol{M}_N(f)$, we obtain

$$\begin{aligned} \boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1} &= \left(\boldsymbol{A}_N\left(\frac{1}{q}\right)\boldsymbol{A}_N(p) + \boldsymbol{C}_N(s_1)\right)\boldsymbol{M}_N^{-1}\left(\frac{p}{q}\right) \\ &= \left(\boldsymbol{A}_N\left(\frac{1}{q}\right)(\boldsymbol{M}_N(p) - \boldsymbol{B}_N(p)) + \boldsymbol{C}(s_1)\right)\boldsymbol{M}_N^{-1}\left(\frac{p}{q}\right) \\ &= \boldsymbol{A}_N\left(\frac{1}{q}\right)\boldsymbol{M}_N^{-1}\left(\frac{1}{q}\right) + \boldsymbol{R}_N(s_1), \end{aligned} \tag{2.17}$$

where only the last s_1 columns of $\boldsymbol{R}_N(s_1)$ are nonzero columns. Similar we obtain by (2.14), (2.9) and definition of $\boldsymbol{M}_N(f)$

$$\begin{aligned} \boldsymbol{I}_N &= \boldsymbol{A}_N\left(\frac{1}{q}\right)\boldsymbol{A}_N(q) + \boldsymbol{C}_N(s_2) \\ &= \boldsymbol{A}_N\left(\frac{1}{q}\right)(\boldsymbol{M}_N(q) - \boldsymbol{B}_N(q)) + \boldsymbol{C}_N(s_2) \\ &= \boldsymbol{A}_N\left(\frac{1}{q}\right)\boldsymbol{M}_N^{-1}\left(\frac{1}{q}\right) + \boldsymbol{R}(s_2), \end{aligned} \tag{2.18}$$

where only the last s_2 columns of $\boldsymbol{R}_N(s_2)$ are nonzero columns. Now the assertion follows by (2.17) and (2.18). ■

By Remark 2.6 we can prove a similar result with respect to the preconditioner $\tilde{\boldsymbol{M}}_N(f)$.

Note that similar results concerning the number of outliers outside $[1-\varepsilon, 1+\varepsilon]$ were obtained for the preconditioned matrices in [9] if $|f| > 0$. The preconditioners in [9] do not require the explicit knowledge of the generating function.

Let $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1} = \boldsymbol{I}_N + \boldsymbol{R}_N(s)$. By [14, p. 195], the residual

$$\boldsymbol{r}^{(k)} := \boldsymbol{b} - (\boldsymbol{I}_N + \boldsymbol{R}_N(s))\boldsymbol{y}^{(k)}$$

of the k-th iteration of GMRES applied to the preconditioned system can be estimated by

$$\frac{||\boldsymbol{r}^{(k)}||_2}{||\boldsymbol{r}^{(0)}||_2} \leq \min_{p \in \Pi_k^0} ||p(\boldsymbol{I}_N + \boldsymbol{R}_N(s))||_2 ,$$

where Π_k^0 denotes the space of polynomials of degree $\leq k$ with $p(0) = 1$. Assume that $\boldsymbol{R}_N(s)$ has the pairwise distinct nonzero eigenvalues $\lambda_1, \ldots, \lambda_q$ with multiplicities $n_1, \ldots, n_q$, where $n_1 + \ldots + n_q = s$. Let

$$\boldsymbol{R}_N = \boldsymbol{X} \operatorname{diag}(\tilde{\boldsymbol{J}}_{1,1}, \ldots, \tilde{\boldsymbol{J}}_{1,m_1}, \ldots, \tilde{\boldsymbol{J}}_{q,1}, \ldots, \tilde{\boldsymbol{J}}_{q,m_q}, \boldsymbol{o}_{N-s}) \boldsymbol{X}^{-1}$$

with $\tilde{\boldsymbol{J}}_{j,k} \in \mathbb{C}^{n_{j,k}, n_{j,k}}$ and nonincreasing sizes $n_{j,k}$ fulfilling $n_{j,1} + \ldots + n_{j,m_j} = n_j$, be the decomposition of $\boldsymbol{R}_N$ into Jordan blocks. Then

$$\begin{aligned} \boldsymbol{I}_N + \boldsymbol{R}_N(s) &= \boldsymbol{X} \operatorname{diag}(\boldsymbol{J}_{1,1}, \ldots, \boldsymbol{J}_{q,m_q}, \boldsymbol{I}_{N-s}) \boldsymbol{X}^{-1}, \quad \boldsymbol{J}_{j,k} := \boldsymbol{I}_{n_{j,k}} + \tilde{\boldsymbol{J}}_{j,k}, \\ p(\boldsymbol{I}_N + \boldsymbol{R}_N(s)) &= \boldsymbol{X} \operatorname{diag}(p(\boldsymbol{J}_{1,1}), \ldots, \ldots, p(\boldsymbol{J}_{q,m_q}), p(\boldsymbol{I}_{N-s})) \boldsymbol{X}^{-1}, \end{aligned} \tag{2.19}$$

where by [7, p. 557]

$$p(\boldsymbol{J}_{j,k}) = \begin{pmatrix} p(\lambda_j+1) & p^{(1)}(\lambda_j+1) & \cdots & \cdots & \frac{p^{(n_{j,k}-1)}(\lambda_j+1)}{(n_{j,k}-1)!} \\ 0 & p(\lambda_j+1) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & p^{(1)}(\lambda_j+1) \\ 0 & \cdots & \cdots & \cdots & p(\lambda_j+1) \end{pmatrix}. \tag{2.20}$$

Choosing $p \in \Pi_{s+1}^0$ as

$$p(x) := (1-x) \prod_{j=1}^{q} \left(1 - \frac{x}{\lambda_j + 1}\right)^{n_{j,1}} \qquad (\lambda_j \neq -1),$$

we see by (2.20) and (2.19) that $p(\boldsymbol{I}_N + \boldsymbol{R}_N(s))$ becomes zero. Consequently, GMRES terminates after at most $s+1$ steps. Applied to our setting this means by Theorem 2.7 that GMRES requires at most $1 + \max\{s_1, s_2\}$ iteration steps.

If we replace the ω-circulant preconditioner $\boldsymbol{M}_N(f)$ by the circulant preconditioner $\tilde{\boldsymbol{M}}_N(f)$ the number of GMRES steps may increase at most to $1 + \max\{s_1, s_2\} + m$.

3. TRIGONOMETRIC PRECONDITIONERS

Since the function $|f|^2$ is even, the matrix $\boldsymbol{A}_N(|f|^2)$ is symmetric. This suggests the application of so-called trigonometric preconditioners. These are matrices which are diagonalizable by trigonometric transforms. In practice, four discrete sine transforms (DST I – IV) and four discrete cosine transforms (DCT I – IV) were used (see [20]). Any of these eight trigonometric transforms can be realized with $\mathcal{O}(N \log N)$ arithmetical operations. Likewise, we can define preconditioners with respect to any of these transforms.

In this paper, we restrict our attention to the so-called discrete cosine transform of type II (DCT-II) and discrete sine transform of type II (DST-II), which are determined by the following transform matrices:

$$\text{DCT–II} \quad : \quad \boldsymbol{C}_N^{II} := \left(\frac{2}{N}\right)^{1/2} \left(\epsilon_j^N \cos \frac{j(2k+1)\pi}{2N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N},$$
$$\text{DST–II} \quad : \quad \boldsymbol{S}_N^{II} := \left(\frac{2}{N}\right)^{1/2} \left(\epsilon_{j+1}^N \sin \frac{(j+1)(2k+1)\pi}{2N}\right)_{j,k=0}^{N-1} \in \mathbb{R}^{N,N},$$

where $\epsilon_k^N := 2^{-1/2}(k = 0, N)$ and $\epsilon_k^N := 1 \ (k = 1, \dots, N-1)$. For (1.2) we propose the preconditioners

$$\begin{aligned} \text{DCT} - \text{II}: \quad & \tilde{\boldsymbol{M}}_N(|f|^2, \boldsymbol{C}_N^{II}) \quad := \quad (\boldsymbol{C}_N^{II})^{\mathrm{T}} \operatorname{diag}\left(|f(\tilde{x}_{N,l})|^2\right)_{l=0}^{N-1} \boldsymbol{C}_N^{II}, \\ \text{DST} - \text{II}: \quad & \tilde{\boldsymbol{M}}_N(|f|^2, \boldsymbol{S}_N^{II}) \quad := \quad (\boldsymbol{S}_N^{II})^{\mathrm{T}} \operatorname{diag}\left(|f(\tilde{x}_{N,l})|^2\right)_{l=1}^{N} \boldsymbol{S}_N^{II}, \end{aligned} \tag{3.21}$$

where

$$\tilde{x}_{N,l} := \begin{cases} \frac{l\pi}{N} & \text{if} \quad \frac{l\pi}{N} \neq t_j \quad (j = 1, \dots, m), \\ \frac{\tilde{l}\pi}{N} & \text{otherwise} \end{cases}$$

and where $\tilde{l} \in \{0, \dots, N-1\}$ is the next higher index to l such that $|f(x_{N,l})| > 0$. See [13].

Then we can prove in a completely similar way as in Section 2 the following

Theorem 3.1. *Let $f \in \mathcal{W}$ given by (2.12) and let $\kappa_2(\boldsymbol{A}_N(f)) = N^\alpha$ $(\alpha > 0)$. Then CG applied to*

$$\boldsymbol{M}_N^{-1}(|f|^2, \boldsymbol{O}_N)\, \boldsymbol{A}_N^*(f)\, \boldsymbol{A}_N(f)\boldsymbol{x} = \boldsymbol{M}_N^{-1}(|f|^2, \boldsymbol{O}_N)\, \boldsymbol{A}_N^*(f)\, \boldsymbol{b} \tag{3.22}$$

with $\boldsymbol{O}_N \in \{\boldsymbol{C}_N^{II}, \boldsymbol{S}_N^{II}\}$ requires $\mathcal{O}(\log N)$ iteration steps to produce a solution of prescribed precision.

4. NUMERICAL RESULTS AND AN APPLICATION TO QUEUEING NETWORKS

In this section, we test our ω-circulant and trigonometric preconditioners on a SGI O2 work station. As transform length we chose $N = 2^n$ and as right-hand side $\boldsymbol{b}$ of (1.1) the vector consisting of N entries "1". The iterative methods started with the zero vector and stopped if $\|\boldsymbol{r}^{(j)}\|_2 / \|\boldsymbol{r}^{(0)}\|_2 < 10^{-7}$, where $\boldsymbol{r}^{(j)}$ denotes the residual vector after j iterations.

We compare four different iterative methods. We apply CGS, BICGSTAB and restarted GMRES(20) to

$$\tilde{\boldsymbol{M}}_N(f)^{-1}\,\boldsymbol{A}_N(f)\,\boldsymbol{x} = \tilde{\boldsymbol{M}}_N(f)^{-1}\,\boldsymbol{b}$$

with circulant preconditioner $\tilde{\boldsymbol{M}}_N(f)$.

Further we use CG for solving

$$\boldsymbol{M}_N^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f)\boldsymbol{x} = \boldsymbol{M}_N^{-1}\boldsymbol{A}_N^*(f)\boldsymbol{b}\,,$$

where $\boldsymbol{M}_N$ denotes one of the following preconditioners:

$$\begin{array}{lll} \boldsymbol{M}_N(|f|^2,\boldsymbol{F}_N) \;=\; \tilde{\boldsymbol{M}}_N(|f|^2) & \text{given by} & (2.15), \\ \tilde{\boldsymbol{M}}_N(|f|^2,\boldsymbol{C}_N^{II})\;,\; \tilde{\boldsymbol{M}}_N(|f|^2,\boldsymbol{S}_N^{II}) & \text{given in} & (3.21) \end{array}$$

and the so-called *optimal trigonometric preconditioners*

$$\boldsymbol{M}_N^O(\boldsymbol{C}_N^{II})\;,\;\boldsymbol{M}_N^O(\boldsymbol{S}_N^{II})$$

of $\boldsymbol{A}_N^*(f)\boldsymbol{A}_N(f)$ introduced in [12].

Although CGS is not very common, we have included this method to compare our results with the results in [3]. In [3], CGS is used without an analysis of the number of necessary iteration steps. Moreover, the application of our circulant preconditioner requires less arithmetical operations than the use of the preconditioner consisting of the product of a banded Toeplitz and an optimal circulant matrix proposed in the above paper.

The following remark shortly prescribes the effort per iteration step of the proposed methods.

Remark 4.1. Each iteration step of BICGSTAB and CGS requires two matrix vector products with the Toeplitz matrix $\boldsymbol{A}_N(f)$ and two matrix vector products with the preconditioner.

In contrast, GMRES requires only one matrix vector products with the Toeplitz matrix $\boldsymbol{A}_N(f)$ and one matrix vector product with the preconditioner. But the number of inner products grows linearly with the iteration number, up to the restart.

The PCG-method applied to the normal equation requires one matrix vector product with both $\boldsymbol{A}_N(f)$ and $\boldsymbol{A}_N^*(f)$, and one matrix vector product with the preconditioner. □

First we test Toeplitz systems with the following rational functions as generating functions (cf. [3]):

$$\text{(i)}\;\; f_1(z) := \frac{(z^4-1)}{(z-\frac{3}{2})(z-\frac{1}{2})} = \frac{15}{8}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{13}{24} + \frac{7}{36}z - \frac{11}{54}z^2 - \frac{65}{24}\sum_{k=3}^{\infty}\left(\frac{2z}{3}\right)^k.$$

$$\text{(ii)}\;\; f_2(z) := \frac{(z+1)^2(z-1)^2}{(z-\frac{3}{2})(z-\frac{1}{2})} = -\frac{9}{8}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{5}{24} + \frac{47}{36}z - \frac{29}{54}z^2 - \frac{25}{24}\sum_{k=3}^{\infty}\left(\frac{2z}{3}\right)^k.$$

$$\text{(iii)}\;\; f_3(z) := \frac{(z+1)^2(z-1)}{(z-\frac{3}{2})(z-\frac{1}{2})} = \frac{9}{4}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{11}{12} - \frac{7}{18}z - \frac{25}{12}\sum_{k=2}^{\infty}\left(\frac{2z}{3}\right)^k.$$

Tables 1–3 present the number of iteration steps with different iterative methods and different preconditioners.

The first row of each table contains the exponent n of the transform length $N = 2^n$. The corresponding iterative method is listed in the first column and the preconditioner in the second column of each table. The symbol $*$ denotes that the method stopped without converging to the desired tolerance in 500 iteration steps or that the method stagnated.

In addition to the above preconditioners we also test the ω-circulant preconditioner $\boldsymbol{M}_N(f) = \boldsymbol{M}_N(f, \boldsymbol{F}_N\boldsymbol{W}_N)$ determined by (2.8) with $w_N = \frac{\pi}{N}$ in connection with GMRES(20). Since $N \geq 16$ is even, the grid points fulfill (2.13). By Theorem 2.7 we obtain for our three examples that

(i) 4 eigenvalues of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$ are not equal to 1

(Here all these eigenvalues are equal to 1/2 !),

(ii) 4 eigenvalues of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$ are not equal to 1,

(iii) 3 eigenvalues of $\boldsymbol{A}_N(f)\boldsymbol{M}_N(f)^{-1}$ are not equal to 1.

Indeed GMRES requires 2, 4 and 3 iterations, respectively.

For the circulant preconditioner $\tilde{\boldsymbol{M}}_N(f)$ we have to replace those of the above grid points which meet the zeros of the generating function, i.e. in (i) the points $0, \pi/2, \pi, 3\pi/2$ and in (ii),(iii) the points $0, \pi$.

By Remark 2.6 this implies 8, 6 and 5 outliers in (i), (ii) and (iii), respectively. Our numerical results confirm our expectations.

method	$\boldsymbol{M}_N$	4	5	6	7	8	9	10	11	12
CGS	$\boldsymbol{I}_N$	22	75	379	*	*	*	*	*	*
CGS	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	8	8	8	8	8	*	*	*	*
BICGSTAB	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	7.5	7.5	7.5	7.5	7	7	6	7	9
GMRES(20)	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	8	8	8	8	8	8	8	8	8
GMRES(20)	$\boldsymbol{M}_N(f, \boldsymbol{F}_N\boldsymbol{W}_N)$	3	2	2	2	2	2	2	2	2
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{F}_N)$	13	13	15	18	18	19	22	23	28
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{C}_N^{II})$	10	11	11	13	15	15	18	19	22
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{S}_N^{II})$	10	11	12	12	14	15	16	16	19
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{C}_N^{II})$	14	19	22	29	35	45	58	77	104
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{S}_N^{II})$	17	26	35	45	59	77	105	144	200

Table 1: $f(t) = f_1(\mathrm{e}^{\mathrm{i}t}) \quad (t \in [-\pi, \pi))$

The next examples with generating functions

(iv) $f_4(t) := \mathrm{i}t$,

(v) $f_5(t) := t^2 e^{\mathrm{i}t}$

show that PCG applied to the normal equation can outperform the other 3 methods. The function in $f_4(z) = \log(z)$ $(z := e^{\mathrm{i}t})$ in (iv) is of special interest. The first row and column of

method	$\boldsymbol{M}_N$	4	5	6	7	8	9	10	11	12
CGS	$\boldsymbol{I}_N$	28	156	*	*	*	*	*	*	*
CGS	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	6	6	7	7	33	*	*	*	*
BICGSTAB	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	5.5	6.5	6.5	6.5	6	7.5	8	11	11
GMRES(20)	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	6	6	6	6	6	6	6	7	7
GMRES(20)	$\boldsymbol{M}_N(f, \boldsymbol{F}_N \boldsymbol{W}_N)$	5	5	4	4	4	4	4	4	4
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{F}_N)$	11	13	16	20	28	42	53	71	125
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{C}_N^{II})$	8	10	11	14	22	28	34	43	62
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{S}_N^{II})$	11	12	14	18	22	31	42	51	71
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{C}_N^{II})$	14	19	27	38	61	129	276	*	*
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{S}_N^{II})$	17	27	41	75	159	364	*	*	*

Table 2: $f(t) = f_2(\mathrm{e}^{\mathrm{i}t}) \quad (t \in [-\pi, \pi))$

method	$\boldsymbol{M}_N$	4	5	6	7	8	9	10	11	12
CGS	$\boldsymbol{I}_N$	26	70	218	*	*	*	*	*	*
CGS	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	5	6	6	6	6	7	7	7	18
BICGSTAB	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	4.5	5.5	5.5	5.5	5.5	6.5	6.5	7.5	7.5
GMRES(20)	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	5	5	5	5	5	5	5	5	6
GMRES(20)	$\boldsymbol{M}_N(f, \boldsymbol{F}_N \boldsymbol{W}_N)$	4	4	3	3	3	3	3	3	3
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{F}_N)$	9	9	11	13	15	18	22	26	39
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{C}_N^{II})$	7	8	8	9	11	12	14	16	20
PCG	$\tilde{\boldsymbol{M}}_N(\lvert f\rvert^2, \boldsymbol{S}_N^{II})$	9	9	11	11	15	16	17	24	31
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{C}_N^{II})$	13	16	19	26	32	42	57	79	130
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{S}_N^{II})$	18	34	53	99	204	472	*	*	*

Table 3: $f(t) = f_3(\mathrm{e}^{\mathrm{i}t}) \quad (t \in [-\pi, \pi))$

$\boldsymbol{A}_N(\mathrm{i}t)$ are given by

$$\left(0, -1, 1/2, -1/3, 1/4, \dots, \frac{(-1)^{N+1}}{N}\right)$$

and

$$\left(0, 1, -1/2, 1/3, -1/4, \dots, \frac{(-1)^{N}}{N}\right),$$

respectively. Matrices of this kind arise in sinc-collocation methods for initial value problems (see [2]). Note further that (iv) is up to now the only example where the optimal trigonometric preconditioner works well, too.

method	$\boldsymbol{M}_N$	4	5	6	7	8	9	10	11	12
CGS	$\boldsymbol{I}_N$	16	*	*	*	*	*	*	*	*
CGS	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	*	*	*	*	*	*	*	*	*
BICGSTAB	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	*	*	*	*	*	*	*	*	*
GMRES(20)	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	7	8	9	9	10	10	11	11	12
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{F}_N)$	5	5	6	7	7	7	8	11	15
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{C}_N^{II})$	5	5	5	8	8	8	9	11	13
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{S}_N^{II})$	5	5	5	6	6	6	6	7	9
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{C}_N^{II})$	8	10	13	17	20	26	33	42	56
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{S}_N^{II})$	6	7	7	7	8	8	8	9	10

Table 4: $f(t) = \mathrm{i}t \quad (t \in [-\pi, \pi))$

Finally, we apply our methods to Markovian queueing models with batch arrivals considered in [3]: The input of the queueing system will be an exogenous Poisson batch arrival process with mean batch interarrival time λ^{-1}. By $\lambda_k = \lambda p_k$ we denote the batch arrival rate for batches of size k, where p_k is the probability that the arrival batch size is k. The number of servers in the queueing system is s. Finally we assume that the service time of each server is independent of the others and is exponentially distributed with mean μ^{-1}. The waiting room is of size $N - s - 1$ and the queueing discipline is blocked–customers–cleared. If the arrival batch size is larger than the number of waiting places left, then only part of the arrival batch will be accepted, the other customers will be treated as overflows and will be cleared from the system. This kind of queueing system occurs in many applications, such as telecommunication networks [10] and loading dock models [15].

By [3], the probability distribution vector of the queueing system is given by the solution of a system of linear equations

$$(\boldsymbol{A}_N(f) + \boldsymbol{R}_N(s))\boldsymbol{x} = (0, \dots, 0, s\mu)', \tag{4.1}$$

method	M_N	4	5	6	7	8	9	10
CGS	I_N	58	*	*	*	*	*	*
CGS	$\tilde{M}_N(f, F_N)$	17	*	*	*	*	*	*
BICGSTAB	$\tilde{M}_N(f, F_N)$	18.5	37	58.5	84	85	94.5	102
GMRES(20)	$\tilde{M}_N(f, F_N)$	16	176	*	*	*	*	*
PCG	$\tilde{M}_N(\|f\|^2, F_N)$	11	14	15	21	26	*	*
PCG	$\tilde{M}_N(\|f\|^2, C_N^{II})$	12	15	17	20	28	35	40
PCG	$\tilde{M}_N(\|f\|^2, S_N^{II})$	10	11	11	14	14	20	21
PCG	$\tilde{M}_N^O(C_N^{II})$	17	29	53	111	257	631	1812
PCG	$\tilde{M}_N^O(S_N^{II})$	14	17	19	24	31	45	62

Table 5: $f(t) = t^2 \mathrm{e}^{\mathrm{i}t} \quad (t \in [-\pi, \pi))$

where $\boldsymbol{A}_N(f)$ denotes the lower Hessenberg Toeplitz matrix with generating function

$$f(z) := -s\mu\frac{1}{z} + \lambda + s\mu - \sum_{k=1}^{\infty} \lambda_k z^k \quad (z := e^{\mathrm{i}t}).$$

Clearly, our preconditioners $\boldsymbol{M}_n(f)$ also lead to a clustering of the singular values of $\boldsymbol{M}_N(f)^{-1}(\boldsymbol{A}_N(f) + \boldsymbol{R}_N(s))$.

As examples we choose $s \in \{1, 4\}$, $\lambda = 1$, $\mu = s$ and $\lambda_k = 2^k$ (cf. [3]). In this case

$$f(z) = f_6(z) := \frac{(z-1)(2z + s + \mu z - 2s\mu)}{z(z-2)}.$$

By Theorem 2.7, the matrices $\boldsymbol{M}_N(f)^{-1}(\boldsymbol{A}_N(f) + \boldsymbol{R}_N(s))$ have only 3, respectively 6 eigenvalues which are not equal to 1. The numerical results for $f_6(z)$ are reported in Table 6.

REFERENCES

1. O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, Cambridge, 1996.
2. T. S. Carlson, J. Dockery, and J. Lund. A sinc–collocation method for initial value problems. *Math. Comp.*, 66:215 – 235, 1997.
3. 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*, 17:762 – 772, 1996.
4. R. H. Chan and M. K. Ng. Conjugate gradient methods of Toeplitz systems. SIAM *Review*, 38:427 – 482, 1996.
5. R. H. Chan and M.-C. Yeung. Circulant preconditioners for complex Toeplitz systems. SIAM *J. Numer. Anal.*, 30:1193 – 1207, 1993.
6. F. Di Benedetto. Solution of Toeplitz normal equations by sine transform based preconditioning. *Linear Algebra Appl.*, 285:229 – 255, 1998.

s		1					4				
method	$\boldsymbol{M}_N$	4	6	8	10	12	4	6	8	10	12
CGS	$\boldsymbol{I}_N$	15	*	*	*	*	16	*	*	*	*
CGS	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	3	3	3	3	3	6	6	6	6	6
BICGSTAB	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	2.5	2.5	2.5	2.5	2.5	5.5	5.5	5.5	5.5	5.5
GMRES(20)	$\tilde{\boldsymbol{M}}_N(f, \boldsymbol{F}_N)$	3	3	3	3	3	6	6	6	6	6
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{F}_N)$	5	5	6	6	8	8	8	9	9	9
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{S}_N^{II})$	5	5	5	7	7	8	8	8	8	9
PCG	$\tilde{\boldsymbol{M}}_N(\|f\|^2, \boldsymbol{C}_N^{II})$	4	4	4	4	5	7	7	7	7	7
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{C}_N^{II})$	11	21	35	62	116	14	25	39	68	120
PCG	$\tilde{\boldsymbol{M}}_N^O(\boldsymbol{S}_N^{II})$	8	8	7	6	6	11	11	11	10	10

Table 6: $f(t) = f_6(\mathrm{e}^{\mathrm{i}t})$ with $s \in \{1, 4\}$ $\quad (t \in [-\pi, \pi))$

7. G. H. Golub and Ch. F. Van Loan, *Matrix Computations*, The J. Hopkins University Press, third edition, Baltimore and London, 1996.
8. R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.
9. T. Ku and C. Kuo. Spectral properties of preconditioned rational Toeplitz matrices: The nonsymmetric case. SIAM *J. Matrix Anal. Appl.*, 14:521 – 544, 1993.
10. T. Oda. Moment analysis for traffic associated with Markovian queueing systems. *IEEE Trans. Comm.*, 39:737 – 745, 1991.
11. S. V. Parter. On the distribution of singular values of Toeplitz matrices. *Linear Algebra Appl.*, 80:115 – 130, 1986.
12. D. Potts and G. Steidl. Optimal trigonometric preconditioners for nonsymmetric Toeplitz systems. *Linear Algebra Appl.*, 281:265 – 292, 1998.
13. D. Potts and G. Steidl. Preconditioners for ill–conditioned Toeplitz matrices. *BIT*, 39:513 – 533, 1999.
14. Y. Saad. *Iterative Methods for Sparse Linear Systems.* PWS Publ., Boston, 1996.
15. A. Seila. Multivariate estimation of conditional performance measure in regenerative simulation. *Amer. J. Math. Mgt. Sci.*, 10:17 – 45, 1990.
16. P. Sonneveld. A fast Lanczos-type solver for non-symmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 10:36 – 52, 1989.
17. N. Trefethen. Approximation Theory and Numerical Linear Algebra. In J. Mason and M. Cox, editors, *Algorithms for Approximation II*, Chapman and Hall, London, 1990.
18. E. E. Tyrtyshnikov. Circulant preconditioners with unbounded inverses. *Linear Algebra Appl.*, 216:1 – 23, 1995.
19. E. E. Tyrtyshnikov. A unifying approach to some old and new theorems on distribution and clustering. *Linear Algebra Appl.*, 232:1 – 43, 1996.
20. Z. Wang. Fast algorithms for the discrete W transform and for the discrete Fourier transform. IEEE *Trans. Acoust. Speech Signal Process*, 32:803 – 816, 1984.