MULTIGRID METHODS FOR BLOCK TOEPLITZ MATRICES WITH SMALL SIZE BLOCKS

THOMAS HUCKLE[1] and JOCHEN STAUDACHER[2]

[1] *Institut für Informatik, Technische Universität München*
D-80290 München, Germany.
email: huckle@in.tum.de

[2] *School of Engineering, University of Durham*
Durham DH1 3LE, England.
email: j.h.staudacher@durham.ac.uk, staudacj@in.tum.de
Correspondence to Thomas Huckle.

Abstract.

In this paper we discuss multigrid methods for ill-conditioned symmetric positive definite block Toeplitz matrices. Our block Toeplitz systems are general in the sense that the individual blocks are not necessarily Toeplitz, but we restrict our attention to blocks of small size. We investigate how transfer operators for prolongation and restriction have to be chosen such that our multigrid algorithms converge quickly. We point out why these transfer operators can be understood as block matrices as well and how they relate to the zeroes of the generating matrix function. We explain how our new algorithms can also be combined efficiently with the use of a natural coarse grid operator. We clearly identify a class of ill-conditioned block Toeplitz matrices for which our algorithmic ideas are suitable. In the final section we present an outlook to well-conditioned block Toeplitz systems and to problems of vector Laplace type. In the latter case the small size blocks can be interpreted as degrees of freedom associated with a node. A large number of numerical experiments throughout the article confirms convincingly that our multigrid solvers lead to optimal order convergence.
Key words: Multigrid methods, iterative methods, preconditioning, block Toeplitz matrices, vector valued problems.

1 INTRODUCTION

1.1 Toeplitz matrices and generating functions

Let $f(x)$ be a real-valued continuous function on the interval $I = [-\pi, \pi]$ and periodically extended to the whole real axis. On the basis of its Fourier coefficients

$$t_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} d\theta, \qquad \text{for } k \text{ integer},$$

we can define the sequence of Toeplitz matrices $\{T_n \equiv T_n(f)\}_{n\in\mathbf{N}}$ associated with the generating function $f(x)$. Its entries are given by $(T_n)_{\mu,\nu} = t_{\mu-\nu}$:

$$T_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & \cdots & t_{1-n} \\ t_1 & t_0 & t_{-1} & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & t_1 & t_0 & t_{-1} \\ t_{n-1} & \cdots & \cdots & t_1 & t_0 \end{pmatrix}$$

Since $f(x)$ is real-valued the matrices T_n are Hermitian. In case $f(x)$ is an even function, we are dealing with a sequence of real symmetric Toeplitz matrices. Furthermore, we know that the spectrum of T_n is contained in $range(f)$.

Example 1: The well-known matrix $tridiag(-0.5, 1, -0.5)$ – which is nothing but a uniform discretization of the one-dimensional Laplacian – is generated by the function $f(x) = -0.5e^{-ix} + 1 - 0.5e^{ix} = 1 - \cos(x)$. The eigenvalues of T_n are contained in the interval $[0, 2]$. The small eigenvalues of T_n that lead to the large condition numbers are caused by the zero $x_0 = 0$ of f, $f(x_0) = f(0) = 0$, of multiplicity two.

For more information on Toeplitz matrices and their properties we refer to the book [7] and the overview article [2].

1.2 Block Toeplitz Matrices with fixed block size

The focus of this article lies upon symmetric block Toeplitz matrices

$$A_n = \begin{pmatrix} T_0 & T_{-1} & \cdots & T_{1-n} \\ T_1 & T_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{-1} \\ T_{n-1} & \cdots & T_1 & T_0 \end{pmatrix} \tag{1.1}$$

with $k \times k$ matrices T_j. Just like in the Toeplitz case we will assume that they are generated by the matrix function

$$F(x) = \cdots + T_{-2}e^{-2ix} + T_{-1}e^{-ix} + T_0 + T_1e^{ix} + T_2e^{2ix} + \cdots .$$

Furthermore, we would like to stress at this point that the individual blocks T_j are not necessarily Toeplitz and that we assume the block size k to be fixed.

Example 2: Blocks T_j in (1.1) and related generating matrix function given by

$$T_0 = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}, \; T_{-1} = \begin{pmatrix} 0 & 0 \\ 1 & 10 \end{pmatrix}, \; F(x) = \begin{pmatrix} 3 & -1 + e^{ix} \\ -1 + e^{-ix} & 2 + 20\cos(x) \end{pmatrix}.$$

As in the scalar case we can shift a singularity of the generating matrix function $x_0 \in]-\pi, \pi]$ to the origin. This will come out to be very useful later in our article.

Lemma 1: Consider the generating $k \times k$ matrix function $F(x)$ for the block Toeplitz matrix $T_n(F(x))$. Let $det(F(x_0)) = 0$ for $x_0 \in]-\pi, \pi]$. Then the diagonal transformation $D^H * T_n(F(x)) * D$ with $D_0 = exp(ijx_0) * I_k$ and $D = diag((D_0^j)_{j=0}^{n-1})$ leads to the block Toeplitz matrix $T_n(F(x - x_0)) = D^H * T_n(F(x)) * D$, where the singularity x_0

of the generating matrix function is shifted to the origin.

The following lemma also points out an analogy to the Toeplitz case (cf. [11], [12]).

Lemma 2: Let $F(x)$ and $P(x)$ be continuous $k \times k$-matrix functions in $[-\pi, \pi]$ that are generating functions for the block Toeplitz matrices $T_n(F(x))$ and $T_n(P(x))$. Furthermore, let us assume that all entries in $P(x)$ are trigonometric polynomials with maximum degree r. Then it holds

$$T_n(P(x) * F(x)) = T_n(P(x)) * T_n(F(x)) + R ,$$

and the matrix R is of rank $O(k * r)$.
As an immediate corollary with $G(x) := P(x) * F(x)$ it follows that

$$T_n(P(x))^{-1} * T_n(G(x)) = T_n(P(x)^{-1} * G(x)) + \tilde{R} ,$$

with $rank(\tilde{R}) = O(k*r)$ in case the generating matrix functions $G(x)$ and $P(x)^{-1}G(x)$ are continuous. Hence, if for a given ill-conditioned $G(x)$ we can find a trigonometric polynomial such that $P(x)^{-1}G(x)$ is well-conditioned, then we can expect $T_n(P(x))$ to be an efficient preconditioner for $T_n(G(x))$.

1.3 Multigrid algorithms

For many classes of linear systems, multigrid methods are renowned to be among the fastest iterative solvers. Frequently, their computational complexity is of the same order as the multiplication of the system matrix with a vector, i.e., $O(n)$ for a sparse matrix and $O(n \log(n))$ in the case of a Toeplitz or block Toeplitz matrix, where n denotes the number of unknowns in the linear system.

In the following we would briefly like to remind the reader of the basic concepts by giving a compact version of a multigrid cycle for the iterative solution of a linear system $Ax = b$. For more algorithmic details see, e.g., the books by Greenbaum [6], pp. 193, or Briggs [1], pp. 48.

Algorithm 1: Solving $Ax = b$ iteratively by a **multigrid cycle**:
Proceed with the following iteration until the stopping criterion is satisfied.
(a) Smooth (e.g., by the Richardson method) in order to get a new iterate $x^{(j)}$.
(b) Compute the residual $r = Ax^{(j)} - b$.
(c) Restrict the residual $r^{coarse} = Rr$, using the restriction operator R.
(d) Set up the coarse grid matrix, e.g., via the Galerkin approach $A^{coarse} = RAP$. (Hereby P denotes the prolongation operator.)
(e) Solve the residual equation $A^{coarse}y = r^{coarse}$ on the coarse grid – if not already on coarsest level, then apply the multigrid cycle recursively.
(f) Update $x^{(j)} := x^{(j)} - Py$ using the prolongation operator P.

If within the recursive solution in step (e) we use one cycle, we get a so-called V-cycle algorithm. By applying two cycles, we recognize the W-cycle algorithm.

As it is important for the rest of the paper we would also briefly like to comment on the two most common choices for setting up coarse grid matrices A^{coarse} in step (d): As we already mentioned, Galerkin coarsening can be used. In that case, there usually

holds $R = P^T$ and thus $A^{coarse} = P^T AP$.
In certain situations, an alternative is to rediscretize the given problem on a coarser grid. Historically, this coarsening strategy is older than the Galerkin approach and it is often referred to as employing a natural coarse grid operator.
For example, given a uniform discretization of the one-dimensional Laplacian with meshsize h

$$A_h = (1/h^2) * tridiag(-1, 2, -1)$$

we can obtain a natural coarse grid operator A_{2h} on a mesh with size $2h$ as

$$A_{2h} = (1/2h)^2 * tridiag(-1, 2, -1).$$

Observe that the fine grid matrix A_h and the coarse grid matrix A_{2h} bear the different factors $(1/h^2)$ and $(1/2h)^2$, respectively. In practical implementations it is normally preferred to multiply the discrete equation by the factor h^2. Thus care needs to be taken in order to use the correct factor $4h^2$ on the coarse grid. In other words: We must not forget to multiply r^{coarse} in step (c) with the appropriate factor 4.
For more details on natural coarse grid operators, we refer to the book [18], ch. 2, or to the doctoral thesis of the second author [15], ch. 2.

1.4 Multigrid methods and Toeplitz matrices

As we have just pointed out, we need to apply a restriction and a prolongation operator in multigrid algorithms: If we use a Galerkin coarsening we can write the coarse grid matrix for a twogrid step as

$$T_{n/2} = I^T_{n,n/2} * B^H_n * T_n * B_n * I_{n,n/2} = P^H_n * T_n * P_n \tag{1.2}$$

with a Toeplitz matrix B_n related to a function $b(x)$, and the elementary projection matrix

$$I_{n,n/2} = \begin{pmatrix} 1 & & \\ 0 & 0 & \\ 0 & 1 & 0 \\ & 0 & 0 \\ & 0 & 1 \\ & & . \end{pmatrix} = I(1:n, 1:2:n)$$

in MATLAB-notation with the identity matrix I. As a quick and simple motivation let us make use of the following heuristics: With $\tilde{f}(x) = \overline{b(x)} * f(x) * b(x)$ the entries of the matrix $B^H_n * T_n * B_n$ in our model are – up to a perturbation of low rank – given by the coefficients of $\tilde{f}(x)$; therefore the coefficients of $T_{n/2}$ can be found by deleting every second entry in $\tilde{f}(x)$:

$$f_2(x) = (1/2) * \left(|b(\frac{x}{2})|^2 f(\frac{x}{2}) + |b(\frac{x}{2} + \pi)|^2 f(\frac{x}{2} + \pi) \right) . \tag{1.3}$$

(Note that the above formula actually comes from a convolution argument and thus – as pointed out e.g. in [4], [5], [14] – it holds in particular also for matrix algebras.)
Let us assume that $f(x) \geq 0$ has a unique zero x_0 of finite order 2κ in the interval $]-\pi, \pi]$. Now the new matrix $T_{n/2}$ should be closely related to the original T_n. Hence the related function $f_2(x)$ should have a zero with the same multiplicity as $f(x)$. In view of $f(x) \geq 0$ this is only possible if $b(x_0 + \pi) = 0$. Therefore, we can easily motivate to use a prolongation operator of the form

$$b(x) \; = \; (\cos(x_0) + \cos(x))^\kappa. \tag{1.4}$$

Multigrid methods for Toeplitz systems were first proposed by Serra and Fiorentino in [4], [5] – and it was them who first came up with (1.3) and showed that transfer operators corresponding to (1.4) are suitable. More recently, Serra [13] gave a precise analysis of twogrid optimality suggesting lower and upper bounds for κ in (1.4) . In these papers the focus lies on symmetric positive definite problems generated by functions with a single isolated zero $x_0 \in]-\pi, \pi]$.
In [16] and [3] R. Chan and collaborators also studied multigrid for Toeplitz systems: The work [3] presents solutions for symmetric positive definite Toeplitz problems with entries $t_1 = \ldots = t_l = 0$, $l < n$, like, for example, the Toeplitz matrices generated by $f(x) = 1 - \cos((l+1) * x)$. This article is partly based on results disseminated in [3]: In particular, we will present a block interpretation of the transfer operators suggested in [3] and explain how and why these ideas can be extended to certain block Toeplitz matrices with small size blocks.
All the articles mentioned so far employ Galerkin coarse grid operators: However, this may in general result in a loss of Toeplitz structure on the coarser levels. In a recent research paper [10] the authors presented a very straightforward resort for this difficulty using a natural coarse grid operator, i.e. the coarse level representations are nothing but Toeplitz matrices of smaller size generated by the original function f. There we pointed out how natural coarse grid operators can be employed efficiently for non-negative generating functions with a finite number of equidistant zeroes in $]-\pi, \pi]$ – and we note that for this approach the systems need to be diagonally scaled in advance such that one of the zeroes of highest order is shifted to the origin (see [10], sec. 2 and sec. 3, for details). Furthermore, we pointed out in [10] how the natural coarse grid operator needs to be accompanied by appropriate scaling of the residuals and how the scaling factor reflects the orders of the zeroes of the generating functions. (See also [15], ch. 4, for a detailed discussion.) For example, in the case of the one-dimensional Laplacian mentioned at the end of subsection 1.3, the scaling factor $2^2 = 4$ reflects the fact that the zero of the generating function $f(x) = 1 - \cos(x)$ is of order 2.

1.5 Brief outline of this article and a few preliminary remarks

This article proposes new multigrid algorithms for block Toeplitz matrices with small size blocks and it identifies a class of matrices for which our algorithms are very well suited.
We are developing our key ideas in three steps aiming to make the presentation clear and understandable: Therefore we first define a generating matrix function to be simple if it is symmetric non-negative definite and if all of its individual eigenvalue functions have only got a single isolated zero of order at most 2 in the interval $]-\pi, \pi]$. We then start taking a look at block Toeplitz matrices with simple diagonal generating matrix functions. We propose an efficient multigrid strategy employing a natural coarse grid operator and transfer operators for prolongation and restriction based on block matrices for the situation that all the functions on the diagonal have zeroes of order at most 2 at the origin. Then we take a look at more general simple generating matrix functions (which can be diagonalized, but are not necessarily diagonal) restricting our focus to the case where all singularities are at the same point. We see that our algorithmic approach from the previous section carries over without any difficulties if the zeroes of the eigenvalues of the underlying matrix function are all at the origin.
In the fourth section of the paper, we then forgo the restraint of zeroes to be located at the origin. We point out how our algorithmic ideas including the natural coarse grid matrices can be extented to more general situations with multiple zeroes in $]-\pi, \pi]$ of

orders at most 2 among the eigenvalues of the generating matrix function by proper "redoubling" of the latter, i.e. we explain how generating matrix functions can be made simple in the sense of our definition by expanding them. We demonstrate the efficiency of our new multigrid approach via a large number of numerical experiments.
In the final section we show that the new concepts may prove to be very useful for more general block Toeplitz problems: First, we point out that the prolongation and restriction operators based on block matrices we identify are also helpful for well-conditioned symmetric positive block Toeplitz matrices. (Such systems have previously been studied by other authors e.g. in [12]). Then we also take a look at problems of vector Laplace type and see how the small size blocks may be thought of as degrees of freedom associated with a node. Our algorithms work well not only in the sparse and separable case of the vector Laplacian, but also for dense non-separable examples with related properties. We end the article summarizing our conclusions and presenting a brief outlook.
A particularly strong emphasis of this work lies upon comprehensive numerical tests. As in [10], we always employ the following stopping criterion to obtain the iteration counts we list in our tables:

$$\frac{\|r^{(j)}\|_\infty}{\|r^{(0)}\|_\infty} \leq 10^{-6}.$$

Here $r^{(j)}$ denotes the residual after j iterations and $r^{(0)}$ the original residual, i.e., we stop iterating when the relative residual corresponding to the maximum norm is less than or equal 10^{-6}. In all our numerical experiments we use two steps of the Richardson method for pre- and post-smoothing in our multigrid cycles. Following the considerations for the Toeplitz case from [4], [5] and [10] , we use the damping parameters $\omega_1 = 1/\max_{\theta\in[-\pi,\pi]} \|F(\theta)\|_\infty$ for pre-smoothing and $\omega_2 = 2/\max_{\theta\in[-\pi,\pi]} \|F(\theta)\|_\infty$ for post-smoothing, respectively. We report iteration counts for W-cycles as solvers. The coarsest level is independent of the matrix size: We always fix it to be second smallest matrix possible, i.e. in case of a 2-by-2 generating function our coarsest level matrix would be 4-by-4.
Finally, we would like to stress that this is the first paper studying multigrid in the case that the individual blocks T_j are not necessarily Toeplitz. In other words: Our investigations are very much different from the studies on multigrid for block Toeplitz matrices with Toeplitz blocks in [5] or [17] (– where the block size k was variable, but the blocks were Toeplitz).

2 BLOCK TOEPLITZ MATRICES WITH DIAGONAL GENERATING FUNCTIONS

In this paper we are in principle only concerned with the case that the generating matrix functions of our block Toeplitz systems have singularities of order at most 2.
For the course of our presentation we would like to introduce the following definition.

Definition 1: Let G be the k-by-k generating matrix functions of a sequence of symmetric positive definite block Toeplitz matrices. We call the matrix function G **simple** if it satisfies the following two conditions:
1. It is symmetric and non-negative definite.
2. All of its individual eigenvalue functions have only got a single isolated zero in the interval $]-\pi,\pi]$ and the order of any of these zeroes is at most 2.

Note that Definition 1 implies that block Toeplitz matrices generated by simple matrix

functions are always diagonalizable.
As we are basically interested in real symmetric block Toeplitz systems, the reader may safely assume that the singularities are located either at 0 or π. In this section we focus on the case that the matrix functions only have entries on their main diagonal.

2.1 Multigrid transfer operators in the Galerkin case

Let us consider block Toeplitz matrices with n diagonal blocks $T_j = diag((T_{j,m})_{m=1}^k)$ of size $k \times k$. The generating function is then also a diagonal matrix function of the form

$$F(x) = \Lambda(x) = \begin{pmatrix} \lambda_1(x) & 0 & \cdots & 0 \\ 0 & \lambda_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k(x) \end{pmatrix}.$$

In the following we will identify two ways of defining multigrid algorithms employing Galerkin coarse grid operators for the case that the matrix function F is simple according to Definition 1.
Variant A: It is clear that by a straightforward permutation we can transform this block Toeplitz matrix into a block diagonal matrix with k Toeplitz matrices of size $n \times n$ as diagonal blocks. As we are assuming that $\lambda_m(x)$ has only got a single isolated zero $\tilde{x}_m \in]-\pi, \pi]$, we can obviously define a multigrid prolongation according to (1.4) related to the function $\lambda_m(x)$. That way we can reduce our block Toeplitz problem to the scalar Toeplitz case. For each $\lambda_m(x)$ we introduce a scalar function $b_m(x) = \cos(\tilde{x}_m) + \cos(x)$ such that for the single zero $\tilde{x}_m$ of $\lambda_m(x)$ there holds $b_m(\tilde{x}_m + \pi) = 0$.
In other words: The multigrid method for block Toeplitz matrices where the generating function is a diagonal matrix can be seen as k independent multigrid methods for scalar Toeplitz matrices.

For the course of our investigation, it will, however, prove to be much more interesting not to permute our block Toeplitz matrices and to keep their block structure as it is.
Variant B: Forgoing the permutation of the previous variant, the multigrid method for each block can also be interpreted as a multigrid algorithm for the given block Toeplitz matrix with prolongation defined by the block Toeplitz matrix generated by the function

$$B(x) = \begin{pmatrix} b_1(x) & 0 & \cdots & 0 \\ 0 & b_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & b_k(x) \end{pmatrix}, \text{where } I_{n,n/2,k} = \begin{pmatrix} I_k & & \\ 0_k & 0_k & \\ 0_k & I_k & 0_k \\ & 0_k & 0_k \\ & 0_k & I_k \\ & & . \end{pmatrix},$$

is the elementary projection matrix with I_k and 0_k denoting the $k \times k$ identity and the $k \times k$ matrix of zeroes, respectively.
Speaking more concretely, for singularities at 0 and π, we use n_h-by-n_{2h} prolongation

matrices of the form

$$I_{2h}^{h} = \begin{pmatrix} \pm\frac{1}{2}I_k & & \\ I_k & & \\ \pm\frac{1}{2}I_k & \pm\frac{1}{2}I_k & \\ & I_k & \ddots \\ & \pm\frac{1}{2}I_k & \ddots \\ & & \ddots \end{pmatrix}. \tag{2.1}$$

with n_h and n_{2h} standing for the matrix size on the fine and the coarse grid, respectively. As for $\pm\frac{1}{2}I_k$ we need to pick the sign according the scalar functions $\lambda_1, \ldots, \lambda_k$ and use plus in connection to with a zero at the origin and minus for a zero at π.

Example 3: For the block Toeplitz systems generated by

$$G_1 = \begin{pmatrix} 1+\cos(x) & 0 \\ 0 & 1-\cos(x) \end{pmatrix}$$

we employ block transfer operators related to

$$B_1 = \begin{pmatrix} 1-\cos(x) & 0 \\ 0 & 1+\cos(x) \end{pmatrix},$$

i.e. the term $\pm\frac{1}{2}I_2$ in the interpolation operator (2.1) takes a minus sign in the odd columns and a plus sign in the even columns.

Note that in this case there holds $B_1(x) = G_1(x+\pi)$. Note also for the sparse example G_1 the identical multigrid transfer operators would be derived using the authors' Matrix Multilevel Method [9] for sparse banded matrices.

Finally, we would also like to point out that transfer operators of the form (2.1) have first been suggested by R. Chan and collaborators in [3] for symmetric positive Toeplitz matrices generated by functions $f : [-\pi, \pi] \to \mathbb{R}$ satisfying

$$\max_{x\in[-\pi,\pi]} \frac{f(x)}{1 \pm \cos(k * x)} > 0. \tag{2.2}$$

We will come back to this idea and explore it in more detail in section 4.

2.2 An efficient algorithm with natural coarse grid operators

So far, we have only been studying Galerkin coarse grid operators. However, for dense block Toeplitz matrices this will not lead to efficient multigrid algorithms of optimal complexity $O(n\log(n))$, because the block Toeplitz structure will in general be lost on the coarse grids (and this would in turn prevent us from using block variants of the FFT for matrix vector multipications). We would also like to mention that in the block Toeplitz case the loss of structure through the Galerkin approach is usually more severe than for scalar Toeplitz matrices.

Our resort is to employ a natural coarse grid operator, i.e. our coarse grid matrices are block Toeplitz matrices of smaller size generated by the same matrix function as the original problem. However, we can – at this stage – only use this approach if all the singularities are located at the origin (or can be shifted there according to Lemma 1). We are briefly summarizing our algorithm as follows:

Algorithm 1: We would like to solve block Toeplitz systems generated by k-by-k diagonal matrix functions which are simple in the sense of Definition 1 and, in addition, have all their singularities at the origin. We are using a multigrid algorithm with a natural coarse grid operator. We are using transfer operators corresponding to the trigonometric function $b(x) = 1 + \cos(k * x)$, i.e. the prolongation operator is given by

$$P_k = \begin{pmatrix} \frac{1}{2}I_k & & \\ I_k & & \\ \frac{1}{2}I_k & \frac{1}{2}I_k & \\ & I_k & \ddots \\ & \frac{1}{2}I_k & \ddots \\ & & \ddots \end{pmatrix}. \qquad (2.3)$$

and (– just as for the one-dimensional Laplacian from Example 1 –) the restriction operator is given by $R_k = \frac{1}{2}P_k^T$.

As we pointed out in subsection 1.3, we need – due to the fact that we are employing a natural coarse grid operator – to scale the residuals with the appropriate factor $fac = 2^{s/k}$ where s denotes the sum of the orders of the zeroes of the scalar functions $\lambda_1, \ldots, \lambda_k$. (See [10] or [15] for more details.)

3 BLOCK TOEPLITZ MATRICES WITH SIMPLE GENERATING FUNCTIONS AND ONE SINGULARITY

As we have seen in the previous section, the case of diagonal generating matrix functions can always be lead back to the scalar Toeplitz case by straightforward permutations. Hence we are now studying block Toeplitz systems generated by matrix functions which are simple in the sense of Definition 1 but no longer diagonal. In order to keep the presentation clear and concise, we restrict our attention to the case that all singularities occur at the same point. We will see that Algorithm 1 is still applicable if all the singularities are located at the origin.

3.1 Algorithmic strategy for Galerkin and natural coarse grid matrices

For a general block Toeplitz matrix the range of the matrix $T_n(F(x))$ is given by the range of the generating matrix function $F(x)$, respectively by the range of the eigenvalues $\lambda_j(x)$ of $F(x) = U(x)^H \Lambda(x) U(x)$ – see [11], [12]. Hence, the condition number of $T_n(F(x))$ depends on the zeroes of scalar functions $\lambda_j(x) \geq 0$, and the prolongation and restriction operators have to be chosen in view of the location of the zeroes of the eigenvalues of $F(x)$. For motivation let us start with the case of a Galerkin coarse grid representation: Similar to the scalar case for a generating function $F(x)$ and a chosen prolongation/restriction operator – with generating function $B(x)$ and elementary projection $I_{n,n/2,k}$ – the Galerkin coarse grid matrix of half size

$$A_{n/2} = I_{n,n/2,k}^T * B_n^H * A_n * B_n * I_{n,n/2,k}$$

is "associated" (– i.e. following Lemma 2 we simply ignore low rank perturbations for a moment –) to the matrix function

$$F_2(x) = (1/2) * \left(B(\frac{x}{2})^H F(\frac{x}{2}) B(\frac{x}{2}) + B(\frac{x}{2}+\pi)^H F(\frac{x}{2}+\pi) B(\frac{x}{2}+\pi) \right) .$$

To derive the appropriate multigrid transfer operators, we want to reduce the problem to the scalar case which is certainly possible if the two addends in $F_2(x)$ can be treated like diagonal matrices.
Let us furthermore assume that all the singularities of F are located at $\tilde{x} \in]-\pi,\pi]$, i.e. imagine either $\tilde{x} = 0$ or $\tilde{x} = \pi$. Setting $B(x) = b(x) * I$ with $b(x) = 1 + \cos(x)$ in the former and $b(x) = 1 - \cos(x)$ in the latter case we obtain

$$F_2(x) = |b(x/2)|^2 F(x/2) + |b(x/2+\pi)|^2 F(x/2+\pi) =$$

$$= |b(x/2)|^2 U(x/2)^H \Lambda(x/2) U(x/2) + |b(x/2+\pi)|^2 U(x/2+\pi)^H \Lambda(x/2+\pi) U(x/2+\pi) =$$

$$= U(x/2)^H * (|b(x/2)|^2 \Lambda(x/2) + |b(x/2+\pi)|^2 \Lambda(x/2+\pi)) * U(x/2) .$$

Obviously, $F_2(x)$ is positive semidefnite. The only possible singularities are all located at $2\tilde{x}$, because $2\tilde{x}$ is then a zero of both addends in $F_2(x)$. (More concretely, the singularities at π would move to the origin for the Galerkin coarse grid matrix – and this observation, by the way, once again underlines why a natural coarse grid operator could not be used in this case.) Let $u(\tilde{x})$ be the eigenvector of $F(x)$ to λ_j at $\tilde{x}$ with $\lambda(\tilde{x}) = 0$; then

$$F_2(2\tilde{x}) * u(\tilde{x}) = |b(\tilde{x})|^2 F(\tilde{x}) u(\tilde{x}) + |b(\tilde{x}+\pi)|^2 F(\tilde{x}+\pi) u(\tilde{x}) = 0.$$

The first addend becomes zero because $\lambda(\tilde{x}) = 0$ and the second term because of the definition of $b(x)$. Hence, $F_2(x)$ has an eigenvalue at $2\tilde{x}$ with a zero of the same order as $\lambda_j(x)$ at $\tilde{x}$. (And as we assumed all singularities of F to be in the same place $\tilde{x} \in]-\pi,\pi]$, it is obvious that $\tilde{x}+\pi$ can not also be a zero of F.)
Note also that if all the singularities were at π one would certainly prefer to shift them all to the origin as explained in Lemma 1.

Still, we are looking at the Galerkin operator only for theoretical purposes. For practical algorithms we prefer to employ natural coarse grid operators: As we have already mentioned in the first sentence of this subsection the spectral behaviour of block Toeplitz matrices is determined by their eigenvalue functions. Thus we suggest to use Algorithm 1 without any changes for block Toeplitz matrices with small size blocks generated by simple matrix functions (which need not necessarily be diagonal).

3.2 Numerical experiments

At this stage we think it is time to convince the reader that we are on the right algorithmic track and present a few numerical experiments. In the following we are mainly looking at 2-by-2 and 3-by-3 diagonal generating functions (as eigenvalues) which we rotate in the plain or in three-dimensional space.
In the 2-by-2 case we can twist a given diagonal generating functions by an angle α via Givens rotations

$$Rot(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

i.e. given the diagonal matrix function F_{diag} the twisted generating function F then becomes $F = Rot(\alpha)^T * F_{diag} * Rot(\alpha)$.

Similarly, we would like to remind the reader that rotations in three-dimensional space are given as the product of three Givens rotations about the x-, y- and z-axes, i.e.

$$Rot3d(\alpha,\beta,\gamma) = Rot_z(\gamma) * Rot_y(\beta) * Rot_x(\alpha)$$

with

$$Rot_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$
$$Rot_y(\beta) = \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix},$$
$$Rot_z(\gamma) = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us first take a look at two examples with 2-by-2 generating matrix functions, namely

$$F_1 = \begin{pmatrix} x^2 & 0 \\ 0 & 1-\cos(x) \end{pmatrix} \tag{3.1}$$

and

$$F_2 = \begin{pmatrix} x^2 & 0 \\ 0 & |x| \end{pmatrix}. \tag{3.2}$$

Observe that F_2 exhibits singularities of orders 1 and 2 whereas F_1 has only got singularities of order 2. Thus F_1 will lead to more ill-conditioned block Toeplitz systems.

number of unknowns	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}
F_1 with $\alpha = 0$	13	14	13	14	13	14	14	14
F_1 with $\alpha = \pi/3$	13	14	13	14	14	14	14	14
F_2 with $\alpha = 0.5$	12	13	13	13	13	13	12	13
F_2 with $\alpha = 0.8$	12	12	12	12	12	12	12	12

Table 1. Iteration counts for F_1 and F_2 twisted by various rotations.

The iteration counts are in Table 1 are independent of the matrix size and we observe optimal $O(n \log n)$ convergence for all our numerical examples. Not surprisingly, the angle α does not have a major influence. Note also that our W-cycle solvers have no problem at all dealing with cases where not all zeroes are of even order.

Still, we would also briefly like to point out that our algorithms work equally well in case the generating matrix functions are not diagonalizable with a constant orthogonal matrix. Therefore we study the following two examples:

$$F_3 = \begin{pmatrix} 8*x^2 & (\sin(x))^4 \\ (\sin(x))^4 & 8*x^2 \end{pmatrix} \tag{3.3}$$

has first been investigated by Miranda and Tilli in [11], sec. 5, where it is also explained in detail that F_3 has singularities of order 2 at the origin.
On the other hand,

$$F_4 = \begin{pmatrix} \frac{4*|x|}{\pi} + \frac{x^2}{\pi^2} & \frac{2*|x|}{\pi} \\ \frac{2*|x|}{\pi} & 4 \end{pmatrix} \tag{3.4}$$

exhibits only a singularity of order 1 at the origin.
Just as before we observe asymptotically optimal performance of our W-cycle solvers when applied to the corresponding block Toeplitz systems:

number of unknowns	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}
F_3 with $\alpha = 0$	11	12	12	12	12	12	12	12	12
F_4 with $\alpha = 0$	9	10	10	10	10	10	10	10	10

Table 2. Iteration counts for F_3 and F_4, two generating matrix functions which are both not diagonalizable by constant orthogonal matrices.

Finally, we would also like to look at an example with a 3-by-3 generating function

$$F_5 = \begin{pmatrix} |x| & 0 & 0 \\ 0 & x^2 & 0 \\ 0 & 0 & 1-\cos(x) \end{pmatrix}. \tag{3.5}$$

Note that F_5 has singularities of orders both 1 and 2. Again, our W-cycle algorithms work extremely well:

number of unknowns	$3*2^6$	$3*2^7$	$3*2^8$	$3*2^9$	$3*2^{10}$	$3*2^{11}$
F_5 rotated	12	12	12	12	13	13

Table 3. Iteration counts for F_5 rotated with $\alpha = \pi/3$, $\beta = \pi/4$ and $\gamma = \pi/6$ in the x-, y- and z-directions.

We see that employing natural coarse grid operators leads to optimal order convergence in all our examples. Furthermore, the occurence of singularities of odd order does not cause any problems. However, we can so far only deal with the case that all the singularities are located at the origin.

4 BLOCK TOEPLITZ MATRICES WITH MORE GENERAL UNDERLYING FUNCTIONS AND SINGULARITIES OF ORDER AT MOST TWO

4.1 Motivation: The scalar Toeplitz case

We would now like to overcome the algorithmic restriction that all the singularities of our generating function are at the origin. Let us therefore take a look at two examples for which we can not yet apply natural coarse grid operators:
The problem

$$F_6 = G_1 = \begin{pmatrix} 1+\cos(x) & 0 \\ 0 & 1-\cos(x) \end{pmatrix} \tag{4.1}$$

from Example 3 is simple in the sense of Definition 1, but has singularities at π.
On the other hand

$$F_7 = \begin{pmatrix} 1-\cos(2x) & 0 \\ 0 & |x| \end{pmatrix} \tag{4.2}$$

does not satisfy Definition 1.
However, looking at those two matrix functions more closely, we will see that they

actually have a lot in common.
Let us therefore go back to the Toeplitz systems generated by $f(x) = 1 - \cos(l * x)$. As we already mentioned at the end of subsection 2.1, it was R. Chan and collaborators who first pointed out in [3] that these Toeplitz matrices should be treated by prolongation and restriction operators related to $b(x) = 1 + \cos(l * x)$.
Our idea is to interpret $f(x) = 1 - \cos(l * x)$ as a matrix function

$$F = (1 - \cos(x)) * I_l. \tag{4.3}$$

This enables us to understand why a block approach is appropriate for these kinds of linear systems.

Example 4: More specifically, let us take a look at the Toeplitz matrices generated by $f(x) = 1 - \cos(2x)$. Viewing these matrices as 1-by-1 block Toeplitz problems we certainly could not apply a multigrid algorithm with a natural coarse grid matrix. However, what we need to do is to look at these problems from the block Toeplitz point of view (4.3) and understand $f(x) = 1 - \cos(2x)$ as

$$F_{block} = \begin{pmatrix} 1 - \cos(x) & 0 \\ 0 & 1 - \cos(x) \end{pmatrix},$$

which is certainly a simple generating matrix function.

With the same trick of redoubling the generating matrix function we are in a position to apply Algorithm 1 to both F_6 and F_7.

4.2 Proper expansion of generating matrix functions

We would now like to make our multigrid algorithms – including the natural coarse grid operator – available for more general block Toeplitz systems with small blocks of fixed size.
Let us therefore look at the case of redoubling a matrix function in a little more detail and prove the following theorem.

Theorem 1: Let $A(x)$ be a $k \times k$ matrix function relative to the symmetric positive definite block Toeplitz matrix $B = BT(A(x))$ of size $kn \times kn$ with even n. Then we can consider B also as a block Toeplitz matrix with block size $2k$ relative to a generating matrix function $\tilde{A}(x)$. Furthermore, for the singularities of $A(x)$ and $\tilde{A}(x)$ it holds:

$$\tilde{A}(2x) = 0 \quad \text{if and only if} \quad A(x) = 0.$$

Proof: With the generating function

$$A(x) = \begin{pmatrix} a_{11}(x) & \cdots & a_{1k}(x) \\ \vdots & & \vdots \\ a_{k1}(x) & \cdots & a_{kk}(x) \end{pmatrix},$$

the given block Toeplitz matrix

$$B = \left(\begin{array}{ccc|ccc|cccc} a_{11,0} & \cdots & a_{1k,0} & a_{11,-1} & \cdots & a_{1k,-1} & \cdots & \cdots & \cdots & \cdots \\ a_{k1,0} & \cdots & a_{kk,0} & a_{k1,-1} & \cdots & a_{kk,-1} & \cdots & \cdots & \cdots & \cdots \\ -- & -- & -- & -- & -- & -- & -- & -- & -- & \\ a_{11,1} & \cdots & a_{1k,1} & a_{11,0} & \cdots & a_{1k,0} & a_{11,-1} & \cdots & a_{1k,-1} & \cdots \\ a_{k1,1} & \cdots & a_{kk,1} & a_{k1,0} & \cdots & a_{kk,0} & a_{k1,-1} & \cdots & a_{kk,-1} & \cdots \\ -- & -- & -- & -- & -- & -- & -- & -- & -- & \\ . & . & . & . & . & . & . & . & . & . \end{array}\right)$$

can be related to the generating matrix function

$$\tilde{A}(x) = (1/2)\begin{pmatrix} A(x/2) + A(x/2+\pi) & e^{ix/2}(A(x/2) - A(x/2+\pi)) \\ e^{-ix/2}(A(x/2) - A(x/2+\pi)) & A(x/2) + A(x/2+\pi) \end{pmatrix} .$$

This matrix is up to a diagonal transformation equivalent to

$$F = \begin{pmatrix} C+D & C-D \\ C-D & C+D \end{pmatrix}$$

with $C = A(x/2)$ and $D = A(x/2+\pi)$. The matrix F is singular exactly iff

$$0 = \begin{pmatrix} C+D & C-D \\ C-D & C+D \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} C(u+v) + D(u-v) \\ C(u+v) - D(u+v) \end{pmatrix} \quad , \quad \|u\|_2^2 + \|v\|_2^2 = 1 .$$

This can be true only for $C = A(x/2)$ singular or $D = A(x/2+\pi)$ singular. Hence, the zeroes of $\tilde{A}(x)$ are exactly twice the zeroes of $A(x)$.

Theorem 1 implies that we can transform generating matrix functions with zeroes at both 0 and π, like e.g. F_6 or F_7, into simple matrix functions of double block size with zeroes only at the origin. In combination with Lemma 1, the above theorem also allows us to apply the natural coarse grid approach to all generating functions with any single zero or with two zeroes of distance π.

Obviously, we could prove completely analogous theorems for tripling, quadrupling, and so on, of matrix functions. That way we are able to turn generating matrix functions with singularities at $\frac{l*\pi}{k}$ (with k, l integer and $k \leq l$) into simple matrix functions. This enables us to apply our multigrid algorithms with natural coarse grid operators to the corresponding block Toeplitz matrices. We expect our algorithms to work well as long as k is relatively small.

4.3 Numerical tests

In this subsection of numerical experiments we take a look at rotated versions of F_6 and F_7 from (4.1) and (4.2), respectively. Furthermore, we also study

$$F_8 = \begin{pmatrix} x^2 & 0 \\ 0 & 1+\cos(x) \end{pmatrix}$$

and

$$F_9 = \begin{pmatrix} x^2 * (x-\pi)^2 & 0 \\ 0 & |x| \end{pmatrix} .$$

Note that F_6 and F_8 have only got singlarities of order 2, whereas F_7 is similar to F_9 in the sense that both matrix functions are not simple and have singularities of orders 1 and 2.

number of unknowns	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}
F_6 with $\alpha = 0$	11	11	12	12	12	12	12	12	12
F_7 with $\alpha = \pi/2$	11	11	12	11	12	11	12	12	12
F_8 with $\alpha = \pi/6$	20	21	21	20	20	20	20	20	20
F_9 with $\alpha = \pi/4$	12	12	12	12	12	12	12	12	12

Table 4. Iteration counts for F_6 to F_9 twisted by various rotations α .

We observe that our concept of redoubling the generating matrix functions works extremely well: The iterations counts are independent of the matrix size.

Finally, we would also like to look at an example where one needs to triple the generating matrix function and test a wisted version of

$$F_{10} = \begin{pmatrix} 1-\cos(x) & 0 \\ 0 & 1-\cos(3x) \end{pmatrix}.$$

number of unknowns	$6*2^5$	$6*2^6$	$6*2^7$	$6*2^8$	$6*2^9$	$6*2^{10}$
F_{10} with $\alpha = 0.7$	17	17	17	17	17	17

Table 5. Iteration counts for a rotated variant of F_{10}.

As before, Table 5 confirms optimal order multigrid convergence for the trebled generating matrix function.

5 MORE GENERAL BLOCK TOEPLITZ PROBLEMS

So far, the attention of this paper has been confined to ill-conditioned block Toeplitz matrices with singularities of order at most 2. In this final section we would briefly like to show that our new algorithms may also be very useful in more general situations. In particular, we are looking at well-conditioned matrices and problems of vector Laplace type.

5.1 Well-conditioned block Toeplitz matrices

Previous papers on block Toeplitz matrices without Toeplitz blocks have also discussed well-conditioned symmetric matrices – see, in particular, [11] and [12].
Let us start looking at an example with a 3-by-3 generating function

$$F_{11} = \begin{pmatrix} x^2+1 & |x| & x^2 \\ |x| & x^2+1 & |x| \\ x^2 & |x| & 2x^2+1 \end{pmatrix}.$$

The generating function F_{11} is taken from S. Serra Capizzano's paper [12], sec. 4, where the author points out in detail why the corresponding block Toeplitz systems are well-conditioned.
In the following we compare two ways of applying multigrid to these matrices. We can either ignore the block structure completely and use standard transfer operators corresponding to $b(x) = 1+\cos(x)$ or we can use block prolongations and restrictions according to $b(x) = 1+\cos(3x)$. Of course, we still employ a natural coarse grid representation.

number of unknowns	$3*2^6$	$3*2^7$	$3*2^8$	$3*2^9$	$3*2^{10}$	$3*2^{11}$
Block approach	29	29	29	29	29	29
Standard approach	47	47	47	47	47	47

Table 6. Iteration counts for W-cycle solvers for F_{11}.

Table 6 shows that both approaches for prolongation and restriction lead to asymptotically optimal convergence for this well-conditioned problem. Still, it pays to take into account the block structure of the problem. Employing block versions of the transfer operators nearly cuts the iteration counts in half. (Note that applying block prolongations and restrictions does not necessitate more floating point operations than the standard approach.)

Now let us study the influence of the "amount of well-conditioning" (– or ill-conditioning, respectively –) in more detail. The following example is derived from matrix function F_1 from (3.1) and given as

$$F_{12}(c) = \left(Rot(0.75)^T * \begin{pmatrix} x^2 & 0 \\ 0 & 1-\cos(x) \end{pmatrix} * Rot(0.75) \right) + c * I. \tag{5.1}$$

A positive shift c guarantees that the associated block Toeplitz matrices are well-conditioned.

number of unknowns	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}
Block approach, $c = 0.25$	17	17	17	17	17	17	17	17	17
Standard approach, $c = 0.25$	71	77	78	78	78	78	78	78	78
Block approach, $c = 0.5$	13	12	12	12	12	12	12	12	12
Standard approach, $c = 0.5$	40	41	42	42	42	42	42	42	42
Block approach, $c = 1$	9	9	9	9	9	9	9	9	9
Standard approach, $c = 1$	22	22	22	22	22	22	22	22	22

Table 7. Iteration counts for F_{12} with different shifts c.

Table 7 points out that it always pays off to take into account the 2-by-2 block structure of the problems from (5.1) and to employ block prolongations and restrictions according to $b(x) = 1 + \cos(2x)$. However, the standard approach also leads to optimal order convergence, although a significantly larger number of multigrid iterations may be required. Comparing different shifts c we conclude that the more well-conditioned our matrix (i.e. the further we push the singularities away from the origin), the less we benefit from exploiting block structure.

5.2 Problems of vector Laplace type

We would finally like to point out that the concept of block prolongations and restrictions is also useful in a broader context than block Toeplitz systems with blocks of small size. We shall see that we can understand the small size blocks as degrees of freedom associated with a node.

Let us start by discussing the simplest possible introductory example: We are looking at uniform discretizations of the two-dimensional vector Poisson problem

$$\Delta \underline{u} = \underline{f}$$

on a square. If we keep x- and y-components belonging to a node next to each other during the discretization, then we end up with a bivariate block Toeplitz matrix with small 2-by-2 blocks. The resulting linear system looks like the scalar two-dimensional

Laplacian with the scalars replaced by those little blocks: Given the meshsize $h = \frac{1}{N}$, our system matrix

$$A_h = \begin{pmatrix} D & -I & \cdots & \cdots & 0 \\ -I & D & -I & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & -I & D & -I \\ 0 & \cdots & \cdots & -I & D \end{pmatrix} \tag{5.2}$$

is a bivariate block Toeplitz matrix of size $2N^2$-by-$2N^2$ (consisting of N blocks each of which is of size $2N$-by-$2N$). In (5.2) the matrix D is given by

$$D = \begin{pmatrix} 4 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 4 & 0 & -1 & \cdots & \vdots \\ \vdots & & \ddots & & \ddots & \vdots \\ \vdots & \cdots & -1 & 0 & 4 & 0 \\ 0 & \cdots & 0 & -1 & 0 & 4 \end{pmatrix}$$

whereas I stands for the identity.

Obviously, our discretization of the vector Poisson equation has lead to a very special situation, because our 2-by-2 blocks are themselves (diagonal) Toeplitz matrices. (This is physically clear, because in the vector Laplacian the x- and y-components associated with a node are not coupled.) In this particular case the resulting linear system is a BTTB system, i.e. the coefficient matrix is a block Toeplitz matrix with Toeplitz blocks. BTTB matrices are the multivariate counterparts of scalar Toeplitz matrices. For more information on multigrid methods for BTTB problem see e.g. [17] or [10], sec. 5.

We would like to point out that the BTTB matrices we obtain from our uniform discretization of the vector Poisson equation on the square are associated with the generating function

$$f(x, y) = 2 - \cos(2x) - \cos(y). \tag{5.3}$$

Note also that matrices generated by (5.3) can be written as Kronecker sums

$$Mat = I \otimes A + B \otimes I \tag{5.4}$$

with A and B standing for scalar Toeplitz matrices generated by $g(x) = 1 - \cos(2x)$ and $h(x) = 1 - \cos(x)$, respectively.

If we want to treat such matrices by multigrid, we are again left with the two variants we discussed at the beginning of subsection 2.1. The first possibility is to use the fact that the vector Laplacian is a separable problem and to uncouple the x- and y-components by a straightforward permutation. That way we only have to deal with standard two-dimensional Laplace matrices (which, by the way, belong to the function $f(x, y) = 2 - \cos(x) - \cos(y)$). We can simply apply standard multigrid and hence we use

$$P = P_1 \otimes P_1$$

for prolongation with P_1 as defined in (2.3).

The other option is not to apply the permutation. It is plain that then the proper prolongation operator becomes

$$P = P_1 \otimes P_2 \tag{5.5}$$

with P_1 and P_2 as defined in (2.3).
Note again that P_1 corresponds to $1 + \cos(x)$, whereas the block prolongation P_2 is related to $1 + \cos(2x)$. We could say that within (5.5) it is the role of the block prolongation P_2 to keep together the two degrees of freedom corresponding to the x- and y-components of a node in the discretization.
We would also like to mention that prolongation and restriction operators based on blocks have been considered before, in particular in the context of algebraic multigrid for problems in elasticity. (See e.g. [19] or [20]).

However, the focus of our work does not lie upon sparse linear systems arising from the discretization of partial differential equations, but rather on more densely populated multivariate block Toeplitz problems with closely related properties. Hence we would briefly like to show that block prolongations as given by (5.5) are also the proper choice in case our multivariate block Toeplitz systems have small non-Toeplitz blocks. (In other words: We are particularly interested in situations where the individual degrees of freedom associated with a node are actually coupled via the small size blocks.) As usual, we restrict our attention to singularities of order at most 2 and employ natural coarse grid matrices.

In addition to the vector Laplacian generated by (5.3) we also test the following three examples defined as Kronecker sums in terms of (5.4). Let A and B in (5.4) be defined as follows:

Matrix B1: $A = Rot(0.8)^T * F_1 * Rot(0.8)$ and $B = y^2$ with F_1 from (3.1).
Matrix B2: $A = Rot(\pi/2)^T * F_2 * Rot(\pi/2)$ and $B = |y|$ with F_2 from (3.2).
Matrix B3: $A = F_4$ and $B = y^2$ with F_4 from (3.4).

number of unknowns	$2^4 * 2^4$	$2^5 * 2^5$	$2^6 * 2^6$	$2^7 * 2^7$	$2^8 * 2^8$
Vector Laplacian (5.3)	10	11	12	12	12
Matrix B1	13	13	13	13	13
Matrix B2	13	14	14	14	14
Matrix B3	14	14	14	14	14

Table 8. Iteration counts for W-cycle solvers for matrices of vector Laplace type.

Our W-cycle solvers work equally well for both the vector Laplace problem and the three other examples. Note that the bivariate block Toeplitz matrices B1, B2 and B3 do not have BTTB structure and that for these three examples the x- and y-components belonging to a node can not simply be uncoupled by permutation. This outlines that the block transfer operators we suggested are also viable in a wider context.

5.3 Summary, conclusions and outlook

In this paper we proposed new multigrid algorithms for block Toeplitz matrices. There are two main ingredients of our algorithms: One is the use of a natural coarse grid representation, the other is exploiting the block information for the construction of prolongation and restriction operators. We pointed out that our approach is extremely well suited for symmetric positive block Toeplitz problems with small size blocks exhibiting singularities of order at most 2. The only additional information we rely on are the locations and orders of these zeroes. Thereby we also showed how the restriction of the singularities to be at the origin can be overcome by a proper expansion of the generating matrix function. These considerations are also interesting in the light of scalar Toeplitz problems. Furthermore, we demonstrated the feasibility of our multigrid solvers through a very large number of numerical experiments. We would like to stress that all the tables presented in this paper employ a natural coarse grid matrix and that the iteration counts always indicate optimal order $O(n \log n)$ convergence.

The final section points out that our block considerations are applicable in a broader context and we have merely been able to touch on the subjects of well-conditioned structured linear systems and matrices of vector Laplace type.

Furthermore, we have not been considering any singularities of order higher than 2 in this article. For such problems it might be particularly interesting to study the suitability of block prolongations and restrictions of higher order.

Finally, another aspect we have deliberately excluded from this study are different smoothers. As we stated in subsection 1.5, we use scalar Richardson smoothing in all our numerical tests. In our view, this underlines the fact that the optimality we observe in our experiments is due to the block transfer operators we propose. Still, we are aware that our solvers might be further improved by employing block smoothers. For the aforementioned problem of singularities of higher order block smoothing might even be needed in order to cope with the larger degree of ill-conditioning.

The above aspects could be subject to future investigations.

REFERENCES

1. Briggs W. *A Multigrid Tutorial*, SIAM, 1987.
2. Chan R, Ng M. Conjugate gradient methods for Toeplitz systems. *SIAM Review* 1996; **38**(3): 427–482.
3. Chan R, Chang C, Sun H. Multigrid method for ill-conditioned symmetric Toeplitz systems, *SIAM J. Sci. Comp.* 1998; **19**(2): 516–529.
4. Fiorentino G, Serra S. Multigrid methods for Toeplitz matrices. *Calcolo* 1992; **28**: 283–305.
5. Fiorentino G, Serra S. Multigrid methods for symmetric positive definite Block Toeplitz matrices with nonnegative generating functions. *SIAM J. Sci. Comp.* 1996; **17**(5): 1068–1081.
6. Greenbaum A. *Iterative Methods for Solving Linear Systems*, SIAM, 1997.
7. Grenander U, Szegö G. *Toeplitz Forms and Their Applications.* Second Edition, Chelsea, 1984.
8. Hackbusch W. *Multigrid methods and applications.* Springer, 1985.
9. Huckle T, Staudacher J. Matrix multilevel methods and preconditioning. *BIT* 2002; **42**(3): 541–560.
10. Huckle T, Staudacher J. Multigrid preconditioning and Toeplitz matrices. *ETNA* 2002; **13**: 81–105.
11. Miranda M, Tilli P. Asymptotic Spectra of Hermitian Block Toeplitz matrices and preconditioning results, *SIAM J. Matrix Anal. Appl.* 2000; **21**(3): 867–881.
12. Serra S. Asymptotic results on the spectra of Block Toeplitz preconditioned matrices, *SIAM J. Matrix Anal. Appl.* 1998; **20**(1): 31–44.
13. Serra Capizzano S. Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs Matrix-sequences, *Numer. Math.* 2002; **92**(3): 433–465.

14. Serra Capizzano S, Tablino Possio C. Multigrid methods for multilevel circulant matrices. Preprint, 2001, submitted to *SIAM J. Sci. Comp.*, 16 pages.
15. Staudacher J. *Multigrid methods for matrices with structure and applications in image processing.* Ph.D. thesis, Department of Computer Science, Technical University Munich, 2002. Available online via *http://tumb1.biblio.tu-muenchen.de/publ/diss/in/2002/staudacher.pdf*
16. Sun H, Chan R, Chang Q. A note on the convergence of the two-grid method for Toeplitz systems. *Comp. Math. Appl.* 1997: **34**(1): 11–18.
17. Sun H, Jin X, Chang Q. Convergence of the multigrid method for ill-conditioned Block Toeplitz systems, *BIT* 2001; **41**(1): 179–190.
18. Trottenberg U, Oosterlee C, Schüller K. *Multigrid.* Academic Press, 2001.
19. Vanek P, Brezina M, Mandel J. Convergence of Algebraic Multigrid Based on Smoothed Aggregation, *Numerische Mathematik* 2001; **88**: 559–579.
20. Vanek P, Mandel J, Brezina M. Algebraic Multigrid by Smoothed Aggregation for Second and Fourth Order Elliptic Problems, *Computing* 1996; **56**: 179–196.