A Note on the Convergence of the Two-Grid Method for Toeplitz Systems[*]

Hai-Wei Sun[†]   Raymond H. Chan[‡]   Qian-Shun Chang[§]

Abstract

In this paper, we consider solutions of Toeplitz systems $Au = b$ where the Toeplitz matrices A are generated by nonnegative functions with zeros. Since the matrices A are ill-conditioned, the convergence factor of classical iterative methods, such as the Richardson method, will approach 1 as the size n of the matrices becomes large. In [1, 2], convergence of the two-grid method with Richardson method as smoother was proved for band τ matrices and it was conjectured that this convergence result can be carried to Toeplitz systems. In this paper, we show that the two-grid method with Richardson smoother indeed converges for Toeplitz systems that are generated by functions with zeros, provided that the order of the zeros are less than or equal to 2. However, we illustrate by examples that the convergence results of the two-grid method cannot be readily extended to multi-grid method for n that are not of the form $2^\ell - 1$.

AMS(MOS) subject classifications. 65F10.

Key Words: Multigrid Method, Toeplitz Matrices, Richardson Method, Damped-Jacobi Method.

1   Introduction

An n-by-n matrix A_n is said to be a Toeplitz matrix if it has the form

$$
A_n = \begin{bmatrix}
a_0 & a_{-1} & \cdots & a_{2-n} & a_{1-n} \\
a_1 & a_0 & a_{-1} & \cdots & a_{2-n} \\
\vdots & a_1 & a_0 & \ddots & \vdots \\
& & \ddots & \ddots & \\
a_{n-2} & \cdots & \ddots & \ddots & a_{-1} \\
a_{n-1} & a_{n-2} & \cdots & a_1 & a_0
\end{bmatrix},
$$

i.e. A_n is constant along its diagonals. Given a function $f(\theta)$, let

$$
a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} d\theta, \qquad k = 0, \pm 1, \pm 2, \cdots,
$$

[†]Department of Mathematics & Physics, Guangdong University of Technology, Guangzhou, People's Republic of China.

[‡]Department of Mathematics, Chinese University of Hong Kong, Shatin, Hong Kong.

[§]Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, People's Republic of China.

be the Fourier coefficients of f. For all $n \geq 1$, let A_n be the n-by-n Toeplitz matrix with entries $a_{j,k} = a_{j-k}$, $0 \leq j,k < n$. The function $f(\theta)$ is called the generating function of the sequence of Toeplitz matrices A_n. We will denote $A_n = \mathcal{T}_n[f(\theta)]$. When f is an even function, the matrices A_n are real symmetric. In this paper we study the solutions of Toeplitz systems $A_n u = b$ by the multigrid method.

In Grenander and Szegö [3, pp.64-65], it is shown that the eigenvalues $\lambda_j(A_n)$ of A_n lie in the range of $f(\theta)$, i.e.

$$\min_{\theta \in [-\pi,\pi]} f(\theta) \leq \lambda_j(A_n) \leq \max_{\theta \in [-\pi,\pi]} f(\theta), \qquad 1 \leq j \leq n. \tag{1}$$

Moreover, the maximum and minimun eigenvalues of A_n satisfy

$$\lim_{n \to \infty} \lambda_{\max}(A_n) = \max_{\theta \in [-\pi,\pi]} f(\theta) \qquad \text{and} \qquad \lim_{n \to \infty} \lambda_{\min}(A_n) = \min_{\theta \in [-\pi,\pi]} f(\theta).$$

Consequently, when $f(\theta)$ is nonnegative and vanishes at some points $\theta_0 \in [-\pi, \pi]$, then the condition number $\kappa(A_n)$ of A_n is unbounded as n tends to infinity, i.e. A_n is ill-conditioned.

Classical iterative methods such as the Jacobi, Gauss-Seidel or SOR methods are not applicable when the generating function has zeros. Since $\lim_{n \to \infty} \kappa(A_n) = \infty$, the convergence factor is expected to approach 1 for large n. In [1, 2], Fiorentino and Serra proposed to use multigrid method (MGM) with Richardson method as smoother for solving Toeplitz systems. Their numerical results show that the multigrid method gives very good convergence rate for Toeplitz systems generated by nonnegative functions. The cost per iteration of MGM is of $O(n \log n)$ operations.

However, in [1, 2], the convergence of two-grid method (TGM) is only proved for the so-call band τ matrices. These are matrices that can be diagonalized by sine transform matrices. In general, τ matrices are not Toeplitz matrices and vice versa. Thus the author in [2] has posed the question of whether the convergence of TGM for τ matrices can be carried over to Toeplitz matrices. In this paper, we give partial answer to this question.

We first prove that with the smoother and projection operators proposed by Fiorentino and Serra in [1, 2], TGM indeed converges for Toeplitz systems generated by functions with zeros that are of order 2 or less. More precisely, using the theory of algebraic multigrid method, we are able to prove that the convergence factor of TGM for this class of Toeplitz systems is uniformly bounded below 1 independent of n. However, examples will be given to illustrate that the convergence results for TGM cannot be readily extended to MGM unless n is of the form $2^\ell - 1$. This is because the coarse grid operator will not be Toeplitz in this case and it is difficult to estimate the spectral radius of the coarse grid operator to be used in the Richardson smoother. As a remedy, one can use the damped-Jacobi method as smoother instead of the Richardson method. The convergence of MGM method with damped-Jacobi method as smoother can be found in Chan, Chang and Sun [4].

The paper is organized as follows. In §2, we recall the method proposed in [1, 2]. In §3, we analyze the convergence of TGM for Toeplitz systems where the generating functions have a zero of order 2 or less. In §4, numerical examples are given to illustrate the effectiveness of MGM method with Richardson smoother when n is of the form $2^\ell - 1$. In §5, we give some remarks on the choice of the smoothers, when n is not of the form $2^\ell - 1$. In particular, we give examples to show that the Richardson smoother does not work when n is not of this form. Comparison with damped-Jacobi smoother will be given too. Concluding remarks are given in §6.

2 The Two-Grid Method for Toeplitz Matrices

In this section, we recall the two-grid method proposed by Fiorentino and Serra in [1, 2] for Toeplitz matrices. Given $Au = b$, with $u, b \in \mathbb{R}^n$, the smoother is defined as

$$u^{(j+1)} = Su^{(j)} + b_1 = S(u^{(j)}, b_1),$$

where $S = I - M^{-1}A$ is the iteration matrix, M depends on the iterative scheme and $b_1 = M^{-1}b \in \mathbb{R}^n$. Let P be the projection operator. The TGM algorithm is given by:

> $\mathbf{TGM}(S, P)$
> $u^{(j,\nu)} = S^{\nu}(u^{(j)}, b_1)$;
> $d_n = Au^{(j,\nu)} - b$;
> $d_k = Pd_k$;
> $A_k = PAP^T$;
> $\texttt{Solve}\ A_k y = d_k$;
> $u^{(j+1,\nu)} = u^{(j,\nu)} - P^T y$.

We note that the global iteration matrix of TGM is

$$G = [I - P^T (PAP^T)^{-1} PA] S^{\nu}. \tag{2}$$

In [1, 2], S is chosen to be the Richardson iteration, i.e.

$$S = I - \frac{1}{\rho}A, \tag{3}$$

where $2\rho > \rho(A)$, the spectral radius of A. For a Toeplitz matrix A generated by an even function f, we see from (1) that $\rho(A) \leq \max_{\theta \in [-\pi,\pi]} f(\theta)$. In applications where f is not known a priori, we can estimate $\rho(A)$ by the matrix ∞-norm of A. The estimate can be computed in $O(n)$ operations.

In §3, we will give the convergence proof of TGM for Toeplitz systems where their generating functions $f(\theta)$ are even and satisfy

$$\min_{\theta \in [-\pi,\pi]} \frac{f(\theta)}{1 - \cos\theta} > 0 \tag{4}$$

or

$$\min_{\theta \in [-\pi,\pi]} \frac{f(\theta)}{1 + \cos\theta} > 0. \tag{5}$$

Since $1 \pm \cos\theta$ are functions with zeros that are of order 2, any generating function $f(\theta)$ that satisfies (4) or (5) will have zeros of order less than or equal to 2. For if otherwise, then it can easily be shown that the minimum is zero at the zeros of $f(\theta)$.

For these Toeplitz systems, the projection P defined in Fiorentino and Serra [1] reduced to the classical projectors, namely

$$P = \begin{bmatrix} 1 & 2 & 1 & & & \\ & & 1 & 2 & 1 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} \tag{6}$$

3

or

$$P = \begin{bmatrix} -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{7}$$

depending on whether (4) or (5) holds.

3 Convergence Analysis

In this section, we discuss the convergence of TGM for Toeplitz matrices whose generating functions satisfy (4) or (5). Let us begin by introducing the following notations. We say $A > B$ (respectively $A \geq B$) if $A - B$ is a positive (respectively semi-positive) definite matrix. In particular, $A > 0$ means that A is positive definite. For $A > 0$, we define the following inner products which are useful in the convergence analysis of multigrid methods, see Ruge and Stuben [5, p.77–78]:

$$\langle u, v \rangle_0 = a_0 \langle u, v \rangle, \quad \langle u, v \rangle_1 = \langle Au, v \rangle, \quad \langle u, v \rangle_2 = \frac{1}{a_0} \langle Au, Av \rangle. \tag{8}$$

Here $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and a_0 is the main diagonal entry of A. The respective norms of the inner products defined in (8) are denoted by $\| \cdot \|_i$, $i = 0, 1, 2$. We first note that for the Richardson smoother defined in (3), $\|S\|_1 \leq 1$ if ρ is properly chosen.

Theorem 1 ([5, p.84]) *Suppose $A > 0$ and $\rho > \rho(A)/2$. Then*

$$\|Se\|_1^2 \leq \|e\|_1^2 - \alpha \|e\|_2^2, \qquad \forall e \in \mathbb{R}^n, \tag{9}$$

where

$$\alpha = (2 - \frac{\rho(A)}{\rho}) \frac{a_0}{\rho} > 0. \tag{10}$$

Inequality (9) is called the *smoothing condition*. We see from the theorem that $\|S\|_1 \leq 1$. We see also from (10) that $\rho = \rho(A)$ is the best choice since α will be the largest.

   Let $A_n = A^h$ and $PA^h P^T = A^H$, the coarse grid operator. For TGM, the correction operator is given by

$$T = I - P^T (A^H)^{-1} P A^h$$

and hence by (2), the global iteration matrix of TGM is

$$G = TS^\nu.$$

In the following, we let $\nu = 1$.

Theorem 2 ([5, p.89]) *Let $A = A^h > 0$ and let ρ be chosen such that S satisfies the smoothing condition (9), i.e.*

$$\|Se^h\|_1^2 \leq \|e^h\|_1^2 - \alpha \|e^h\|_2^2, \qquad \forall e^h \in \mathbb{R}^n,$$

4

where α is given by (10). Suppose that the projection operator P has full rank and that there exists a scalar $\beta > 0$ such that

$$\min_{e^H \in \mathbb{R}^{\lfloor n/2 \rfloor}} \|e^h - P^T e^H\|_0^2 \leq \beta \|e^h\|_1^2, \qquad \forall e^h \in \mathbb{R}^n. \tag{11}$$

Then $\beta \geq \alpha$ and the convergence factor of the h-H two-level TGM satisfies

$$\|G\|_1 \leq \sqrt{1 - \frac{\alpha}{\beta}}.$$

Inequality (11) is called the *correcting condition*.

From Theorems 1 and 2, we see that if ρ is chosen as in Theorem 1, then we only have to establish (11) in order to get the convergence results.

Theorem 3 *Let the generating function $f(\theta)$ of A be even and satisfy (4) or (5) and let P be chosen as in (6) or (7) accordingly. Then there exists a scalar $\beta > 0$ independent of n such that (11) holds. In particular, the convergence factor of TGM is bounded uniformly below 1 independent of n.*

Proof: We will prove the theorem for the case of (4). The proof for the case of (5) is similar and is sketched at the end of this proof. We first assume that $n = 2k + 1$ for some k. For any $e^h = (e_1, e_2, \cdots, e_n)^t \in \mathbb{R}^n$, we define

$$e^H = (\tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_k)^t \in \mathbb{R}^k,$$

where

$$\tilde{e}_i = \frac{1}{2} e_{2i}, \qquad 1 \leq i \leq k.$$

For ease of indexing, we set $e_i = 0$ for $i \leq 0$ and $i > n$. We note that with P as defined in (6) and the norm $\|\cdot\|_0$ in (8), we have

$$\|e^h - P^T e^H\|_0^2 = a_0 \sum_{i=0}^{k} \{ e_{2i+1} - \frac{1}{2} e_{2i+2} - \frac{1}{2} e_{2i} \}^2. \tag{12}$$

Thus (11) is proved if we can bound the right hand side above by $\beta \|e^h\|_1$ for some β independent of e^h.

To do so, we observe that for the right hand side in (12), we have

$$a_0 \sum_{i=0}^{k} \{ e_{2i+1} - \frac{1}{2} e_{2i+2} - \frac{1}{2} e_{2i} \}^2$$

$$= a_0 \sum_{i=0}^{k} \{ e_{2i+1}^2 + \frac{1}{4} e_{2i+2}^2 + \frac{1}{4} e_{2i}^2 - e_{2i+2} e_{2i+1} - e_{2i} e_{2i+1} + \frac{1}{2} e_{2i+2} e_{2i} \}$$

$$\leq a_0 \sum_{i=0}^{k} \{ e_{2i+1}^2 + \frac{1}{4} e_{2i+2}^2 + \frac{1}{4} e_{2i}^2 - e_{2i+2} e_{2i+1} - e_{2i} e_{2i+1} + \frac{1}{4} e_{2i+2}^2 + \frac{1}{4} e_{2i}^2 \}$$

5

$$= a_0 \sum_{i=0}^{k} \{e_{2i+1}^2 + \frac{1}{2}e_{2i+2}^2 + \frac{1}{2}e_{2i}^2 - e_{2i+2}e_{2i+1} - e_{2i}e_{2i+1}\}$$

$$= a_0 \sum_{m=1}^{n} (e_m^2 - e_m e_{m+1}) = a_0 \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle$$

where $\mathcal{T}_n[1 - \cos\theta]$ is the n-by-n Toeplitz matrix generated by $1 - \cos\theta$. Thus

$$\min_{e^H \in \mathbb{R}^k} \|e^h - P^T e^H\|_0^2 \leq a_0 \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle, \qquad \forall e^h \in \mathbb{R}^n.$$

Hence to establish (11), we only have to prove that

$$a_0 \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle \leq \beta \|e^h\|_1^2, \qquad \forall e^h \in \mathbb{R}^n$$

for some β independent of e^h. By definition of $\|\cdot\|_1$, see (8), it is equivalent to proving

$$a_0 \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle \leq \beta \langle e^h, A^h e^h \rangle, \qquad \forall e^h \in \mathbb{R}^n$$

for some β independent of e^h.

By (4) and (1), it is obvious that

$$\gamma \mathcal{T}_n[1 - \cos\theta] \leq \mathcal{T}_n[f] = A^h,$$

where

$$\gamma = \min_{\theta \in [-\pi, \pi]} \frac{f(\theta)}{1 - \cos\theta} > 0.$$

Hence

$$a_0 \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle \leq \beta \langle e^h, A^h e^h \rangle, \qquad \forall e^h \in \mathbb{R}^n, \tag{13}$$

where

$$\beta = \frac{a_0}{\gamma}. \tag{14}$$

Thus (11) holds for the case of $n = 2k + 1$.

Next we consider the case where n is not of the form $2k + 1$. In that case, we let $k = n/2$, $\tilde{n} = 2k + 1 > n$. We then embed the vector e^h into longer vectors $e^{\tilde{h}}$ of size $\tilde{n}$ by zeros. Then since

$$\|e^h - P^T e^H\|_0^2 \leq \|e^{\tilde{h}} - \tilde{P}^T e^H\|_0^2$$

and

$$\langle e^{\tilde{h}}, \mathcal{T}_{\tilde{n}}[1 - \cos\theta]e^{\tilde{h}} \rangle = \langle e^h, \mathcal{T}_n[1 - \cos\theta]e^h \rangle$$

we see that the conclusion still holds.

We remark that the case of (5) can be proved similarly. We only have to replace the function $(1 - \cos\theta)$ above by $(1 + \cos\theta)$. Since in this case, $f(\theta) \geq \gamma(1 + \cos\theta)$, we then have

$$\mathcal{T}_n[1 + \cos\theta] \leq \frac{1}{\gamma}A^h.$$

From this, we get (13) and hence (11) with β defined as in (14). $\qquad \square$

6

4   Numerical Results

The analysis presented in §3 is for TGM where A^H, the coarse grid operator, is assumed to be inverted exactly. In the full multigrid method (MGM), A^H is not inverted exactly, but is approximated by using TGM recursively on coarser grids. In this section, we apply MGM for solving ill-conditioned real symmetric Toeplitz systems $A_n u = b$.
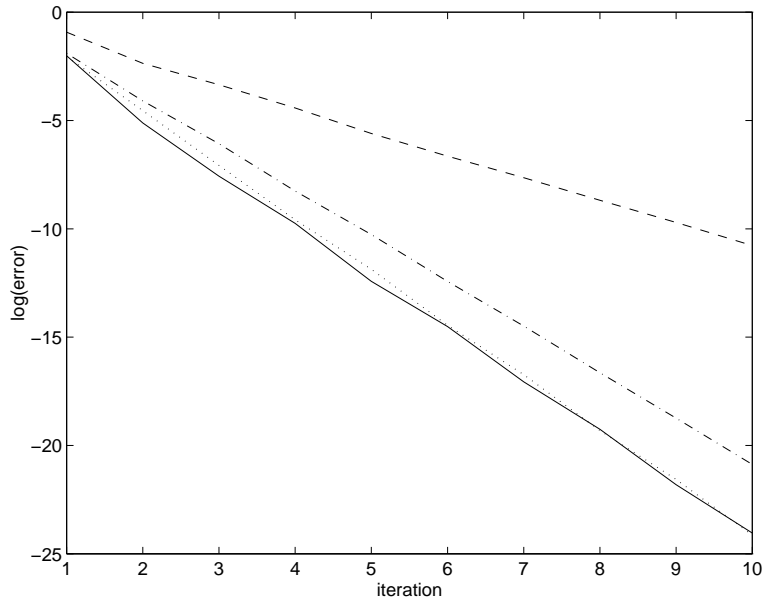
In Table 1, we give the convergence history of four nonnegative generating functions when $n = 2047 = 2^{11} - 1$. The first three are continuous functions and the fourth one, $J(\theta)$, is a function with jump:

$$J(\theta) = \begin{cases} \theta^2 & \text{if } |\theta| \leq \pi/2, \\ 1 & \text{if } |\theta| > \pi/2. \end{cases}$$

The Toeplitz matrices they formed, except for those generated by the first function $f(\theta) = 6 - 4\cos\theta - 2\cos 2\theta$, are dense. Since n is of the form $2^\ell - 1$, the coarse grid operators are also Toeplitz matrices, see for instance Chan, Chang and Sun [4]. We choose as solution a random vector u with $\|u\|_\infty \leq 1$. The right hand side vector b is obtained accordingly. The zero vector is used as the initial guess. When the size of the coarse grid operator is less than 8, we solve the equation exactly. In Table 1, we list the error $\|u - u_j\|_\infty$ where u_j is the approximated solution after the j-th V-cycle. As in [1, 2], we use the Richardson method with $\rho_{pre} = \max f(\theta)$ for pre-smoother and $\rho_{post} = \max f(\theta)/2$ for post-smoother. In the coarser grid level, we also use the Richardson method with $\rho_{pre}^H = a_0^H \rho_{pre}/a_0$ for pre-smoother and $\rho_{post}^H = a_0^H \rho_{post}/a_0$ for post-smoother. We use one pre-smoothing and one post-smoothing on each level. In Figure 1, we plot the log of the error against the number of iterations. We clearly see the linear convergence of the method. Also from the table and the figure, we see that MGM with Richardson smoother works very efficiently.

V-cycle	$6 - 4\cos\theta - 2\cos 2\theta$	$|\theta|$	θ^2	$J(\theta)$
1	1.333962e-01	1.294058e-01	3.972341e-01	1.535674e-01
2	6.023454e-03	1.062903e-02	9.420688e-02	1.663342e-02
3	5.143994e-04	8.542301e-04	3.525454e-02	2.284307e-03
4	5.842898e-05	6.930363e-05	1.193746e-02	2.652951e-04
5	4.015539e-06	7.010466e-06	3.785376e-03	3.537911e-05
6	4.992195e-07	5.342456e-07	1.301305e-03	3.999559e-06
7	3.888069e-08	5.415987e-08	4.876845e-04	5.136616e-07
8	4.339097e-09	4.115722e-09	1.745122e-04	5.910621e-08
9	3.382731e-10	4.293213e-10	6.132923e-05	7.366634e-09
10	3.627324e-11	3.270896e-11	2.144583e-05	8.642967e-10

Table 1: Convergence History for Full Multigrid Method.

solid line: $6 - 4\cos\theta$, dotted line: $|\theta|$, dashed line: θ^2, dotted-and-solid line: $J(\theta)$

Figure 1. The Log Error against the Iteration.

5 Remarks on the Choice of the Smoother

The smoother used in the above discussion is the Richardson method proposed by Fiorentino and Serra in [1, 2]. We note that for Toeplitz matrices, the Richardson method is the same as the damped-Jacobi method which is defined by

$$S_J = I - \omega \cdot (\text{diag}(A))^{-1} A$$

where $\omega^{-1} > \rho((\text{diag}(A))^{-1} A)/2$. In fact if A is Toeplitz, then $\text{diag}(A) = a_0 \cdot I$. Hence

$$S_J = I - \frac{\omega}{a_0} A. \tag{15}$$

Let $\rho = a_0/\omega$, then $S_J = S$ of (3).

If the size n of A is not the form $2^\ell - 1$, the coarse grid operator A^H will no longer be Toeplitz but is a sum of a Toeplitz matrix and a low rank matrix (with rank 2), see Chan, Chang and Sun [4]. In this case, the Richardson method is different from the damped-Jacobi method on the coarser level.

We note that it is cheaper to use the damped-Jacobi smoother than to use the Richardson smoother in general in the full multigrid method. For the damped-Jacobi iteration (15), we can use the same ω in (15) for all coarser level operators. However, the parameter ρ in the Richardson iteration in (3) cannot be computed easily on the coarser grid. And even if it is computed exactly, the performance may not be as good as the damped-Jacobi method with a constant ω.

To illustrate that, we consider Toeplitz matrices $\mathcal{T}_n[1 - \cos\theta]$ and $\mathcal{T}_n[\theta^2]$ for different values of n not of the form $2^\ell - 1$. We solve the systems by the full multigrid method with one pre-smoothing and one post-smoothing as in §4. For the Richardson smoother, the parameter ρ of the pre-smoother

on each coarser level is chosen to be the spectral radius of the coarse grid operator on that level and the parameter of post-smoother is chosen to be $\rho/2$. For the damped-Jacobi smoother, we just use $\omega_{pre} = a_0/\max f(\theta)$ in pre-smoothing and $\omega_{post} = 2a_0/\max f(\theta)$ for all levels. The average convergence rates of the first ten iterations for both smoothers are shown in Table 2. We see that the convergence rate of MGM with damped-Jacobi smoother is about constant independent of n whereas that of MGM with Richardson smoother is approaching 1 as n increases.

$f(\theta)$	$2 - 2\cos\theta$		θ^2	
n	Richardson	damped-Jacobi	Richardson	damped-Jacobi
64	0.29297382	0.11414715	0.40202522	0.31879256
128	0.44943557	0.11809467	0.55478058	0.33836982
256	0.61484019	0.12755175	0.68098468	0.33169762
512	0.74128849	0.12531082	0.77651215	0.34075431
1024	0.81458679	0.12633265	0.84132052	0.33728271

Table 2: Comparison of the Richardson and the Damped-Jacobi smoothers.

We note that here the Richardson or damped Jacobi method is used as the smoother because they can take advantage of the Toeplitz structure of the given matrices. Hence the costs per iteration for these two smoothers are of order $O(n\log n)$ operations for Toeplitz matrices as compared to $O(n^2)$ operations for general matrices. However for the Gauss-Seidel smoother, it is not clear if it can make use of the Toeplitz structure of the matrices and reduce the cost per iteration to the same order.

6 Concluding Remarks

In this paper, we have partially answered the conjecture posed by Serra in [2]. For Toeplitz matrices generated by functions having zeros of order less than or equal to 2, we proved that TGM with Richardson method does converge. However, for MGM, it will be easier to use damped-Jacobi method as smoother than to use Richardson method, especially when n is not of the form $2^\ell - 1$. We remark that the convergence of MGM with damped-Jacobi method as smoother is discussed in Chan, Chang and Sun [4].

References

[1] G. Fiorentino and S. Serra, *Multigrid Methods for Toeplitz Matrices*, Calcolo, 28 (1991).

[2] S. Serra, *Multi-Iterative Methods*, Comp. Math. Applic., **26** 65–87 (1993).

[3] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, 2nd Ed., Chelsea, New York (1984).

[4] R. Chan, Q. Chang and H. Sun, *Multigrid Method for Ill-Conditioned Symmetric Toeplitz Systems*, SIAM J. Sci. Comput., to appear.

[5] J. Ruge and K. Stuben, Algebraic Multigrid, *in Multigrid Methods*, Ed: S. McCormick, 4, SIAM, Philadephia (1987).