Numer. Math. 75: 267–292 (1997)

Numerische Mathematik

On multisplitting methods for band matrices

Götz Alefeld[1], Ingrid Lenhardt[1], Günter Mayer[2]

[1] Institut für Angewandte Mathematik, Universität Karlsruhe, D-76128 Karlsruhe, Germany
[2] Fachbereich Mathematik, Universität Rostock, D-18051 Rostock, Germany

Received July 18, 1994 / Revised version received November 20, 1995
Dedicated to J.W. Schmidt, Dresden, on the occasion of his 65th birthday

Summary. We present new theoretical results on two classes of multisplitting methods for solving linear systems iteratively. These classes are based on overlapping blocks of the underlying coefficient matrix A which is assumed to be a band matrix. We show that under suitable conditions the spectral radius $\rho(H)$ of the iteration matrix H does not depend on the weights of the method even if these weights are allowed to be negative. For a certain class of splittings we prove an optimality result for $\rho(H)$ with respect to the weights provided that A is an M–matrix. This result is based on the fact that the multisplitting method can be represented by a single splitting $A = M - N$ which in our situation surprisingly turns out to be a regular splitting. Furthermore we show by numerical examples that weighting factors $\alpha \notin [0, 1]$ may considerably improve the convergence.

Mathematics Subject Classification (1991): 65F10

1. Introduction

To solve quadratic linear systems of equations

$$Ax = b\ , \quad A \text{ non-singular}, \tag{1}$$

iteratively, one often starts with a *splitting*

$$A = M - N\ , \quad M \text{ non-singular}. \tag{2}$$

One replaces A in (1) by (2) and ends up with the fixed point equation

$$x = M^{-1}Nx + M^{-1}b\ . \tag{3}$$

This equation forms the base for the iterative process

Correspondence to: Any of the authors

$$x^{k+1} = M^{-1}Nx^k + M^{-1}b\ , \quad k = 0, 1, \ldots\ . \tag{4}$$

Among many well–known methods like the Jacobi method, the Gauß–Seidel method, the relaxation methods JOR, SOR, SSOR, AOR, one also finds the *block Jacobi method* for which M has the form

$$M = \begin{pmatrix} A_{11} & & O \\ & A_{22} & \\ & & \ddots \\ O & & A_{pp} \end{pmatrix}$$

with the quadratic diagonal blocks A_{ii} being chosen from the same block partitioning

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pp} \end{pmatrix} \tag{5}$$

of A. Assuming $A_{\ell\ell} \in \mathbb{R}^{n_\ell \times n_\ell}$, $\ell = 1, \ldots, p$, to be nonsingular, M^{-1} exists, N is determined by $N := M - A$, and (4) can be realized in practice by solving in each step of iteration the p smaller subsystems

$$A_{\ell\ell} f_\ell(x^{k+1}) = f_\ell(Nx^k + b)\ , \quad \ell = 1, \ldots, p,$$

with $f_\ell : \mathbb{R}^n \to \mathbb{R}^{n_\ell}$ denoting the projection $f_\ell(x) := (x_{q_\ell+1}, \ldots, x_{q_\ell+n_\ell})^{\mathrm{T}} \in \mathbb{R}^{n_\ell}$ with $x = (x_i) \in \mathbb{R}^n$ and with $q_\ell := \sum_{i=1}^{\ell-1} n_i$. Choosing the lower triangular part of $A_{\ell\ell}$ for M yields a method which was called *Gauß–Seidel like method* or *block Gauß–Seidel type method* (cf. [18]). Both classes of methods show a natural parallelism since each of the occuring smaller subsystems can be solved independently from each other. These methods can also be viewed as particular multisplitting methods, a class of methods for solving (1) iteratively on a parallel computer. Multisplitting methods were introduced by O'Leary and White in [27] in the following way. (Cf. also [4], [6], [8]–[22], [24]–[26], [29]–[33], [35]–[39], and, in particular, [8], [9], [13], [35], [36] for systems of nonlinear equations.)

Definition 1. Let $A = M_\ell - N_\ell \in \mathbb{R}^{n\times n}$, $\ell = 1, \ldots, p$, be splittings of A, and let $E_\ell \in \mathbb{R}^{n\times n}$, the so–called weights, satisfy $\sum_{\ell=1}^{p} E_\ell = I$, where I denotes the $n \times n$ identity matrix. Then the iterative process

$$\begin{cases} M_\ell y^{k,\ell} = N_\ell x^k + b\ , & \ell = 1, \ldots, p \\ x^{k+1} = \sum_{\ell=1}^{p} E_\ell y^{k,\ell} & , \quad k = 0, 1, \ldots, \end{cases} \tag{6}$$

is called a *multisplitting method.*

Note that $\sum_{\ell=1}^{p} E_\ell = I$ ensures that in case of convergence the fixed point of the multisplitting method is the solution of (1).

We set

$$M_\ell = \begin{pmatrix} D_1^{(\ell)} & 0 & 0 \\ 0 & A_{\ell\ell} & 0 \\ 0 & 0 & D_2^{(\ell)} \end{pmatrix}, \quad E_\ell = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_\ell & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7}$$

with $A_{\ell\ell} \in \mathbb{R}^{n_\ell \times n_\ell}$ from (5), with I_ℓ being the $n_\ell \times n_\ell$ identity matrix, and with $D_1^{(\ell)}$, $D_2^{(\ell)}$ being any non–singular matrices of the appropriate dimension. In theoretical investigations we usually choose diagonal matrices for $D_1^{(\ell)}$ and $D_2^{(\ell)}$. One recognizes at once that one gets the iterates of the block Jacobi method, and it is a trivial task to verify that the iterates of the Gauß–Seidel like method can be obtained as a particular multisplitting method, too. (It is clear that (7) has to be modified appropriately in the cases $\ell = 1$ and $\ell = p$.) Of course, (6) is much more general. For example, one can expand $A_{\ell\ell}$ for each M_ℓ, allowing the $n_\ell \times n_\ell$ diagonal blocks $A_{\ell\ell}$ of A now to overlap with $A_{\ell-1,\ell-1}$ and with $A_{\ell+1,\ell+1}$. The weights E_ℓ are then assumed to have the form

$$E_\ell = E_\ell(\alpha) := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & (1-\alpha)I_1^{(\ell)} & 0 & 0 & 0 \\ 0 & 0 & I_2^{(\ell)} & 0 & 0 \\ 0 & 0 & 0 & \alpha I_3^{(\ell)} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \alpha \in \mathbb{R}, \tag{8}$$

where $I_j^{(\ell)}$, $j = 1, 2, 3$, are identity matrices corresponding to the upper overlapping part, the non–overlapping part and the lower overlapping part of $A_{\ell\ell}$, respectively, with the usual modifications for $\ell \in \{1, p\}$. In the sequel we will refer to this method (i. e., M_ℓ from (7), E_ℓ from (8)) as the *overlapping block Jacobi method*. Analogously one can define the *overlapping Gauß–Seidel like method*.

For these two classes of methods and for band matrices A we address the following questions.

a) How does the choice of the parameter α in (8) influence the speed of convergence of (6)?
b) How does (6) behave if the overlapping part of $A_{\ell\ell}$ increases?
c) Up to now only $\alpha \in [0, 1]$ was considered in the literature. In this case $E_\ell \geq 0$ holds. How does (6) behave if one chooses α outside $[0, 1]$ in (8)?

For particular situations we will answer some of these questions theoretically, and we will add a part of our numerical experiments to show what may happen. Only for this reason the discussed examples are chosen as simple as possible. As a consequence they are (perhaps with the exception of the last one) not of great practical relevance. (This was also pointed out by one of the referees.) For the overlapping block Jacobi method we will prove that the choice of the weighting parameter α does not influence the iteration matrix H if the bandwidth β, the

block size m of the non–overlapping method and the size ovl of the overlapping part are related in a certain manner. This holds also for $\alpha \notin [0,1]$. For the same relation between the bandwidth β, the blocksize m of the non–overlapping method and the size ovl of the overlapping part, and for the overlapping Gauß–Seidel like method we will find that for M–matrices A the multisplitting method (6) can be represented as a single so–called *regular* splitting (cf. Sect. 2). This is a surprising result since even for M–matrices generally a multisplitting method cannot be described by a single *regular* splitting. Now, due to a result of Elsner [6], comparisons are possible which show that for fixed ovl the spectral radius $\rho(H(\alpha))$ of the iteration matrix $H = H(\alpha)$ decreases with increasing $\alpha \in [0,1]$. Numerical examples indicate that $\rho(H(\alpha))$ can apparently become smaller if α quits the range $[0,1]$. (The situation is similar to that of the well–known successive overrelaxation method (SOR–method): It is fairly easy to show by using nonnegative matrix theory that the spectral radius of the SOR–matrix is decreasing if the relaxation parameter ω is chosen from the interval $[0,1]$. However, beyond $\omega = 1$ this theory is not applicable, and it is well–known that the optimal ω is in most cases greater than one. At the moment we do not have ideas how to investigate the case $\alpha \notin [0,1]$ in the case of the Gauß–Seidel like multisplitting method or for other multisplitting methods. It is one of the main purposes of this paper to show – by numerical examples – that $\alpha \notin [0,1]$ can improve the speed of convergence considerably.)

We have arranged our paper as follows. In Sect. 2 we recall several known results on multisplittings, in Sect. 3 we consider the overlapping block Jacobi method, and in Sect. 4 we deal with the Gauß–Seidel like method. Our numerical experiments are incorporated in both of these sections.

We stress the fact that promising results with respect to the practical relevance of multisplitting methods have been shown in different contexts. For example, one can use multisplitting methods as smoothers in multigrid methods or as preconditioners for conjugate gradient methods. Earlier results can be found in [1], [29], [38]. For a very recent result in which the importance of multisplitting methods as preconditioners has been shown see [3]. Moreover, multisplitting methods play an important role in connection with the so–called *multisplitting waveform relaxation algorithm*. This is an algorithm for solving the linear initial value problem

$$x'(t) + Ax(t) = f(t), \quad x(0) = x_0,$$

where $A \in \mathbb{R}^{n\times n}$ is very large. The function f is assumed to be continuous, $t \in [0,T]$, $x \in C^1([0,T];\mathbb{R}^n)$, $x_0 \in \mathbb{R}^n$. The given system is decomposed in a finite number of subsystems and these are discretized by appropriate methods on the interval $[0,T]$. For details see [18], [20], [23], [30], [31] and [34]. In this way one gets the multisplitting waveform relaxation method for solving ordinary differential equations on parallel computers. In order to compare this method with algorithms for solving ordinary differential equations on a serial computer, it is essential to solve the arising subsystems by an efficient serial ODE solver. See [5] for an implementation of the method and for time comparisons with a fast serial ODE solver.

Our theoretical results are independent of a particular parallel computer. It is clear, however, that the architecture of the computer and therefore the implementation of multisplitting methods both influence the practical behaviour of these methods. It is not the purpose of this paper to discuss this important practical problem.

2. Preliminaries

We first list some notations which we will use throughout the paper without further reference.

If not otherwise stated we denote the entries of a vector $x \in \mathbb{R}^n$ by x_i, those of a matrix $A \in \mathbb{R}^{n\times n}$ by a_{ij}. We always reserve the letter A for a real $n \times n$ matrix. If all entries of x or A are non–negative, we write $x \geq 0$ and $A \geq 0$, respectively, calling x and A *non–negative*; $A \leq B$ is defined by $B - A \geq 0$. We term A an *M–matrix* if $a_{ij} \leq 0$ for $i \neq j$ and if A^{-1} exists and is non–negative.

By the non–negative integer β we abbreviate the *(half–) bandwidth* of A, i.e., $\beta := \max\{|i-j| \mid a_{ij} \neq 0\}$. By $\rho(A)$ we mean the spectral radius of A, by $|S|$ we denote the number of elements contained in a finite set S.

Definition 2. a) The representation $A = M - N$, M non–singular, is called a *splitting* of A.

b) The splitting $A = M - N$ is called a *weak regular* one if $M^{-1} \geq 0$ and $M^{-1}N \geq 0$. It is called a *regular splitting* if $M^{-1} \geq 0$ and $N \geq 0$.

It is obvious that a regular splitting is a weak regular one. We now list several auxiliary results starting with a well–known theorem on M–matrices.

Theorem 1. *([2], [7], [28]) a) If $a_{ij} \leq 0$ for $i \neq j$, then A is an M–matrix if and only if there exists a vector $x > 0$ such that $Ax > 0$ holds.*

b) Let A be an M–matrix and let $A \leq B$ with $b_{ij} \leq 0$ for $i \neq j$. Then B is an M–matrix, too.

Some of our results in Sect. 4 are based on the following comparison theorem due to Elsner [6].

Theorem 2. *Let A be non–singular with $A^{-1} \geq 0$, and let $A = M - N = \hat{M} - \hat{N}$ be two weak regular splittings of which at least one is a regular splitting.*

a) If $\hat{M}^{-1} \leq M^{-1}$ then $\rho(M^{-1}N) \leq \rho(\hat{M}^{-1}\hat{N}) < 1$.

b) If $M \leq \hat{M}$, or, equivalently, $N \leq \hat{N}$, then $\hat{M}^{-1} \leq M^{-1}$, hence $\rho(M^{-1}N) \leq \rho(\hat{M}^{-1}\hat{N}) < 1$.

In view of comparison theorems it is an interesting question whether the multisplitting method (6) can be represented by means of a single splitting $A = M - N$. If this is possible, M must have the form

$$M = \left(\sum_{\ell=1}^{p} E_\ell M_\ell^{-1} \right)^{-1}, \tag{9}$$

and $N := M - A$. This is seen from (6) and (4) when eliminating $y^{k,\ell}$ in (6) in order to get

$$x^{k+1} = \underbrace{\left(\sum_{\ell=1}^{p} E_\ell M_\ell^{-1} N_\ell\right)}_{= M^{-1}N} x^k + \underbrace{\left(\sum_{\ell=1}^{p} E_\ell M_\ell^{-1}\right)}_{= M^{-1}} b \tag{10}$$

with the iteration matrix

$$H = M^{-1}N = \sum_{\ell=1}^{p} E_\ell M_\ell^{-1} N_\ell \,. \tag{11}$$

The representation (10) is, of course, only suited for theoretical considerations. It is a standard result in numerics that (10) converges for all starting vectors x^0 to the solution of (1) if and only if $\rho(H) < 1$.

The following theorem gathers results of [6] and [27], cf. also [18].

Theorem 3. *Let $A = M_\ell - N_\ell$, $\ell = 1, \ldots, p$, be weak regular splittings and let $H := \sum_{\ell=1}^{p} E_\ell M_\ell^{-1} N_\ell$, $G := \sum_{\ell=1}^{p} E_\ell M_\ell^{-1}$, $E_\ell \geq 0$, $\ell = 1, \ldots, p$.*

a) *If A^{-1} is non–negative then $\rho(H) < 1$.*
b) *If $\rho(H) < 1$ then G is non–singular, $A^{-1} \geq 0$, and the splitting $A = M - N$ with $M := G^{-1}$, $N := M - A$ is a weak regular one. In particular, $H = M^{-1}N \geq 0$.*

Proof. a) is proved in [27]; cf. also [18].
b) With the exception of $A^{-1} \geq 0$ the assertions are also proved in [18]. We will shortly repeat the steps.
Since $N_\ell = M_\ell - A$ we get $H = I - GA$, hence $\rho(H) < 1$ implies G to be non–singular. Thus the splitting $A = M - N$ exists, and (9), (11) together with the hypotheses of the theorem show $M^{-1} \geq 0$, $H = M^{-1}N \geq 0$ which proves the splitting to be weak regular. The non–negativity of A^{-1} follows from $A^{-1} = (I - H)^{-1}M^{-1}$ when representing $(I - H)^{-1}$ as a Neumann series. □

Unfortunately, regular splittings $A = M_\ell - N_\ell$, $\ell = 1, \ldots, p$, do not imply a regular splitting $A = M - N$ for the representation of A in Theorem 3. This was already shown in [6] by a counterexample.

For multisplitting methods there exist, among others, the following comparison theorems.

Theorem 4. ([6]) *Let $A^{-1} \geq 0$, $E_\ell \geq 0$ for $\ell = 1, \ldots, p$, H as in (11), and let $A = M_\ell - N_\ell$, $\ell = 1, \ldots, p$, be weak regular splittings.*

a) *If $A = \underline{M} - \underline{N}$ is a regular splitting satisfying $\underline{M} \leq M_\ell$, $\ell = 1, \ldots, p$, then $\rho(\underline{M}^{-1}\underline{N}) \leq \rho(H) < 1$.*

b) *If* $A = \overline{M} - \overline{N}$ *is a regular splitting satisfying* $M_\ell \leq \overline{M}$, $\ell = 1, \ldots, p$, *then* $\rho(H) \leq \rho(\overline{M}^{-1}\overline{N}) < 1$.

Corollary 1. ([6], [17], [26]) *Let* $E_\ell \geq 0$ *for* $\ell = 1, \ldots, p$, *and let* $A = D - L - U$ *be an M–matrix for which* D, L, U *denote the diagonal part, the lower triangular part and the upper triangular part, respectively. If* $A = M_\ell - N_\ell$, $\ell = 1, \ldots, p$, *are splittings of* A *satisfying* $D - L \leq M_\ell \leq D$, $\ell = 1, \ldots, p$, *then the splittings* $A = M_\ell - N_\ell$ *are regular, and*

$$\rho(H_{GS}) \leq \rho(H) \leq \rho(H_J) < 1$$

holds, where $H_{GS} := (D - L)^{-1}U$, $H_J := D^{-1}(L + U)$ *are the iteration matrices of the standard Gauß–Seidel method and the standard Jacobi method, respectively.*

Theorem 5. ([18]) *Let* A *be an M–matrix and let* H, $\hat{H}$ *be the iteration matrices of the non–overlapping and the corresponding overlapping block Jacobi multisplitting method, using in both cases the same weighting matrices* E_ℓ *from* (7), *i.e., the weights of the non–overlapping block Jacobi multisplitting method. Then*

$$\rho(\hat{H}) \leq \rho(H) \,. \tag{12}$$

By (12) one can expect a faster convergence for the overlapping method as compared with the non–overlapping one. This justifies overlapping blocks. But note the particular form of E_ℓ ! Theorem 5 says nothing on the behaviour of $\rho(\hat{H})$ when the overlapping parts increase.

Theorem 6. ([18]) *The assertions of Theorem* 5 *hold if the block Jacobi multisplitting method is replaced by the Gauß–Seidel like multisplitting method.*

3. The overlapping block Jacobi multisplitting method

In this section we prove a result on the overlapping block Jacobi multisplitting method which we already described in Sect. 1 and for which we now want to repeat a more precise definition.

Definition 3. ([11]) 3 Let $\{S_1, \ldots, S_p\}$ be a partition of $\{1, \ldots, n\}$, i.e., the S_ℓ are pairwise disjoint non–empty subsets of $\{1, \ldots, n\}$ so that $\bigcup_{\ell=1}^{p} S_\ell = \{1, \ldots, n\}$. Assume that the condition

$$i \in S_\ell \,, \quad j \in S_{\ell'} \,, \quad i < j \quad \Rightarrow \quad \ell \leq \ell' \tag{13}$$

holds. In addition, let $S_\ell \subseteq T_\ell \subseteq \{1, \ldots, n\}$ for $\ell = 1, \ldots, p$, with T_ℓ satisfying

$$i \in T_\ell \,, \quad j \in T_{\ell'} \,, \quad i < j \quad \Rightarrow \quad \ell \leq \ell' \tag{14}$$

and

$$T_\ell \cap T_{\ell'} = \emptyset \quad \text{if} \quad |\ell - \ell'| > 1 \,. \tag{15}$$

Let $T_0 := T_{p+1} := \emptyset$.

a) Define $M_\ell = ((M_\ell)_{ij})$, N_ℓ, $E_\ell = ((E_\ell)_{ij})$ for $\ell = 1, \ldots, p$ by

$$(M_\ell)_{ij} := \begin{cases} a_{ij} & \text{if } i, j \in S_\ell \\ a_{ii} & \text{if } i = j \notin S_\ell \\ 0 & \text{otherwise} \end{cases},$$

$$N_\ell := M_\ell - A,$$

$$(E_\ell)_{ij} := \begin{cases} 1 & \text{if } i = j \in S_\ell \\ 0 & \text{otherwise} \end{cases}.$$

If the matrices M_ℓ all are non–singular, then the corresponding multisplitting method (6) is called a *(non–overlapping) block Jacobi multisplitting method.*

b) Define $M_\ell = ((M_\ell)_{ij})$, N_ℓ, $E_\ell = ((E_\ell)_{ij})$ for $\ell = 1, \ldots, p$ by

$$(M_\ell)_{ij} := \begin{cases} a_{ij} & \text{if } i, j \in T_\ell \\ a_{ii} & \text{if } i = j \notin T_\ell \\ 0 & \text{otherwise} \end{cases},$$

$$N_\ell := M_\ell - A,$$

$$(E_\ell)_{ij} := \begin{cases} 0 & \text{if } i \neq j \text{ or if } i = j \notin T_\ell \\ 1 & \text{if } i = j \in T_\ell \backslash (T_{\ell-1} \cup T_{\ell+1}) \\ \alpha_i & \text{if } i = j \in T_\ell \cap T_{\ell+1} \\ 1 - \alpha_i & \text{if } i = j \in T_{\ell-1} \cap T_\ell \end{cases},$$

where $\alpha_i \in \mathbb{R}$.

If the matrices M_ℓ all are non–singular then the corresponding multisplitting method (6) is called an *overlapping block Jacobi multisplitting method.*

Note that (13), (14) mean in particular that the sets S_ℓ and T_ℓ, respectively, contain only successive integers from $\{1, \ldots, n\}$. Condition (15) means that only two successive blocks may overlap. In this respect our definition differs slightly from that in [18], although in practice (15) is fulfilled. The matrices E_ℓ in Definition 1b), are defined a little bit more general than $E_\ell(\alpha)$ in (8). We will return to (8) in our examples.

Theorem 7. *Let A be a band matrix with bandwidth β and let S_ℓ, T_ℓ, M_ℓ, N_ℓ, E_ℓ be defined as in Definition* 3b), *and H as in* (11). *Assume*

$$\min S_\ell = \min T_\ell, \quad \ell = 1, \ldots, p. \tag{16}$$

If

$$\beta \leq |S_\ell| - |S_\ell \cap T_{\ell-1}|, \quad \ell = 1, \ldots, p, \tag{17}$$

then for the overlapping block Jacobi method the spectral radius $\rho(H)$ does not depend on the values of the α_i.

If all the sets S_ℓ have the same number of elements and if the overlapping parts have the same size then Theorem 7 implies at once the subsequent corollary.

Corollary 2. *With the notations of Theorem 7 let the following conditions hold.*

$$\left.\begin{array}{l}\min S_\ell = \min T_\ell \\ |S_\ell| = m = \text{constant}\end{array}\right\} \quad \ell = 1, \ldots, p,$$

$$|S_\ell \cap T_{\ell-1}| = ovl = \text{constant}\,, \quad \ell = 2, \ldots, p.$$

If

$$\beta \leq m - ovl$$

then $\rho(H)$ does not depend on the values α_i of E_ℓ.

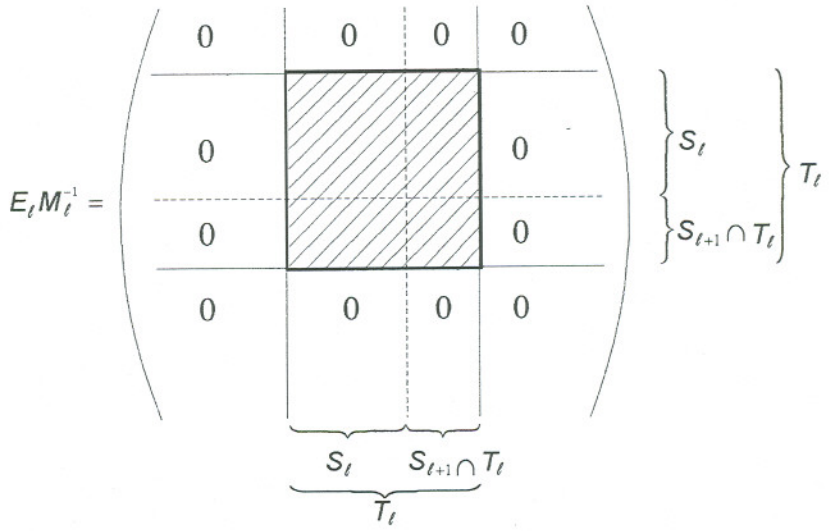

Fig. 1.

Proof of Theorem 7. By the Definition 3b) the matrix $E_\ell M_\ell^{-1}$ has the block form in Fig. 1, where here and in the subsequent figures the dashed area indicates entries which are (in general) non–zero. The matrix N_ℓ has the block form in Fig. 2, where the parts of N_ℓ outside the horizontal strip with row indices from T_ℓ are irrelevant in view of the multiplication with $E_\ell M_\ell^{-1}$. The dashed areas result from (17) which implies

$$\beta \leq |S_{\ell-1}| - |S_{\ell-1} \cap T_{\ell-2}| \tag{18}$$

and

$$\beta \leq |S_{\ell+1}| - |S_{\ell+1} \cap T_\ell|\,. \tag{19}$$

The inequalities (17) prevent the band of N_ℓ from running into the area $T_\ell \times (S_{\ell-1} \cap T_{\ell-2})$ and $T_\ell \times (S_{\ell+2} \cap T_{\ell+1})$, respectively, where '$\times$' denotes the cartesian product. It is easy to see that the product $E_\ell M_\ell^{-1} N_\ell$ has the block form in Fig. 3, where only the double dashed areas may contain numbers which are multiplied by the weighting parameters α_i and $1 - \alpha_i$, respectively. Due to this fact $H = \sum_{\ell=1}^{p} E_\ell M_\ell^{-1} N_\ell$ has the block form in Fig. 4. The integers ℓ indicate, from which splitting $A = M_\ell - N_\ell$ the dashed and double dashed regions come.

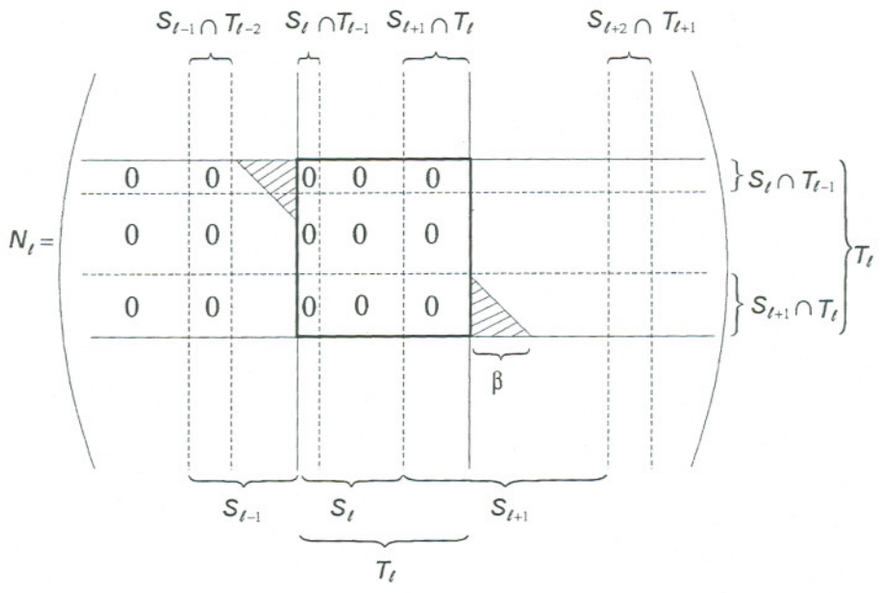

Fig. 2.

Evaluating $\det(\lambda I - H)$ along the columns with indices from $\bigcup_{\ell=1}^{p-1} (S_{\ell+1} \cap T_\ell)$ shows that $\det(\lambda I - H)$ does not depend on the elements of the double dashed areas, which proves the assertion. □

Theorem 7 normally becomes false if (17) does no longer hold. This can be seen by the numerical results of the following example which we will also use in Sect. 4.

We obtained our numerical multisplitting results on the Parsytec parallel computers of the University of Paderborn. The spectral radii were computed on the IBM 3090 serial computer of the University of Karlsruhe using the NAG library.

Example 1. This example was already published in [15] and was used in a modified form in [18]. It is based on a linear system with a banded symmetric Toeplitz matrix. For such particular matrices there are, of course, more favorable solution methods available. (This was also pointed out by one of the referees.) However, this system is also used in Example 3 in order to show that the spectral radius of a Gauß–Seidel like multisplitting method can be dramatically decreased for values of α outside the interval $[0, 1]$.

Let $A \in \mathbb{R}^{n \times n}$ be a band matrix which is defined by

$$a_{ij} := \begin{cases} 2 & \text{if } i = j \\ -2^{-|i-j|} & \text{if } 0 < |i-j| \leq \beta \\ 0 & \text{otherwise} \end{cases} . \tag{20}$$

The matrix A is a Stieltjes matrix, i.e., a symmetric M–matrix. Let $b \in \mathbb{R}^n$ be defined by

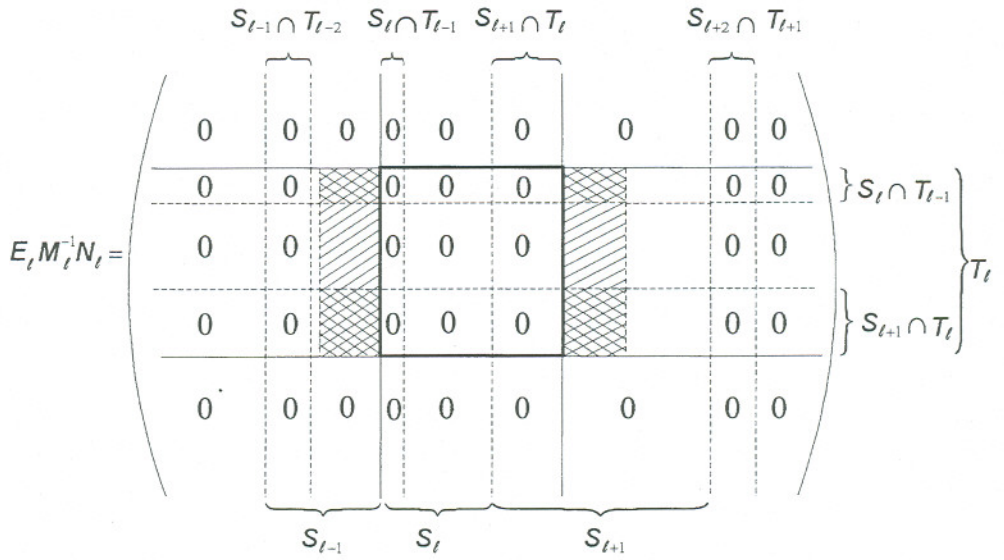

Fig. 3.

$$b_i := \begin{cases} \dfrac{1}{2^\beta} + \dfrac{1}{2^{i-1}}\,, & 1 \le i \le \beta \\ \dfrac{1}{2^{\beta-1}}\,, & \beta < i \le n-\beta \\ \dfrac{1}{2^\beta} + \dfrac{1}{2^{n-i}}\,, & n-\beta < i \le n \end{cases} \quad . \tag{21}$$

Then the solution x of (1) is $x = e := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$.

If J denotes the iteration matrix of the standard Jacobi method associated with A (i.e., in particular, $|S_\ell| = 1$, $\ell = 1, \ldots, n$) we have

$$\rho(J) \le \|J\|_\infty = 1 - 2^{-\beta}$$

with the row sum norm $\| \cdot \|_\infty$. Note that $\|J\|_\infty$ can be made arbitrarily close to the critical value 1 if the bandwidth β is chosen large enough. Because of Theorem 2, $\rho(J)$ and therefore $\|J\|_\infty$ certainly represent upper bounds for $\rho(\hat{H})$ where $\hat{H}$ denotes the iteration matrix of the non–overlapping block Jacobi method.

Let $\alpha_i = \alpha$ for all α_i in E_ℓ, i. e., let $E_\ell = E_\ell(\alpha)$ as in (8). In order to illustrate the influence of α on the iteration matrix $H = H(\alpha)$ of the overlapping block Jacobi multisplitting method we chose

(22) $n = 16384$, $\beta = 5$, $p = 128$, $|S_\ell| = 128$, T_ℓ as in Corollary 2, $\ell = 1, \ldots, p$.

Then all Jacobi blocks have the same block size m, and Corollary 2 applies with $m = |S_\ell| = 128$. As long as the number ovl of the overlapping rows fulfills

$$ovl \le m - \beta = 123\,,$$

the spectral radius $\rho(H(\alpha))$ is independent of α by Theorem 7. If one increases ovl beyond 123, the weighting parameter α influences $\rho(H(\alpha))$. This is indicated

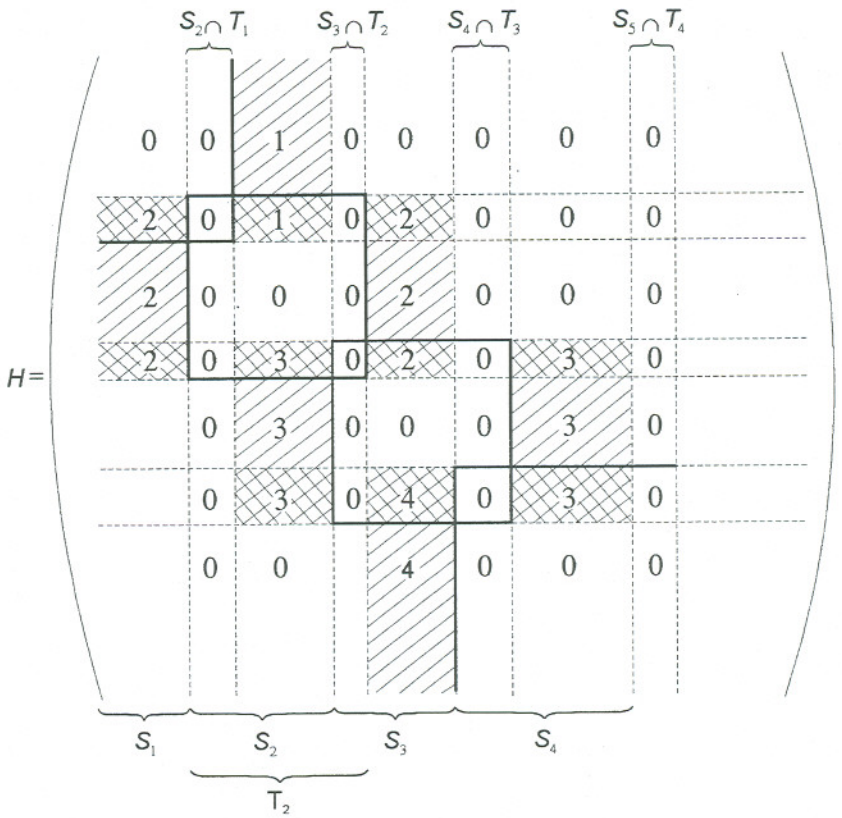

Fig. 4.

by the following Fig. 5 for which we chose b as in (21), $x^0 := 0$, and $ovl = 40$, 125, 127, and 128, respectively. We stopped the iteration when $\|x^k - e\|_\infty \leq 10^{-5}$ was fulfilled and plotted the corresponding number of iterates versus the weighting parameter $\alpha \in [-2, 3]$. Figure 5 shows a symmetry with respect to $\alpha = \frac{1}{2}$.

We now consider an analogous situation for $n = 256$, $\beta = 5$ and $p = 16$. Here $m = |S_\ell| = 16$, hence $\rho(H(\alpha))$ is independent of α if $ovl \leq 11$, but it depends on ovl. How it depends on α for $ovl = 14$, $ovl = 15$, and $ovl = 16$, respectively, is shown in Fig. 6.

Since Theorem 7 implies $\rho(H(\alpha)) = \rho(H(0))$ for $\alpha \in \mathbb{R}$ and $0 \leq ovl \leq 11$, it can be seen from Fig. 7 of the subsequent example that the straight line diminishes its height if ovl increases from $ovl = 0$ to $ovl = 11$.

Example 2. With the same matrix A and the same vector b as in Example 1, and with the parameters from (22) we illustrate first the content of Theorem 5. To this end we set $\alpha = \alpha_i = 0$ and we vary the number ovl of the overlapping rows (and columns, respectively). As in Example 1 we started the iterations with $x^0 = 0$, and we stopped it when $\|x^k - e\|_\infty \leq 10^{-5}$ was fulfilled. We want to compare the number of iterates for $ovl \neq 0$ with that for $ovl = 0$. The results in Table 1 confirm $\rho(\hat{H}) \leq \rho(H)$ if H, $\hat{H}$ denote the iteration matrices of the non–overlapping

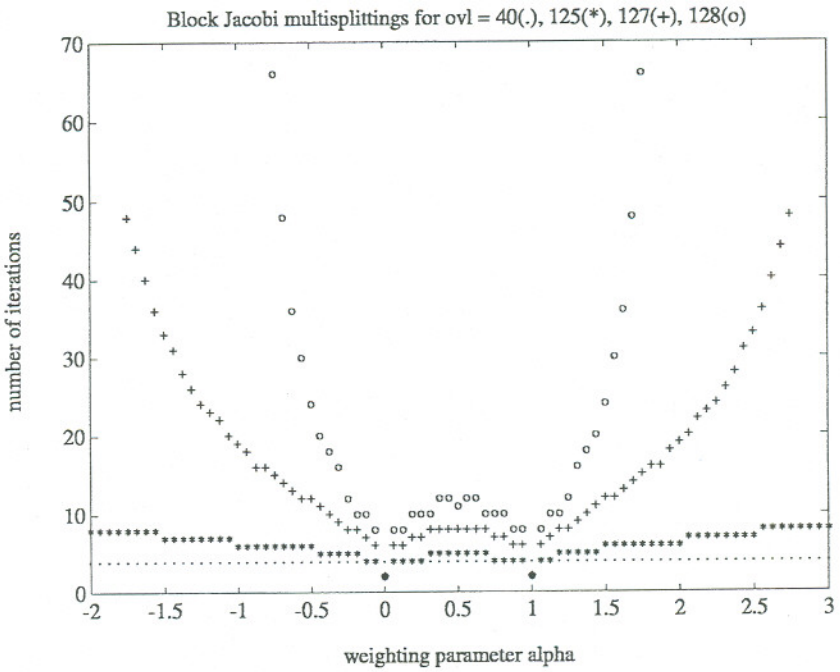

Fig. 5. Influence of the weighting parameter $\alpha = \alpha_i$ on the maximal number of iterations for the overlapping block Jacobi multisplitting method with $n = 16\,384$, $\beta = 5$, $p = m = 128$, and $ovl \in \{40, 125, 127, 128\}$

and the corresponding overlapping block Jacobi multisplitting method with the weights $E_\ell = E_\ell(0)$ in both cases. These results also indicate a monotonous behaviour of $\rho(H)$ with respect to *ovl* just as in one of the examples in [18]. It was A. Frommer who asked whether this monotonicity holds in general. We answer this question negatively by Table 2 in which we changed β from 5 to 11 . (One of the referees claimed that the non–monotonicity may be caused by effects due to rounding errors in this example. We could neither prove nor disprove this conjecture.) The number of iterations shows a non–monotonous behaviour with respect to *ovl*. The same phenomenon can also be seen from Fig. 7 in which we plotted the spectral radius $\rho(\hat{H})$ versus *ovl* for $n = 256$, $p = m = 16$, $\beta = 5$ and $\alpha = \alpha_i = 0$.

4. The overlapping Gauß–Seidel like multisplitting method

Analogously to the previous section we start with a precise definition of the method which we want to consider.

Definition 4. Let $\{S_1, \ldots, S_p\}$, $\{T_0, \ldots, T_{p+1}\}$ be defined as in Definition 3.

a) Let $M_\ell = ((M_\ell)_{ij})$, N_ℓ, $E_\ell = ((E_\ell)_{ij})$ be given by

$$(M_\ell)_{ij} := \begin{cases} a_{ij} & \text{if} \quad i,j \in S_\ell \quad \text{and} \quad i \geq j \\ a_{ii} & \text{if} \quad i = j \notin S_\ell \\ 0 & \text{otherwise} \end{cases},$$

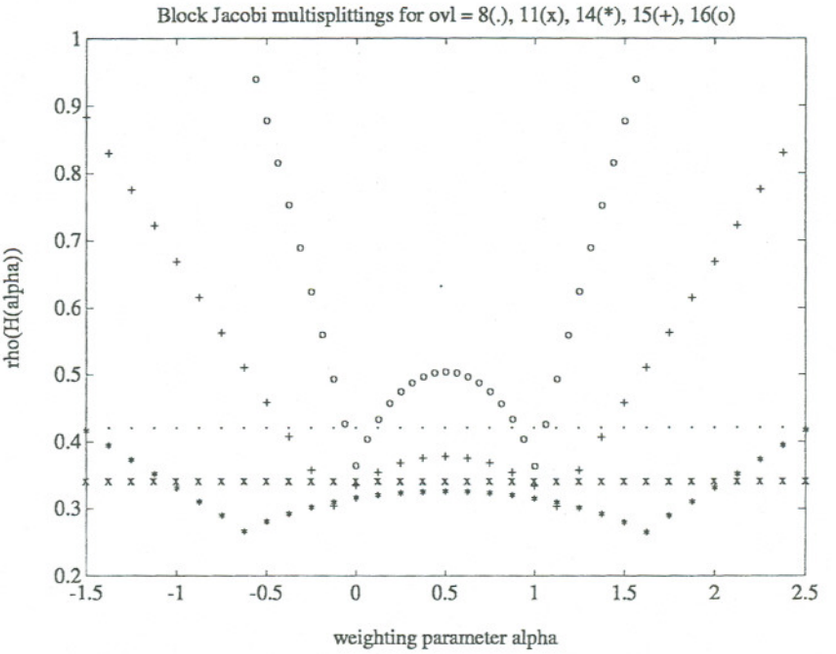

Fig. 6. Influence of the weighting parameter $\alpha = \alpha_i$ on the spectral radius $\rho(H(\alpha))$ for the overlapping block Jacobi multisplitting method with $n = 256$, $\beta = 5$, $p = m = 16$, and $ovl \in \{8, 11, 14, 15, 16\}$

$$N_\ell := M_\ell - A,$$

$$(E_\ell)_{ij} := \begin{cases} 1 & \text{if} \quad i = j \in S_\ell \\ 0 & \text{otherwise} \end{cases}$$

for $\ell = 1, \ldots, p$. If the matrices M_ℓ all are non–singular, then the corresponding multisplitting method *(6)* is called a *(non–overlapping) Gauß–Seidel like multisplitting method.*

b) Let $M_\ell = ((M_\ell)_{ij})$, N_ℓ, $E_\ell = ((E_\ell)_{ij})$ be given by

$$(M_\ell)_{ij} := \begin{cases} a_{ij} & \text{if} \quad i,j \in T_\ell \quad \text{and} \quad i \geq j \\ a_{ii} & \text{if} \quad i = j \notin T_\ell \\ 0 & \text{otherwise} \end{cases},$$

$$N_\ell := M_\ell - A,$$

$$(E_\ell)_{ij} := \begin{cases} 0 & \text{if} \quad i \neq j \quad \text{or if} \quad i = j \notin T_\ell \\ 1 & \text{if} \quad i = j \in T_\ell \backslash (T_{\ell-1} \cup T_{\ell+1}) \\ \alpha_i & \text{if} \quad i = j \in T_\ell \cap T_{\ell+1} \\ 1 - \alpha_i & \text{if} \quad i = j \in T_{\ell-1} \cap T_\ell \end{cases}, \quad \alpha_i \in \mathbb{R},$$

where $\ell = 1, \ldots, p$.

If all the matrices M_ℓ are non–singular then the corresponding multisplitting method *(6)* is called an *overlapping Gauß–Seidel like multisplitting method.*

It is obvious that the remarks following Definition 3 also hold for Definition 4. We gather the parameters α_i in Definition 4b) in a vector $\alpha \in \mathbb{R}^q$ where $q =$

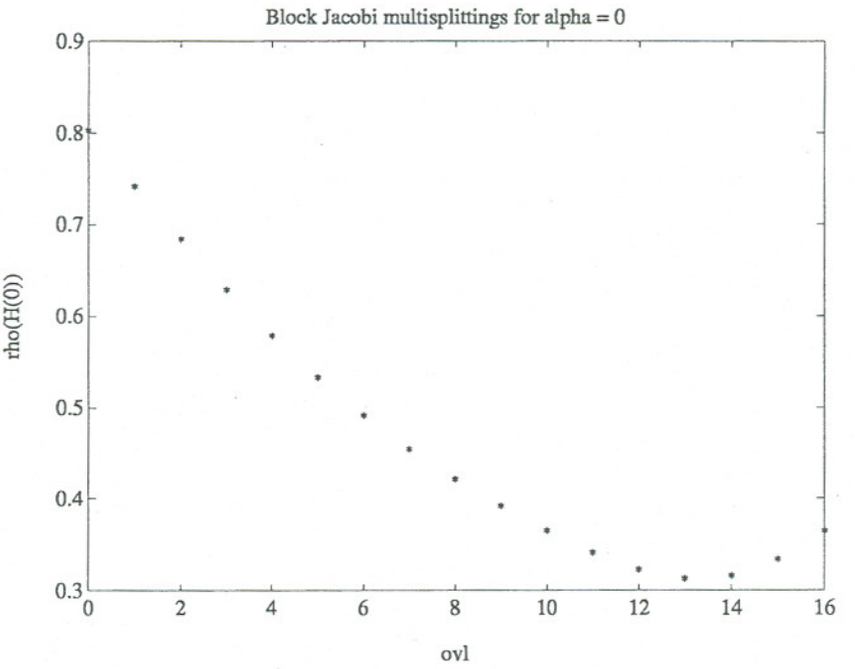

Fig. 7. Influence of the overlapping parameter *ovl* on the spectral radius $\rho(H(\alpha))$ for the overlapping block Jacobi multisplitting method with $n = 256$, $\beta = 5$, $p = m = 16$, $\alpha = \alpha_i = 0$

$|\bigcup_{\ell=2}^{p}(S_\ell \cap T_{\ell-1})|$. The indices of the components α_i are taken from the set $\bigcup_{\ell=2}^{p}(S_\ell \cap T_{\ell-1}) = \bigcup_{\ell=2}^{p}(T_\ell \cap T_{\ell-1})$.

Theorem 8. *Let A be an M–matrix and a band matrix with bandwidth β. Let S_ℓ, T_ℓ, M_ℓ, N_ℓ, E_ℓ be defined as in Definition* 4b) *and $H = H(\alpha)$ as in* (11). *Assume*

$$\min S_\ell = \min T_\ell\,, \quad \ell = 1, \ldots, p\,. \tag{23}$$

If

$$\beta \le |S_\ell| - |S_\ell \cap T_{\ell-1}|\,, \quad \ell = 1, \ldots, p\,, \tag{24}$$

then with

$$q := \left|\bigcup_{\ell=2}^{p}(S_\ell \cap T_{\ell-1})\right|$$

the following assertions hold for the overlapping Gauß–Seidel like multisplitting method.

a) *The splittings $A = M_\ell - N_\ell$, $\ell = 1, \ldots, p$, are regular.*
b) *For any $\alpha = (\alpha_i) \in \mathbb{R}^q$ with $\alpha_i \in [0, 1]$ for each i, the Gauß–Seidel like multisplitting method can be based on a single splitting*

$$A = M(\alpha) - N(\alpha)\,. \tag{25}$$

This splitting is regular.

Table 1. Influence of the overlapping parameter *ovl* on the maximal number of iterations for the overlapping block Jacobi multisplitting method with $n = 16384$, $\beta = 5$, $p = m = 128$, $\alpha = \alpha_i = 0$

ovl	*iterations*	*time*	*% iterations*	*% time*
0	36	52.85 s	100.00	100.00
1	27	40.21 s	75.00	76.08
2	22	34.19 s	61.11	64.69
3	18	28.62 s	50.00	54.15
4	16	25.03 s	44.44	47.36
5	14	22.38 s	38.89	42.35
7	12	19.67 s	33.33	37.22
9	10	18.54 s	27.78	35.08
12	8	15.79 s	22.22	29.88
15	7	12.96 s	19.44	24.52
20	6	12.85 s	16.67	24.31
30	4	9.65 s	11.11	18.26
70	3	10.02 s	8.33	18.96
100	2	8.92 s	5.56	16.88
120	2	10.13 s	5.56	19.17
125	2	10.51 s	5.56	19.89
128	2	10.63 s	5.56	20.11

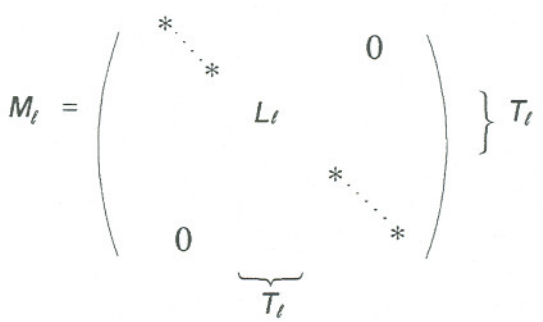

Fig. 8.

c) *For any* $\alpha = (\alpha_i)$, $\hat{\alpha} = (\hat{\alpha}_i) \in \mathbb{R}^q$ *with* α_i, $\hat{\alpha}_i \in [0, 1]$ *for each* i *and with* $\alpha \leq \hat{\alpha}$ *we get*

$$\rho(H(\hat{\alpha})) \leq \rho(H(\alpha))$$

for the iteration matrices $H = H(\alpha)$ *and* $H = H(\hat{\alpha})$. *In particular, the Gauß–Seidel like multisplitting methods are optimal (with respect to* $\alpha_i \in [0, 1]$*) if* $\alpha_i = 1$ *holds for each* i.

Proof. a) By Theorem 1a) there is a positive vector x such that $Ax > 0$, whence $M_\ell x > 0$. By the same theorem, M_ℓ is an M–matrix, hence $M_\ell^{-1} \geq 0$. By the sign pattern of A, the matrix N_ℓ is non–negative, therefore $A = M_\ell - N_\ell$ is a regular splitting.

b) By Theorem 3a) we have $\rho(H) < 1$. Therefore the matrix G of Theorem 3 is non–singular by part b) of this theorem. Hence (25) exists with $M(\alpha) := G^{-1}$

Table 2. Influence of the overlapping parameter *ovl* on the maximal number of iterations for the overlapping block Jacobi multisplitting method with $n = 16384$, $\beta = 11$, $p = m = 128$, $\alpha = \alpha_i = 0$

ovl	*iterations*	*time*	*% iterations*	*% time*
0	306	699.47 s	100.00	100.00
1	238	551.20 s	77.78	78.80
2	195	457.38 s	63.73	65.39
3	166	395.13 s	54.25	56.49
4	145	348.84 s	47.39	49.87
5	129	314.36 s	42.16	44.94
7	106	264.90 s	34.64	37.87
9	90	231.19 s	29.41	33.05
12	74	197.37 s	24.18	28.22
15	63	174.23 s	20.59	24.91
18	55	157.41 s	17.97	22.50
21	49	145.21 s	16.01	20.76
25	43	133.57 s	14.05	19.10
30	37	121.46 s	12.09	17.36
40	30	110.79 s	9.80	15.84
50	25	101.18 s	8.17	14.47
60	22	98.22 s	7.19	14.04
70	19	93.15 s	6.21	13.32
80	17	92.60 s	5.56	13.24
90	16	93.40 s	5.23	13.35
100	15	95.17 s	4.90	13.59
110	14	95.64 s	4.58	13.67
115	13	92.54 s	4.25	13.23
119	13	95.29 s	4.25	13.62
124	13	98.75 s	4.25	14.12
125	14	106.71 s	4.58	15.26
126	14	107.50 s	4.58	15.37
127	16	123.60 s	5.23	17.67
128	18	139.80 s	5.88	19.99

and $N(\alpha) := M(\alpha) - A$. By Definition 4, M_ℓ has the block form in Fig. 8 with a lower triangular matrix L_ℓ which we devide into blocks according to Fig. 9. The stars in Fig. 8 indicate the corresponding diagonal elements of A.

Note that L_1 starts with M_{22} and L_p ends with $M_{2p,2p}$ since $T_0 = T_{p+1} = \emptyset$ by definition. It is obvious that M_{ii}, $i = 2, 3, \ldots, 2p$, are quadratic lower triangular matrices. One shows as in a) that L_ℓ and M_{ii} are M–matrices. In particular, they are invertible, and (24) implies

$$M_{2\ell+1,2\ell-1} = 0\,, \quad \ell = 2, \ldots, p-1\,. \tag{26}$$

The special structure of M_ℓ yields for $M(\alpha)^{-1}$ the block form in Fig. 10. We wrote brackets in $M_{ij}^{(-1)}$, $i \neq j$, in order to indicate that this block is a part of some inverse L_ℓ^{-1}, while M_{ii}^{-1} (without such brackets) denotes an inverse in the usual sense.

$$L_\ell = \begin{pmatrix} M_{2\ell-1,2\ell-1} & 0 & 0 \\ M_{2\ell,\,2\ell-1} & M_{2\ell,\,2\ell} & 0 \\ M_{2\ell+1,\,2\ell-1} & M_{2\ell+1,\,2\ell} & M_{2\ell+1,2\ell+1} \end{pmatrix} \begin{matrix} \} S_\ell \cap T_{\ell-1} \\ \} S_\ell \setminus T_{\ell-1} \\ \} S_{\ell+1} \cap T_\ell \end{matrix}$$

$$\underbrace{\qquad}_{S_\ell \cap T_{\ell-1}} \quad \underbrace{\qquad}_{S_\ell \setminus T_{\ell-1}} \quad \underbrace{\qquad}_{S_{\ell+1} \cap T_\ell}$$

Fig. 9.

Let $e := (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^q$, $M^{-1} := M^{-1}(e)$, and $D(\alpha) := \operatorname{diag}(I_2, \alpha_1 I_3, \alpha_1 I_4, \alpha_1\alpha_2 I_5, \alpha_1\alpha_2 I_6, \ldots, \prod_{i=1}^{p-1} \alpha_i \cdot I_{2p-1}, \prod_{i=1}^{p-1} \alpha_i \cdot I_{2p})$ with I_i being the identity matrix of the same size as M_{ii}. Then $M(\alpha)^{-1} = D(\alpha) M^{-1} D(\alpha)^{-1}$, hence $M(\alpha) = D(\alpha) M\, D(\alpha)^{-1}$, i.e., the parameters α_i occur in $M(\alpha)$ at the same places as in $M(\alpha)^{-1}$, and possibly at those blocks of M which correspond to zero blocks of M^{-1}. We now show that M has the block form in Fig. 11 with

$$G_{2\ell,2\ell-2} = M_{2\ell,2\ell-1} M^{-1}_{2\ell-1,2\ell-1} M_{2\ell-1,2\ell-2} \geq 0 \,. \tag{28}$$

Dividing A into the same blocks as M, the non–negativity of $G_{2\ell,2\ell-2}$ follows from $M_{2\ell,2\ell-1} = A_{2\ell,2\ell-1} \leq 0$, $M_{2\ell-1,2\ell-2} = A_{2\ell-1,2\ell-2} \leq 0$ and $M^{-1}_{2\ell-1,2\ell-1} \geq 0$, since A and $M_{2\ell-1,2\ell-1}$ are M–matrices. We will now show that

$$M^{-1} M = I \,. \tag{29}$$

Because of the relation between L_ℓ and the framed blocks of $M(\alpha)^{-1}$ (with $\alpha = e$), and because of the zero structure of M^{-1} and M, it only remains to prove that the elements $(M^{-1}M)_{2\ell,2\ell-2}$, $\ell = 2, \ldots, p$, and $(M^{-1}M)_{2\ell+1,2\ell-2}$, $\ell = 2, \ldots, p-1$, are zero where the indices are block indices which correspond to those of M and M^{-1}. (In particular, this means that $M^{-1}M$ begins with $(M^{-1}M)_{22}$ in the left uppermost corner, and $(M^{-1}M)_{42}$ is the block scalar product of the third (!) block row of M^{-1} and the first (!) block column of M.) With (27) and with the known relations $(MM^{-1})_{2\ell,2\ell-1} = 0$, $(M^{-1}M)_{2\ell+1,2\ell-1} = 0$ we obtain

$$\begin{aligned} (M^{-1}M)_{2\ell,2\ell-2} &= M^{(-1)}_{2\ell,2\ell-1} M_{2\ell-1,2\ell-2} + M^{-1}_{2\ell,2\ell} G_{2\ell,2\ell-2} \\ &= M^{-1}_{2\ell,2\ell} \left(M_{2\ell,2\ell} M^{(-1)}_{2\ell,2\ell-1} + M_{2\ell,2\ell-1} M^{-1}_{2\ell-1,2\ell-1} \right) M_{2\ell-1,2\ell-2} \\ &= M^{-1}_{2\ell,2\ell} (MM^{-1})_{2\ell,2\ell-1} M_{2\ell-1,2\ell-2} = 0 \,, \end{aligned}$$

(27)

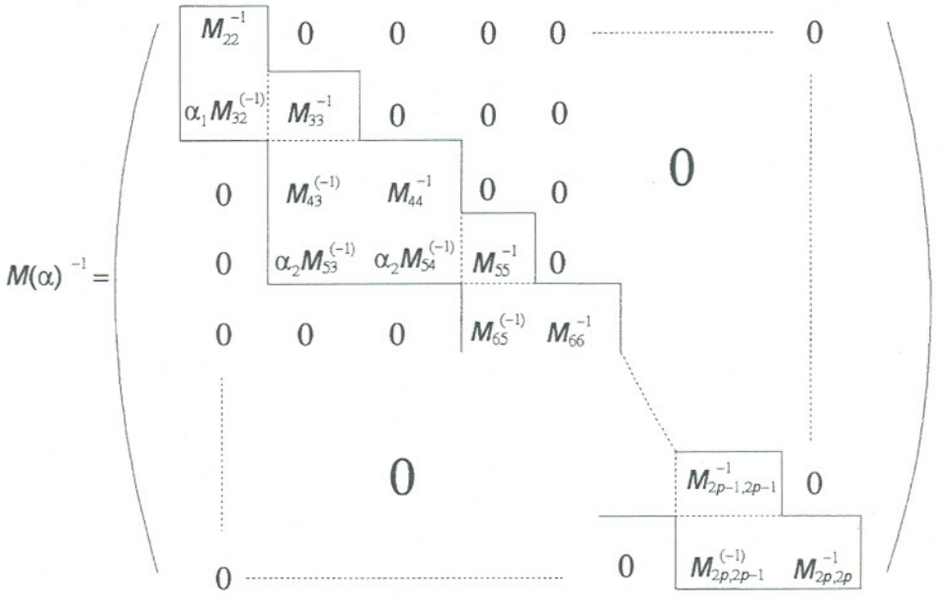

Fig. 10.

and

$$
\begin{aligned}
(M^{-1}M)_{2\ell+1,2\ell-2} &= M^{(-1)}_{2\ell+1,2\ell-1} M_{2\ell-1,2\ell-2} + M^{(-1)}_{2\ell+1,2\ell} G_{2\ell,2\ell-2} \\
&= \left(M^{(-1)}_{2\ell+1,2\ell-1} M_{2\ell-1,2\ell-1} + M^{(-1)}_{2\ell+1,2\ell} M_{2\ell,2\ell-1} \right) M^{-1}_{2\ell-1,2\ell-1} M_{2\ell-1,2\ell-2} \\
&= (M^{-1}M)_{2\ell+1,2\ell-1} M^{-1}_{2\ell-1,2\ell-1} M_{2\ell-1,2\ell-2} = 0 \ .
\end{aligned}
$$

Thus M has the block form above. Since $G_{2\ell,2\ell-2} \geq 0$, the same holds for the block $N_{2\ell,2\ell-2}$ of $N := M - A$ which reads

$$
N_{2\ell,2\ell-2} = G_{2\ell,2\ell-2} \underbrace{-A_{2\ell,2\ell-2}}_{\leq 0} \ .
$$

Hence $N \geq 0$, and $N(\alpha) := M(\alpha) - A = D(\alpha) M D(\alpha)^{-1} - M + N \geq 0$ if $\alpha_i \in [0,1]$ for each i.

c) Since (27) shows that $0 \leq \alpha \leq \hat{\alpha} \leq e$ implies $0 \leq M(\alpha)^{-1} \leq M(\hat{\alpha})^{-1}$, the assertion follows from b) and Theorem 2a). □

Analogously to Corollary 2 we get at once our next result.

Corollary 3. *With the notations of Theorem 8 let the following conditions hold.*

$$
\left.
\begin{array}{l}
\min S_\ell = \min T_\ell \\
|S_\ell| = m = \text{constant}
\end{array}
\right\} \ \ell = 1, \ldots, p \ ,
$$

$$
|S_\ell \cap T_{\ell-1}| = ovl = \text{constant} \ , \quad \ell = 2, \ldots, p \ .
$$

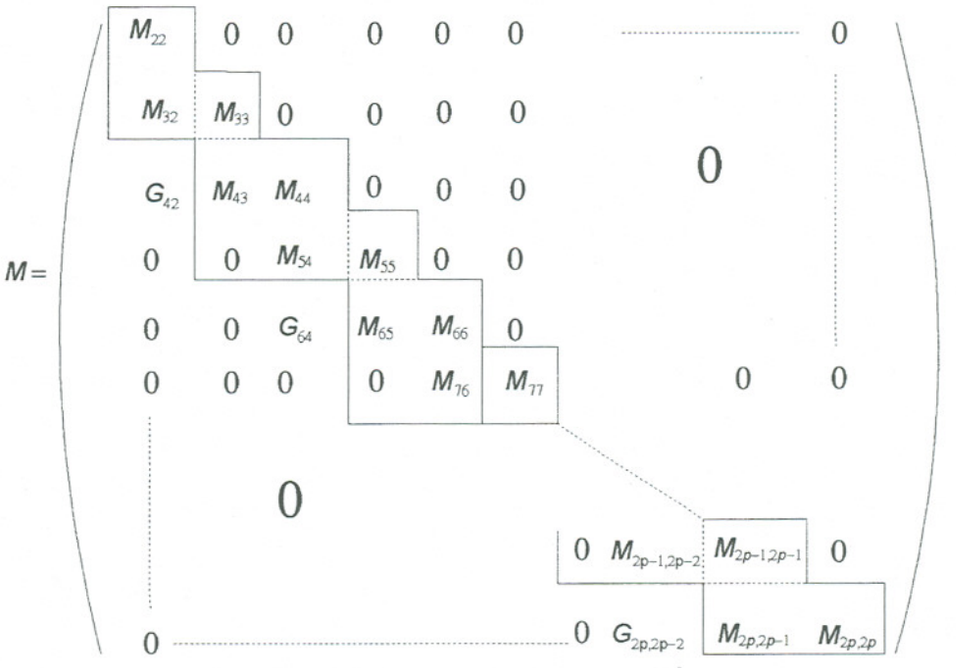

Fig. 11.

If

$$\beta \leq m - ovl$$

then the assertions of Theorem 8 hold.

We will illustrate Corollary 3 by an example using the matrix of Example 1 and choosing $\alpha_i = \alpha$ for all α_i in E_ℓ. Although the results in this corollary and in Theorem 8 refer to values $\alpha \in [0,1]$ we will perform $\rho(H(\alpha))$ for $\alpha \in [-1,7]$. We will see that $\rho(H(\alpha))$ has a relative minimum in $[-1,7]\backslash[0,1]$.

Example 3. Let A and b be defined as in (20), (21) with $n = 256$, $\beta = 5$, $p = 16$, $m = 16$, $\alpha_i = \alpha$. We started the Gauß–Seidel like multisplitting method with $x^0 = 0$ and stopped it when $\|x^k - e\|_\infty \leq 10^{-5}$ was fulfilled. Under these conditions we varied the weighting parameter α from -1 to 7 for $ovl = 4$ and $ovl = 12$. In Fig. 12 we plotted $\rho(H(\alpha))$ versus α. The figure shows that the speed of convergence can be improved by choosing α outside from the interval $[0,1]$.

Our next example shows that for the *overlapping block Jacobi multisplitting method* we cannot expect that the splitting $A = M - N$ from Theorem 3 is necessarily regular *even if the conditions* (16) *and* (17) *of Theorem* 7 *hold* – in contrast to the overlapping block Gauß–Seidel like multisplitting method. Therefore, Theorem 2 seems not to be applicable here.

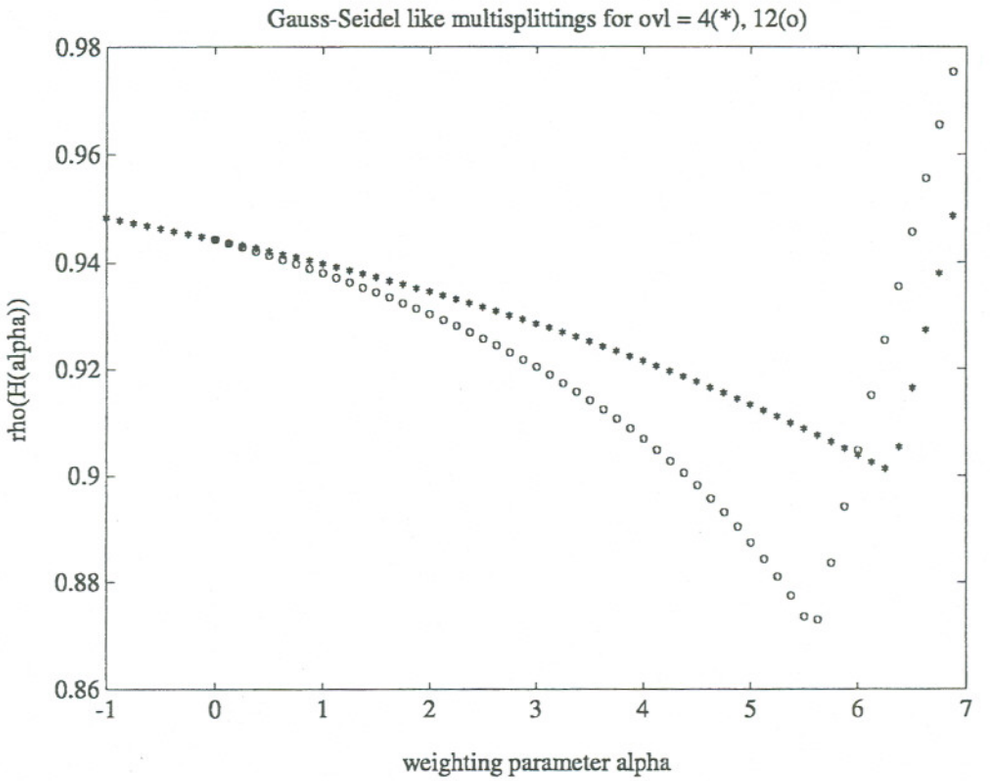

Fig. 12. Influence of the weighting parameter $\alpha = \alpha_i$ on the spectral radius $\rho(H(\alpha))$ for the overlapping Gauß–Seidel like multisplitting method with $n = 256$, $\beta = 5$, $p = m = 16$, $ovl \in \{4, 12\}$

Example 4. Let

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix},$$

hence $\beta = 1$. Choose $S_1 := \{1,2\}$, $S_2 := \{3,4\}$, $T_1 := \{1,2,3\}$, $T_2 := S_2$, $T_0 := T_3 := \emptyset$. Define M_ℓ, N_ℓ as in Definition 3b), and let $E_1 := \operatorname{diag}(1,1,0,0)$, $E_2 := \operatorname{diag}(0,0,1,1)$. Then (16) and (17) hold, and

$$M^{-1} = \begin{pmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix} \tag{30}$$

with M^{-1} being obtained as in (10). Since (30) implies

$$M = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & \frac{3}{2} & -1 & \frac{1}{2} \\ 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix},$$

we get

$$N = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore, $A = M - N$ is only a weak regular splitting, and not a regular one. Nevertheless, one finds the same monotonous behaviour as in Example 2. Choosing $T_1^{(1)} = \{1,2\}$, $T_1^{(2)} = \{1,2,3\}$ and $T_1^{(3)} = \{1,2,3,4\}$, respectively, yields $\rho(H^{(1)}) = \frac{2}{3} = 0.666\ldots$, $\rho(H^{(2)}) = \frac{1}{\sqrt{6}} = 0.408\ldots$ and $\rho(H^{(3)}) = 0$, respectively. This can be seen from a simple computation with the matrices E_ℓ and with the remaining index sets as above.

In our final examples we present further experimental results on the optimality of $\rho(H(\alpha))$ with varying parameter $\alpha \in \mathbb{R}$.

Example 5. Let

$$A = \begin{pmatrix} B & -I & & & O \\ -I & B & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & B & -I \\ O & & & -I & B \end{pmatrix} \in \mathbb{R}^{n^2 \times n^2} \tag{31}$$

with $B := \begin{pmatrix} 4 & -1 & & & O \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1 \\ O & & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{n \times n}$, and with I denoting the $n \times n$ identity matrix. As is well–known A is the matrix resulting from the five–point discretization of the Laplace equation

$$\Delta u = 0 \tag{32}$$

on the unit square $\Omega = (0,1) \times (0,1)$ where the values of u are prescribed at the boundary $\partial\Omega$ and where an equidistant grid of mesh size $h = \frac{1}{n+1}$ is used. We chose $n = 64$ whence A has $n^2 = 4\,096$ rows and columns, respectively. The boundary condition for (32) was adjusted such that the righthand side b of the resulting linear system $Ax = b$ has the solution $x = e$.

Table 3. Influence of the weighting parameter $\alpha = \alpha_i$ on the maximal number of iterations for the overlapping Gauß–Seidel like multisplitting method applied to the discretized Laplace equation with $A \in \mathbb{R}^{4096\times 4096}$, $\beta = 64$, $p = 32$, $m = 128$, $ovl = 64$

α	*iterations*	*time*	*% iterations*	*% time*
0.0	3644	839.05 s	100.00	131.64
0.5	3345	766.74 s	91.79	120.29
1.0	3108	715.65 s	85.29	112.27
2.0	2747	632.59 s	75.38	99.25
3.0	2349	540.95 s	64.46	84.87
4.0	1913	440.54 s	52.50	69.12
5.0	1437	330.92 s	39.43	51.92
6.0	906	208.63 s	24.86	32.73
6.5	599	137.92 s	16.44	21.64
6.625	511	117.03 s	14.02	18.36
6.75	408	93.46 s	11.20	14.66
6.78125	376	86.82 s	10.32	13.62
6.8125	335	76.73 s	9.19	12.04
6.828125	310	71.72 s	8.51	11.25
6.84375	291	66.64 s	7.99	10.46
6.860275	324	74.21 s	8.89	11.64
6.875	331	75.81 s	9.08	11.89
6.9375	344	78.80 s	9.44	12.36
7.0	374	85.66 s	10.26	13.44
7.125	473	108.37 s	12.98	17.00
7.25	698	160.72 s	19.15	25.22
7.375	1146	264.77 s	31.45	41.54

We tested the Gauß–Seidel like multisplitting method for $p = 32$ and $ovl = 64$. To this end we started it with $x^0 = 0$ and stopped it when $\|x^k - e\| \leq 10^{-5}$ was fulfilled, varying α from 0 to 7.375. The results in Table 3 show that the smallest number of iterates is attained for $\alpha = \alpha_{opt} = 6.84375$. For $\alpha = 0$ we needed 3 644 iterations. The number of iterations must be the same for $\alpha = 0$ and any value of ovl since the components in $T_\ell \backslash S_\ell$ which are computed by the ℓ–th processor are not needed for x^{k+1} because of the particular form of $E(0)$. Since for $ovl = 0$ and $\alpha = 0$ no superfluous component is computed we needed here only 637.39 seconds. This is an essentially smaller time as in the case $ovl = 64$, $\alpha = 0$. We used this smaller time as basis for the time percentage in the last column of Table 3.

We conclude our paper with another example in which we modified the partial differential equation to be discretized.

Example 6. We consider the boundary value problem

$$\begin{cases} xu_{xx} + yu_{yy} = f, & (x,y) \in \Omega := (0,1) \times (0,1) \\ u = g, & (x,y) \in \partial\Omega \end{cases} \tag{33}$$

Discretizing (33) as in Example 5 yields a linear system

(34) $$Ax = b$$

in which the matrix A has the same zero pattern as in (31). We chose f and g such that $x = e$ solves (34). For n, p, ovl, and x^0 as in Example 5 and for $\alpha \in [0\,,\,4.05]$ we found again an optimal α outside $[0, 1]$. For $ovl = 0$ the same remark holds as in the foregoing example. We needed 984.71 seconds for the 5 639 iterations of the case $\alpha = 0$, $ovl = 0$ and used this time again as basis for the last column of Table 4.

Table 4. Influence of the weighting parameter $\alpha = \alpha_i$ on the maximal number of iterations for the overlapping Gauß–Seidel like multisplitting method applied to the discretized PDE $xu_{xx} + yu_{yy} = f$ on the unit square with $A \in \mathbb{R}^{4096\times 4096}$, $\beta = 64$, $p = 32$, $m = 128$, $ovl = 64$

α	*iterations*	*time*	*% iterations*	*% time*
0.0	5639	1296.27 s	100.00	131.64
0.5	5203	1196.04 s	92.27	121.46
1.0	4839	1111.74 s	85.81	112.90
2.0	4231	969.22 s	75.03	98.43
3.0	3593	807.95 s	63.72	82.05
3.5	3265	767.54 s	57.90	77.95
3.75	3101	696.95 s	54.99	70.78
3.875	3018	691.63 s	53.52	70.24
3.9375	2977	685.57 s	52.79	69.62
3.96875	2956	679.51 s	52.42	69.01
4.0	2936	673.55 s	52.07	68.40
4.015625	2925	670.77 s	51.87	68.12
4.0234375	2920	657.44 s	51.78	66.76
4.03125	2915	656.09 s	51.69	66.63
4.0390625	3714	850.77 s	65.86	86.40
4.046875	5448	1226.54 s	96.61	124.56

Acknowledgement. We thank A. Frommer, Wuppertal, for sending us a preprint of [18] at a very early stage. We also want to express our gratitude to the University of Paderborn for giving us access to its Parsytec parallel computer systems. Finally we acknowledge the comments and remarks of two referees which helped to improve this paper.

References

1. Adams, L., Ong, E. (1988/89): Additive Polyomial Preconditioners for Parallel Computing. Parallel Comput. **9**, 333–345
2. Berman, A., Plemmons R.J.: Nonnegative Matrices in the Mathematical Sciences. Academic Press, New York, 1979
3. Bru, R., Corral, C., Martinez, A., Mas, J. (1995): Multisplitting Preconditioners Based on Incomplete Choleski Factorizations. SIAM J. Matrix Anal. Appl. **16**, 1210–1222
4. Bru, R., Elsner, L., Neumann, M. (1988): Models of Parallel Chaotic Iteration Methods. Linear Algebra Appl. **103**, 175–192
5. Burrage, K., Pohl, B. (1994): Implementing an ODE–Code on Distributed Memory Computers. Computers Math. Applic. **28**, 235–252
6. Elsner, L. (1989): Comparisons of Weak Regular Splittings and Multisplitting Methods. Numer. Math. **56**, 283–289
7. Fan, Ky (1958): Topological Proof for Certain Theorems on Matrices with Non–negative Elements. Monatsh. Math. **62**, 219–237

8. Frommer, A. (1989): Parallel Nonlinear Multisplitting Methods. Numer. Math. **56** 269–282
9. Frommer, A. (1989): Lokale Kovergenzaussagen bei nichtlinearen Mehrfachzerlegungen. Z. angew. Math. Mech. **69**, T101–T102
10. Frommer, A.: Lösung linearer Gleichungssysteme auf Parallelrechnern. Vieweg, Braunschweig, 1990
11. Frommer, A., Mayer, G. (1989): Convergence of Relaxed Parallel Multisplitting Methods. Linear Algebra Appl. **119**, 141–152
12. Frommer, A., Mayer, G. (1989): Parallel Interval Multisplittings. Numer. Math. **56**, 255–267
13. Frommer, A., Mayer, G. (1989): Safe Bounds for the Solutions of Nonlinear Problems Using a Parallel Multisplitting Method. Computing **42**, 171–186
14. Frommer, A., Mayer, G. (1989): Zur Lösungseinschließung bei linearen Gleichungssystemen auf einem Parallelrechner. Z. angew. Math. Mech. **69**, T102–T103
15. Frommer, A., Mayer, G. (1990): Theoretische und praktische Ergebnisse zu Multisplitting–Verfahren auf Parallelrechnern. Z. angew. Math. Mech. **70**, T600 – T602
16. Frommer, A., Mayer, G. (1991): Iterationsverfahren für lineare Gleichungssysteme auf Parallelrechnern. Z. angew. Math. Mech. **71**, T799–T801
17. Frommer, A., Mayer, G. (1992): On the Theory and Practice of Multisplitting Methods in Parallel Computation. Computing **49**, 63–74
18. Frommer, A., Pohl, B. (1995): A Comparison Result for Multisplittings and Waveform Relaxation Methods. Numerical Linear Algebra with Applications **2**, 335–346. Also available as Research Report No. 93–05, Seminar für Angewandte Mathematik, ETH Zürich, May 1993
19. Huang, C.-M., O'Leary, D.P. (1993): A Krylov Multisplitting Algorithm for Solving Linear Systems of Equations. Linear Algebra Appl. **194**, 9–29
20. Jeltzsch, R., Pohl, B. (1995): Waveform Relaxation with Overlapping Splittings. SIAM J. Sci. Comp. **16**, 40–49
21. Jones, M.T., Szyld, D.B.: Two–stage Multisplitting Methods with Overlapping Blocks. Report 94–31, March 1994, Temple University, Department of Mathematics, Philadelphia
22. Kavanagh, J.P., Neumann, M. (1989): Consistency and Convergence of the Parallel Multisplitting Method for Singular M–Matrices. SIAM J. Matrix Anal. Appl. **10**, 210–218
23. Lelarasmee, E.: The Waveform Relaxation Method for Time Domain Analysis of Large Scale Nonlinear Dynamical Systems. Ph. D. Dissertation, University of California, Berkeley, CA, 1982
24. Mayer, G., Frommer, A.: A Multisplitting Method for Verification and Enclosure on a Parallel Computer. In Ullrich, C. (ed.): Contributions to Computer Arithmetic and Self–Validating Numerical Methods. IMACS Annals on Computing and Applied Mathematics, **7**. Baltzer, Basel, 1990, 483–497
25. Nabben, R. (1996): A Note on Comparison Theorems of Splittings and Multisplittings of Hermitian Positive Definite Matrices. Linear Algebra Appl. **233**, 67–80
26. Neumann, M., Plemmons, R.J. (1987): Convergence of Parallel Multisplitting Iterative Methods for M–matrices. Linear Algebra Appl. **88/89**, 559–573
27. O'Leary, D., White, R.E. (1985): Multi–splittings of Matrices and Parallel Solution of Linear Systems. SIAM J. Alg. Disc. Meth. **6**, 630–640
28. Ortega, J., Rheinboldt, W.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970
29. Papatheodorou, T., Saridakis, Y. (1989): Parallel Algorithms and Architectures for Multisplitting Iterative Methods. Parallel Comput. **12**, 171–182
30. Pohl, B. (1992): Ein Algorithmus zur Lösung von Anfangswertproblemen auf Parallelrechnern. Informatik–Dissertationen ETH Zürich **37**
31. Pohl, B. (1993): On the Convergence of the Discretized Multisplitting Waveform Relaxation Algorithm. Applied Numerical Mathematics **11**, 251–258
32. Szyld, D.B.: Synchronous and Asynchronous Two–stage Multisplitting Methods. In: Lewis, J.G. (ed.): Applied Linear Algebra. Proceedings of the Fifth SIAM Conference held in Snowbird, UT, USA, June 15–18, 1994. SIAM, Philadelphia, PA, 1994, 39–44
33. Wang, D. (1991): On the Convergence of the Parallel Multispitting AOR Algorithm, Linear Algebra Appl. **154–156**, 473–486
34. White, J., Sangiovanni–Vincentelli, A.: Relaxation Techniques for the Simulation on VLSI Circuits. Kluwer Academic Publishers, Boston, MA, 1987

35. White, R.E. (1986): A Nonlinear Parallel Algorithm with Application to the Stefan Problem. SIAM J. Numer. Anal. **23**, 639–652
36. White, R.E. (1986): Parallel Algorithms for Nonlinear Problems. SIAM J. Alg. Discr. Meth. **7**, 137–149
37. White, R.E. (1987): Multisplittings and Parallel Iterative Methods. Computer Meth. Appl. Mech. Engineering **64**, 567–577
38. White, R.E. (1989): Multisplitting with Different Weighting Schemes. SIAM J. Matrix Anal. Appl. **10**, 481–493
39. White, R.E. (1990): Multisplittings of a Symmetric Positive Definite Matrix. SIAM J. Matrix Anal. Appl. **11**, 69–82

This article was processed by the author using the LaTeX style file *pljour1m* from Springer-Verlag.