A New Preconditioner for Toeplitz Matrices

María Elena Domínguez-Jiménez and Paulo J. S. G. Ferreira

***Abstract*—In this paper we introduce and analyze a new preconditioner for Toeplitz matrices that exhibits excellent spectral properties: the eigenvalues of the preconditioned matrix are highly clustered around the unity. As a result, it yields very rapid convergence when used to solve Toeplitz equations via the preconditioned conjugate gradient method. The new preconditioner can be regarded as a refinement of preconditioners built by embedding the Toeplitz matrix in a positive definite circulant. Necessary and sufficient conditions that ensure that the positive definite embedding is possible are given.**

***Index Terms*—PCG, preconditioners, Toeplitz matrices.**

I. Introduction

The stability of direct methods and the convergence rate of iterative methods for the solution of the linear problem $Tx = b$ depend on the condition number of T. This suggests the replacement of $Tx = b$ by an equivalent problem (for example, $P^{-1}Tx = P^{-1}b$) with better spectral properties (that is, a smaller condition number or a more favorable eigenvalue distribution) and hence numerically easier to solve. This process is known as preconditioning, and matrices such as P are called preconditioners [1].

A preconditioner P represents a trade-off between contradictory requirements: P should be much easier to invert than T, yet very close to it. The first condition ensures that the computational demands posed by preconditioning do not significantly add to the overall workload; as for the second condition, the idea is that if P is somehow "close" to T, then $P^{-1}T$ will be "close" to the identity matrix, and therefore well-conditioned.

Among the linear problems $Tx = b$, those with a Toeplitz matrix occur in many signal processing applications and have been the subject of much interest. In addition to $O(n^2)$ methods such as Levinson's [2], $O(n \log^2 n)$ methods have been described [3]. However, the conjugate gradient method, introduced over half a century ago by Hestenes and Stiefel [2], [4], is among the methods most often used. The required Toeplitz multiplications can be performed efficiently using the FFT, and preconditioning can speed up the convergence rate very effectively, yielding a method that, in practice, is of overall computational complexity $O(n \log n)$.

The preconditioner should satisfy the following conditions.

1) For any vector v, the product $P^{-1}v$ should be easily computed (in at most $O(n \log n)$ time).
2) $P^{-1}T \approx I$ in the sense that *almost all* the eigenvalues of $P^{-1}T$ should be clustered around the unity.

Several preconditioners have been proposed for Toeplitz equations [5]–[13]. Their impact on the convergence rate of the preconditioned conjugate gradient method depends on the eigenvalue distribution of $P^{-1}T$. For example, when T is positive definite the error in the T-norm is $O(\alpha^m)$ after m iterations, where $\alpha = (\sqrt{\kappa} - 1/\sqrt{\kappa} + 1)$ and κ is the spectral condition number of T [2], [14]. The error bound decreases as $\sqrt{\kappa}\alpha^m$ [15, p. 154] in the Euclidean norm.

In the spectral norm, κ is the ratio of the largest and smallest eigenvalues; if the ratio is not too large (that is, if the matrix is well-conditioned) the method will converge rapidly. Preconditioning, which effectively replaces T by $P^{-1}T$, is one way of achieving that.

The convergence rate may in fact be far more rapid than the error bounds suggest. If the extremal eigenvalues of the preconditioned matrix $P^{-1}T$ are well separated, the method exhibits superlinear convergence [14]. This happens because the error components in the directions of the eigenvectors associated with these eigenvalues tends to be eliminated first, and then the method is able to proceed as if it was dealing with an increasingly well-conditioned matrix. This stands in contrast with the multilevel case [16], in which superlinear convergence cannot be achieved with circulant-like preconditioners. The results in [14] can also be used to show [5] that when the eigenvalues of the preconditioned matrix lie in $(1-\epsilon, 1+\epsilon)$ except for r outliers, the error in an appropriate norm will decrease at least by ϵ^2 per iteration after r initial iterations.

The goal of this paper is to introduce a new preconditioner for Toeplitz matrices which improves the performance of the preconditioned conjugate gradient method due to its excellent eigenvalue clustering properties. The preconditioner is defined by

$$P^{-1} = C_1(2I - TC_1)$$

where C_1 is the matrix in (2) below. Alternative expressions are given in Section II-A [see (4)–(6)]. It will be shown that the spectrum of $P^{-1}T$ tends to be extremely concentrated around the unity, leading to very rapid convergence of the preconditioned conjugate gradient method.

II. The Preconditioner

Any Toeplitz matrix T can be embedded in a circulant matrix of the form

$$C = \begin{pmatrix} T & S \\ S & T \end{pmatrix} \tag{1}$$

where $T_{ij} = t_{i-j}$, $S_{ij} = s_{i-j}$, $s_i = t_{i-n}$ for $i > 0$ and $s_i = t_{i+n}$ for $i < 0$. The element s_0 is arbitrary.

The embedding idea is of interest for at least two reasons. First, it is possible to solve the Toeplitz equations by using the larger circulant set [17], effectively trading-off matrix size by

Manuscript received April 03, 2009; revised May 14, 2009. First published June 05, 2009; current version published July 01, 2009. This work was supported in part by MICINN (Spain) under Grant José Castillejo. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Andreas Jakobsson.

M. E. Domínguez-Jiménez is with the GI TACA in the Department Matemática Aplicada, ETSII, Universidad Politécnica de Madrid, 28006 Madrid, Spain (E-mail: edominguez@etsii.upm.es).

P. J. S. G. Ferreira is with the Signal Processing Laboratory, DETI/IEETA, Universidade de Aveiro, 3810-193 Aveiro, Portugal (E-mail: pjf@ua.pt).

Digital Object Identifier 10.1109/LSP.2009.2024735

matrix structure. Second, the eigenvalues of C can be computed using the FFT and yield bounds for the eigenvalues of T. In this case, s_0 can be chosen to make the bounds as tight as possible [18]. In fact, the connection between the distribution of the eigenvalues of Toeplitz and circulant matrices is a classic theme [19]; a recent work on the block multilevel case is [20].

We write C^{-1}, which is also circulant, as

$$C^{-1} = \begin{pmatrix} C_1 & C_2 \\ C_2 & C_1 \end{pmatrix} \tag{2}$$

so that

$$\begin{aligned} TC_2 = -SC_1, \quad TC_1 + SC_2 = I, \\ C_2T = -C_1S, \quad C_1T + C_2S = I \end{aligned}. \tag{3}$$

From this point onwards we assume that T is positive definite (we denote this by $T > 0$). The need to identify the conditions under which $T > 0$ can be embedded in a positive definite C has not been adequately recognized in the literature. The usual condition $T > 0$ does not imply $C > 0$ (although, obviously, $C > 0$ does imply $T > 0$). For example, the condition $C > 0$ explicitly mentioned in our Remarks 1 and 2 below cannot be omitted nor replaced by $T > 0$. We discuss this issue in Section II-B, which contains necessary and sufficient conditions for $C > 0$ that can be efficiently tested.

We now examine the potential of certain matrices involving T, C and C^{-1} as preconditioners for T. As in [5], [8], we take $s_0 = 0$ except when otherwise stated.

Remark 1: If C is positive definite, then the matrix $T + S$ is a good preconditioner for T.

The essence of the argument is in [5]: the matrix $K_1 = T + S$ is circulant, so the problem of inverting it or applying it to a vector can be solved efficiently using the FFT in $O(n \log n)$ flops. On the other hand, $K_1^{-1}T = (I + T^{-1}S)^{-1}$ has eigenvalues clustered around the unity.

Remark 2: If C is positive definite, then the matrix C_1^{-1} is a good preconditioner for T.

To see that this is true, first note that C_1 is also positive definite and can be computed in $O(n \log n)$ flops using the FFT. Second, by using $C_1^{-1} = T - ST^{-1}S$ it follows that

$$T^{-1}C_1^{-1} = T^{-1}\left(T - ST^{-1}S\right) = I - \left(T^{-1}S\right)^2.$$

But, as seen in Remark 1, the eigenvalues of $(I + T^{-1}S)^{-1}$ are clustered around the unity, and so those of $T^{-1}S$ are clustered around zero. The same goes for the eigenvalues of $\left(T^{-1}S\right)^2$, since they are the squares of the former. Hence, the eigenvalues of $T^{-1}C_1^{-1}$ are clustered around 1, and it happens the same to its inverse C_1T. This shows that C_1^{-1} is indeed a preconditioner for T.

In fact, C_1^{-1} is one of the preconditioners discussed in [8].

A. The Proposed Preconditioner

The fact that C_1 is already "close" to T^{-1} suggests that further corrections to C_1^{-1} may lead to even better spectral behavior. Since $C_1^{-1} = T - ST^{-1}S$ and $T^{-1} \approx C_1$, we consider the possibility

$$\begin{aligned} M &= T - SC_1S \\ &= T + TC_2S = 2T - TC_1T \end{aligned}$$

where the possible preconditioner M was written in three equivalent ways, using (3). To check if M is an adequate preconditioner we need to investigate its spectral properties and the computational procedure. Recalling that C_1T has eigenvalues clustered around the unity, the expression

$$T^{-1}M = I + C_2S = 2I - C_1T$$

shows that the same happens with the eigenvalues of $2I - C_1T = T^{-1}M$. From the spectral point of view M is a potential preconditioner, but it has a serious drawback: M^{-1} is not easily computed. For this reason, we abandon M and seek an alternative.

The fact that M itself is easier to compute suggests one possible path. We note that M can be expressed in terms of Toeplitz matrices as $M = T - SC_1S$. Moreover, the expression

$$M = 2T - TC_1T = T\left(2I - C_1T\right)$$

involves only T (not S) and C_1, and so the evaluation of $Mz = T\left(2z - C_1Tz\right)$ requires only three Toeplitz products.

Exchanging T with C_1, and S with C_2, and noting that now $T^{-1} = C_1 - C_2C_1^{-1}C_2$, we arrive at the following new potential approximation to the inverse of T:

$$N = C_1 - C_2TC_2. \tag{4}$$

Note that, using (3), N can also be written as

$$N = C_1 + C_2SC_1 = 2C_1 - C_1TC_1 = C_1\left(2I - TC_1\right). \tag{5}$$

This corresponds to the new preconditioner

$$P = N^{-1} = \left(C_1 - C_2TC_2\right)^{-1} \tag{6}$$

which is a correction to the initial preconditioner C_1^{-1}.

Now, the action of $P^{-1} = N$ is easily computed: it takes three Toeplitz products to compute $Nz = C_1\left(2z - TC_1z\right)$, and two of the three products involve the same matrix C_1. As for the eigenvalues of $P^{-1}T = NT$, note that

$$NT = 2\left(C_1T\right) - \left(C_1T\right)^2 = \left(C_1T\right)\left(2I - C_1T\right).$$

We have seen that the eigenvalues of C_1T are clustered around 1, and NT is a deformation $\mu\left(2 - \mu\right)$ of that matrix. So, its eigenvalues are also clustered around 1.

In fact, $NT = C_1T - \left(C_2T\right)^2$ which suggests that N is a correction to C_1. Before confirming its excellent performance through numerical simulations, we address one issue that has been overlooked in the literature: the conditions under which T can be embedded in a positive definite circulant.

B. When is C Positive Definite?

Let T be a positive definite Toeplitz $n \times n$ matrix. Let C be the family of $2n \times 2n$ circulants given by (1), which depend on the parameter s_0.

We write the eigenvalues of C as $\{\lambda_i(C)\}_{0 \le i < 2n}$, where $\lambda_i(C)$ is the ith element of the FFT of the first row of C; this establishes an ordering of the eigenvalues. We denote by C_0 the specific circulant matrix C obtained by taking $s_0 = 0$. The easily computed quantities

$$L_0 = \min_{0 \le i < n} \lambda_{2i}(C_0), \quad L_1 = \min_{0 \le i < n} \lambda_{2i+1}(C_0)$$

play the crucial role in the following theorem.

Theorem 1: Let T be a positive definite Toeplitz matrix. Then T can be embedded in a positive definite circulant C given by (1) if and only if $L_0 + L_1 > 0$. If this condition holds, C will be positive definite if and only if s_0 is chosen in the interval $(-L_0, L_1)$.

Proof: The elementary equality

$$C = \frac{1}{2}\begin{pmatrix} I & I \\ I & -I \end{pmatrix}\begin{pmatrix} T+S & 0 \\ 0 & T-S \end{pmatrix}\begin{pmatrix} I & I \\ I & -I \end{pmatrix}$$

shows that C is unitarily equivalent to a diagonal block matrix with blocks $T+S$ and $T-S$. Hence, its eigenvalues are the union of those of $T+S$ and $T-S$.

The eigenvalues of $T+S$ are $\{\lambda_{2i}(C)\}_{0 \le i < n}$ and those of $T-S$ are $\{\lambda_{2i+1}(C)\}_{0 \le i < n}$, as can be shown [18] by expressing the FFT of length $2n$ as two FFTs of length n.

These eigenvalues are real, since C is hermitian under the conditions of the theorem.

The matrix S is Toeplitz and by construction the element in its main diagonal is s_0. To simplify the notation we set $\alpha = s_0$, so that $S = S_0 + \alpha I$, where the main diagonal elements of S_0 are all zero.

Now, $T+S$ and $T-S$ can be written as $T + S_0 + \alpha I$ and $T - S_0 - \alpha I$, the eigenvalues of which are $\lambda_{2i}(C_0) + \alpha$ and $\lambda_{2i+1}(C_0) - \alpha$.

Clearly, C will be positive definite if and only if these numbers are all positive, that is, if and only if $-\lambda_{2i}(C_0) < \alpha$ and $\alpha < \lambda_{2i+1}(C_0)$ for all i, that is, if $-L_0 < \alpha < L_1$. The interval $(-L_0, L_1)$ is nonempty if and only if $L_0 + L_1 > 0$, as asserted. ■

III. Numerical Experiments

We have considered several systems $Tx = b$ where T is a positive definite Toeplitz matrix, and $b = [1, \ldots, 1]^T$. For each problem, several preconditioners have been applied: Strang's and Chan's [8, Section II], the circulant preconditioner K_1 [5], $\left(P^{(4)}\right)^{-1}$ [6], [8], the Toeplitz preconditioner C_1^{-1} (which is the same as $\left(P^{(2)}\right)^{-1}$ in the terminology of [6], [8]), and the proposed preconditioner $P = N^{-1}$. Recall that all these preconditioners can be applied in $O(n \log n)$ flops: circulant preconditioners require two FFTs of length n per iteration, Toeplitz preconditioners require two FFTs of length $2n$, and the proposed preconditioner requires four additional FFTs. In all cases we have set $s_0 = 0$, as done in [8]. It would be possible to further improve the performance of P by tuning s_0, but for brevity we did not do so.

1) Problem 1: The entries of T are $t_k = (1 + |k|)^{-1.1}$. This problem is also studied in [6]. For $n = 100$, the eigenvalues of $P^{-1}T$ are displayed in Fig. 1; we can see that most of them are clustered around 1. The best preconditioner for the conjugate gradient method is the proposed preconditioner $P = N^{-1}$, as depicted in Fig. 2. Table I shows the results as a function of the matrix size n.

2) Problem 2: We also considered a random positive definite Toeplitz matrix T of size 100 and studied the behavior of the considered preconditioners. Fig. 3 displays the corresponding eigenvalue clusters, and Fig. 4 shows the performance of the conjugate gradient algorithm for the problem $Tx = b$. As in the previous case, the best results are obtained with the proposed preconditioner $P = N^{-1}$. Very often, a random matrix will not fulfill the necessary and sufficient condition of Theorem 1. Hence, it cannot be embedded in a circulant positive definite matrix. For this reason, neither Remark 1 nor Remark 2 apply, and nothing can be said about the preconditioners K_1 and C_1^{-1} *a priori*. The matrix in Problem 2 is an example of this. However, even in this case, the proposed preconditioner $P = N^{-1}$ yields the best results.

3) Problem 3: The matrix T is defined by the Fourier coefficients of x^2 in $[0, 2\pi]$. This matrix is very ill-conditioned, and the relative performance of each preconditioner varies as the matrix size grows. Fig. 5 illustrates the performance of the

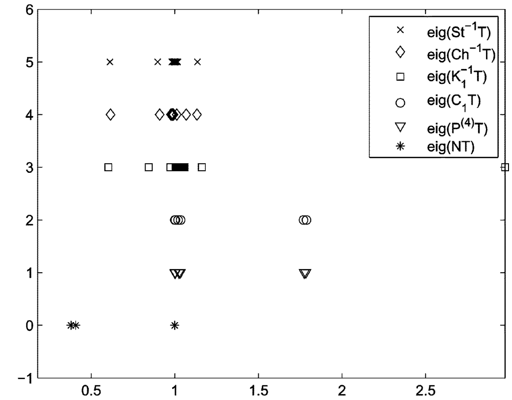

Fig. 1. Problem 1: eigenvalues $\lambda_i(P^{-1}T)$ for each of the considered preconditioners P: Strang's, Chan's, K_1, C_1^{-1}, $(P^{(4)})^{-1}$ and the proposed preconditioner N^{-1}.

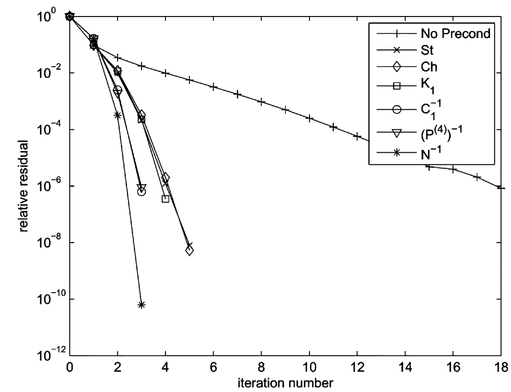

Fig. 2. Problem 1: Convergence of the preconditioned conjugate gradient method for each of the considered preconditioners.

TABLE I
Problem 1. Problem Size and Number of Iterations for Each of the Considered Preconditioners

n	No Precond	St	Ch	K_1	C_1^{-1}	$(P^{(4)})^{-1}$	N^{-1}
100	18	5	5	4	3	3	3
200	23	5	5	5	4	4	3
300	25	5	5	5	4	4	3
400	26	5	5	5	4	4	3
500	27	5	5	5	4	4	3
1000	30	5	5	5	4	4	3

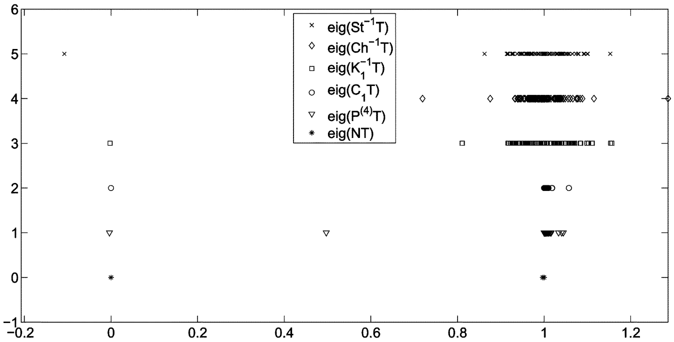

Fig. 3. Problem 2: Eigenvalues $\lambda_i(P^{-1}T)$ for each of the considered preconditioners P: Strang's, Chan's, K_1, C_1^{-1}, $(P^{(4)})^{-1}$ and the proposed preconditioner N^{-1}.

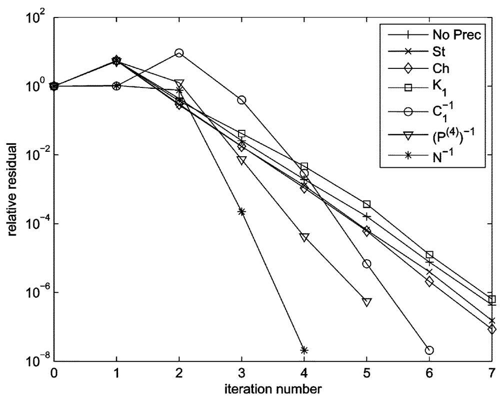

Fig. 4. Problem 2: Convergence of the preconditioned conjugate gradient method for each of the considered preconditioners.

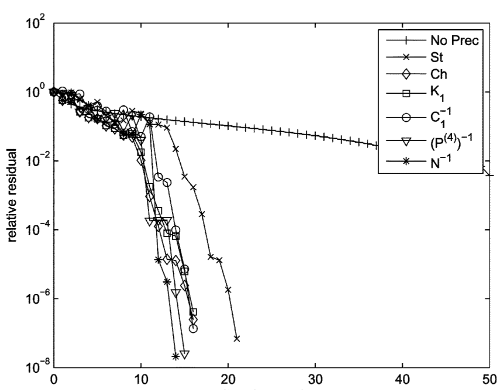

Fig. 5. Problem 3: Convergence of the preconditioned conjugate gradient method for each of the considered preconditioners.

preconditioners for $n = 50$. The proposed preconditioner required the least number of iterations. However, the design of an effective preconditioner for very ill-conditioned Toeplitz matrices across a wide range of sizes remains open.

IV. Conclusions

The principal matrix C_1 of the circulant extension of a Toeplitz matrix

$$C = \begin{pmatrix} T & S \\ S & T \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} C_1 & C_2 \\ C_2 & C_1 \end{pmatrix}$$

is an approximation to the inverse of T and has been used as a preconditioner. We proposed the alternative approximation

$$N = C_1 - C_2 T C_2,$$

which corresponds to the preconditioner

$$P = N^{-1} = (C_1 - C_2 T C_2)^{-1}.$$

This preconditioner has excellent spectral properties and when used with the PCG method it leads to very fast convergence rate and an overall computational workload of $O(n \log n)$ in practice. We have confirmed this with numerical experiments.

The new preconditioner, as well as those against which it is compared, is effective when C is positive definite. To be able to test this, we gave a necessary and sufficient condition under which this holds, which can be efficiently tested.

References

[1] K. Chen, *Matrix Preconditioning Techniques and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[2] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
[3] D. Bini and V. Y. Pan, *Polynomial and Matrix Computations*. Boston, MA: Birkhäuser, 1994.
[4] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bureau Stand.*, vol. 49, no. 6, pp. 409–436, Dec. 1952.
[5] T.-K. Ku and C.-C. J. Kuo, "Design and analysis of Toeplitz preconditioners," *IEEE Trans. Signal Process.*, vol. 40, pp. 129–141, Jan. 1992.
[6] R. H. Chan and K.-P. Ng, "Toeplitz preconditioners for hermitian Toeplitz-systems," *Linear Algebra Applicat.*, vol. 190, pp. 181–208, 1993.
[7] T. Huckle, "Some aspects of circulant preconditioners," *SIAM J. Sci. Comput.*, vol. 14, no. 3, pp. 531–541, May 1993.
[8] R. H. Chan and M. K. Ng, "Conjugate gradient methods for Toeplitz systems," *SIAM Rev.*, vol. 38, no. 3, pp. 427–482, Sep. 1996.
[9] T. Huckle, "Iterative methods for ill-conditioned Toeplitz matrices," *Calcolo*, vol. 33, pp. 177–190, 1996.
[10] S. Serra, "Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems," *Math. Comput.*, vol. 66, no. 218, pp. 651–665, Apr. 1997.
[11] D. Potts and G. Steidl, "Preconditioners for ill-conditioned Toeplitz matrices," *BIT*, vol. 39, no. 3, pp. 513–533, 1999.
[12] R. H. Chan, A. M. Yip, and M. K. Ng, "The best circulant preconditioners for hermitian Toeplitz systems," *SIAM J. Numer. Anal.*, vol. 38, no. 3, pp. 876–896, 2000.
[13] R. H. Chan, M. K. Ng, and A. M. Yip, "The best circulant preconditioners for hermitian Toeplitz systems II: The multiple-zero case," *Numer. Math.*, vol. 92, pp. 17–40, 2002.
[14] D. G. Luenberger, *Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley, 1984.
[15] T. Kailath and A. H. Sayed, *Fast Reliable Algorithms for Matrices With Structure*. Philadelphia, PA: SIAM, 1999.
[16] S. S. Capizzano and E. Tyrtyshnikov, "Any circulant-like preconditioner for multilevel matrices is not superlinear," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 2, pp. 431–439, 1999.
[17] P. J. S. G. Ferreira and M. E. Domínguez, "Trading-off matrix size and matrix structure: Handling Toeplitz Equations by embedding on a larger circulant set," , submitted for publication.
[18] P. J. S. G. Ferreira, M. Blanke and T. Söderström, Eds., "Localization of the eigenvalues of Toeplitz matrices using additive decomposition, embedding in circulants, and the Fourier transform," in *Proc. SysID'94, 10th IFAC Symposium on System Identification*, Copenhagen, Denmark, Jul. 1994, vol. III, pp. 271–276.
[19] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Los Angeles, CA: Chelsea, 1958.
[20] M. Oudin and J. P. Delmas, "Asymptotic generalized eigenvalue distribution of block multilevel Toeplitz matrices," *IEEE Trans. Signal Process.*, vol. 57, pp. 382–387, Jan. 2009.