CONVERGENCE OF THE MULTIPLICATIVE SCHWARZ METHOD FOR SINGULARLY PERTURBED CONVECTION-DIFFUSION PROBLEMS DISCRETIZED ON A SHISHKIN MESH[*]

CARLOS ECHEVERRÍA[§], JÖRG LIESEN[§], DANIEL B. SZYLD[‡], AND PETR TICHÝ[†]

Abstract. We analyze the convergence of the multiplicative Schwarz method applied to nonsymmetric linear algebraic systems obtained from discretizations of one-dimensional singularly perturbed convection-diffusion equations by upwind and central finite differences on a Shishkin mesh. Using the algebraic structure of the Schwarz iteration matrices we derive bounds on the infinity norm of the error that are valid from the first step of the iteration. Our bounds for the upwind scheme prove rapid convergence of the multiplicative Schwarz method for all relevant choices of parameters in the problem. The analysis for the central difference is more complicated, since the submatrices that occur are nonsymmetric and sometimes even fail to be M-matrices. Our bounds still prove the convergence of the method for certain parameter choices.

Key words. singularly perturbed problems, Shishkin mesh discretization, multiplicative Schwarz method, iterative methods, convergence analysis

AMS subject classifications. 15A60, 65F10, 65F35

1. Introduction. It is well known that, due to the presence of boundary layers in the analytic solution, singularly perturbed convection-diffusion problems require special discretization techniques in order to guarantee the stability of a numerical method. A general overview can be found, e.g., in the survey article [22]. One widely accepted discretization technique in this context is given by upwind or central finite difference schemes on a Shishkin mesh as described, e.g., in [22, Section 5] or [6, 12, 14, 19]. In short, Shishkin meshes are formed by an overlapping union of piecewise uniform meshes, with their respective sizes and mesh transition (or interface) points adapted to the expected width of the layers in the solution.

The matrix in a linear algebraic system obtained from a Shishkin mesh discretization of a singularly perturbed convection-diffusion equation is nonsymmetric and often highly nonnormal and ill-conditioned. Standard iterative solvers like the (unpreconditioned) GMRES method converge very slowly when applied to such a system; see Figures 2.2–2.4 in this paper for examples. On the other hand, the Shishkin mesh discretization naturally leads to a *decomposition of the domain*, which suggests to solve the discretized problem by the multiplicative Schwarz method. This is the approach we explore in this paper for one-dimensional model problems.

We consider both upwind and central difference schemes on the Shishkin mesh. Using the algebraic structure of the iteration matrices, we derive bounds on the infinity norm of the error that are valid from the first step of the corresponding multiplicative Schwarz iteration. Thus, unlike asymptotic convergence results based on bounding the spectral radius of the iteration matrix, our results apply to the transient rather than the asymptotic behavior. For the upwind scheme we prove rapid convergence of the multiplicative Schwarz iteration for all relevant

---

[§]Institute of Mathematics, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany
({echeverria,liesen}@math.tu-berlin.de).
[‡]Department of Mathematics, Temple University, Philadelphia, PA 19122–6094, USA
(szyld@temple.edu).
[†]Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, and Institute of Computer Science, Academy of Science of the Czech Republic (ptichy@karlin.mff.cuni.cz).

parameters in the problem. The analysis of the central difference scheme is more complicated, since some of the submatrices that occur in this case are not only nonsymmetric but also fail to be M-matrices. This reminds of the analysis in [1], which showed that in this case the difference scheme itself does not satisfy a discrete maximum principle. Nevertheless, we can prove the convergence of the multiplicative Schwarz method for problems discretized by central differences on a Shishkin mesh under the assumption that the number of discretization points in each of the local subdomains is even. If this assumption is not satisfied, then the method may diverge, which we also explain in our analysis.

The multiplicative Schwarz method as a solution technique, which we analyze here and which is often called the alternating Schwarz method (see [8] for a historical survey), is for the most part used as a preconditioner for Krylov subspace methods such as CG or GMRES. Many of the convergence results for multiplicative Schwarz have been presented in that context; see, e.g., the treatment in the books [5, 24] and many references therein. When the method is considered from an algebraic point of view, as we do here, it is commonly treated as a solution method; see, e.g., [2]. One of our goals in this paper is to bring an understanding on why this solution technique is so effective for solving problems arising from the Shishkin mesh discretizations. Moreover, the convergence bounds provided in this paper shed light on an apparent contradiction: if the continuous problem becomes more difficult (a smaller diffusion coefficient is chosen), then the convergence of the multiplicative Schwarz method for the discretized problem becomes faster.

Several authors have previously applied the alternating (or multiplicative) Schwarz method to the continuous problem based on the partitioning of the domain into overlapping subdomains and subsequently discretized by introducing uniform meshes on each subdomain; see, e.g., [6, 7, 15, 16, 17, 18, 19]. However, as clearly explained in [18], significant numerical problems including very slow convergence and accumulation of errors (up to the point of non-convergence of the numerical solution) can occur when layer-resolving mesh transition points are used in this setup. These problems are avoided in our approach since we first discretize and then apply the multiplicative Schwarz method to the linear algebraic system. To the best of our knowledge, this approach has not been studied in the literature so far.

The paper is organized as follows. Section 2 specifies the model problem and its Shishkin mesh discretization. In Section 3 we describe the multiplicative Schwarz method for the discretized problem. In Section 4 we present the convergence analysis of the multiplicative Schwarz method, first for the upwind scheme and then for the central difference scheme. Numerical examples are shown in Section 5, and a concluding discussion, which also puts our results into a broader context, is given in Section 6.

2. Model problem and its Shishkin mesh discretization. We consider a one-dimensional convection-diffusion model problem with constant coefficients and Dirichlet boundary conditions

$$(2.1) \qquad -\epsilon u''(x) + \alpha u'(x) + \beta u(x) = f(x) \text{ in } (0, 1), \quad u(0) = u_0, \ u(1) = u_1,$$

where $\alpha \gg \epsilon > 0$ and $\beta \geq 0$. We assume that the parameters of the problem, i.e., ϵ, α, β, $f(x), u_0, u_1$, are chosen so that the solution $u(x)$ has one boundary layer at $x = 1$.

A common approach for discretizing such a problem is to resolve the boundary layer using a finer mesh close to $x = 1$. Here we will focus on the Shishkin mesh discretization. This technique has been described in detail, for example, in the book [19] or in the survey article [22, Section 5]. We therefore only state the facts that are relevant for our analysis.
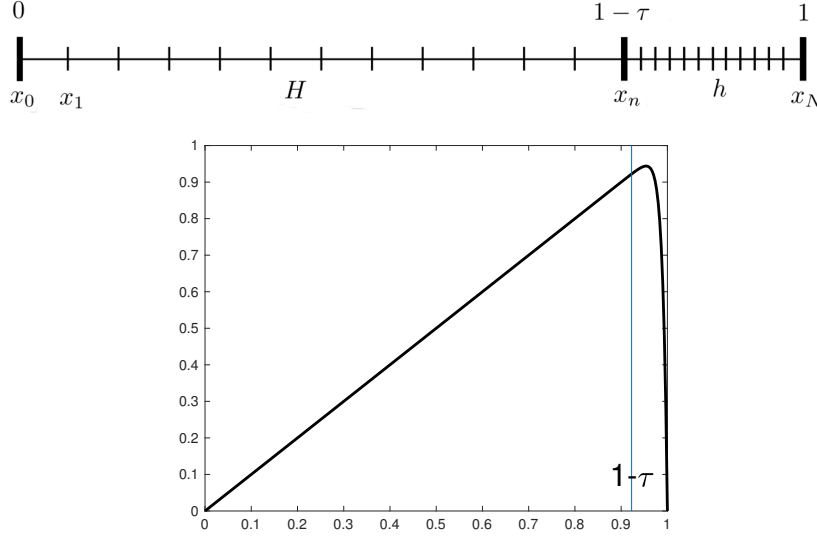
Suppose that an even integer $N \geq 4$ defining the number of intervals constituting the Shishkin mesh is given and suppose that

$$(2.2) \qquad \tau \equiv \frac{2}{\alpha} \epsilon \ln N \leq \frac{1}{2}.$$

The inequality in (2.2) means that

$$(2.3) \qquad \epsilon \leq \frac{\alpha}{4 \ln N},$$

which is a natural assumption since $\alpha \gg \epsilon$ and the number of mesh points usually is not exponentially large relative to ϵ. The *mesh transition point* $1 - \tau$ then will be close to $x = 1$ and the boundary layer will be contained in the (small) interval $[1 - \tau, 1]$.

The idea of the Shishkin mesh discretization of the interval $[0, 1]$ is to use the same number of equidistantly distributed mesh points in each of the subintervals $[0, 1 - \tau]$ and $[1 - \tau, 1]$. Thus, if we denote

$$n \equiv \frac{N}{2}, \quad H \equiv \frac{1 - \tau}{n}, \quad h \equiv \frac{\tau}{n},$$

then the $N + 1$ mesh points of the Shishkin mesh are given by

$$x_i \equiv iH, \quad i = 0, \ldots, n, \qquad x_i \equiv 1 - (N - i)h, \quad i = n + 1, \ldots, N.$$

Here $x_0 = 0$ and $x_N = 1$, so that the mesh consists of $N - 1$ interior mesh points, where the mesh point x_n is exactly at the transition point $1 - \tau$. The ratio between the mesh sizes in the two subdomains is

$$\frac{h}{H} = \frac{\tau}{1 - \tau} = \tau + \mathcal{O}(\tau^2),$$

which is usually much less than 1.

An illustration of a Shishkin mesh and a plot of the (explicitly known) analytic solution of the problem (2.1) with $\epsilon = 0.01$, $\alpha = 1$, $\beta = 0$, $f(x) \equiv 1$, and $u_0 = u_1 = 0$ are shown in Figure 2.1; cf. [22, Example 3.1]. Choosing, for example, $N = 48$ gives the mesh transition point $1 - \tau = 0.9226$.

We consider two different finite difference schemes on the Shishkin mesh: upwind and central differences. Using the standard difference operators (see, e.g., [22, Section 4]), both schemes yield a linear algebraic system $Au^N = f^N$ with the tridiagonal and nonsymmetric matrix

$$(2.4) \qquad A = \begin{bmatrix} a_H & b_H & & & & & & \\ c_H & \ddots & \ddots & & & & & \\ & \ddots & \ddots & b_H & & & & \\ & & c_H & a_H & b_H & & & \\ \hline & & & c & a & b & & \\ \hline & & & & c_h & a_h & b_h & \\ & & & & & c_h & \ddots & \ddots \\ & & & & & & \ddots & \ddots & b_h \\ & & & & & & & c_h & a_h \end{bmatrix} \in \mathbb{R}^{(N-1) \times (N-1)}.$$

FIG. 2.1. *Illustration of a Shishkin mesh (top) and plot of the analytic solution of the problem* (2.1) *with* $\epsilon = 0.01$, $\alpha = 1$, $\beta = 0$, $f(x) \equiv 1$, *and* $u_0 = u_1 = 0$ *(bottom). For* $N = 48$ *the mesh transition point is* $1 - \tau = 0.9226$.

For the upwind scheme, the entries of A are given by

$$
(2.5) \quad
\begin{aligned}
c_H &= -\frac{\epsilon}{H^2} - \frac{\alpha}{H}, & a_H &= \frac{2\epsilon}{H^2} + \frac{\alpha}{H} + \beta, & b_H &= -\frac{\epsilon}{H^2}, \\
c &= -\frac{2\epsilon}{H(H+h)} - \frac{\alpha}{H}, & a &= \frac{2\epsilon}{hH} + \frac{\alpha}{H} + \beta, & b &= -\frac{2\epsilon}{h(H+h)}, \\
c_h &= -\frac{\epsilon}{h^2} - \frac{\alpha}{h}, & a_h &= \frac{2\epsilon}{h^2} + \frac{\alpha}{h} + \beta, & b_h &= -\frac{\epsilon}{h^2},
\end{aligned}
$$

and for the central difference scheme by

$$
(2.6) \quad
\begin{aligned}
c_H &= -\frac{\epsilon}{H^2} - \frac{\alpha}{2H}, & a_H &= \frac{2\epsilon}{H^2} + \beta, & b_H &= -\frac{\epsilon}{H^2} + \frac{\alpha}{2H}, \\
c &= -\frac{2\epsilon}{H(H+h)} - \frac{\alpha}{h+H}, & a &= \frac{2\epsilon}{hH} + \beta, & b &= -\frac{2\epsilon}{h(H+h)} + \frac{\alpha}{h+H}, \\
c_h &= -\frac{\epsilon}{h^2} - \frac{\alpha}{2h}, & a_h &= \frac{2\epsilon}{h^2} + \beta, & b_h &= -\frac{\epsilon}{h^2} + \frac{\alpha}{2h}.
\end{aligned}
$$

If $u^N = A^{-1} f^N = [u_1^N, \ldots, u_{N-1}^N]^T$ is the exact algebraic solution and $u(x)$ is the solution of (2.1), then there exist constants $c_1, c_2 > 0$ such that

$$
\max_{1 \leq i \leq N-1} |u(x_i) - u_i^N| \leq c_1 \frac{\ln N}{N}
$$

for the upwind scheme and

$$
\max_{1 \leq i \leq N-1} |u(x_i) - u_i^N| \leq c_2 \left(\frac{\ln N}{N} \right)^2
$$

for the central difference scheme. Thus, the convergence of both schemes is ϵ-uniform, and the central difference scheme is more accurate than the upwind scheme. As pointed out
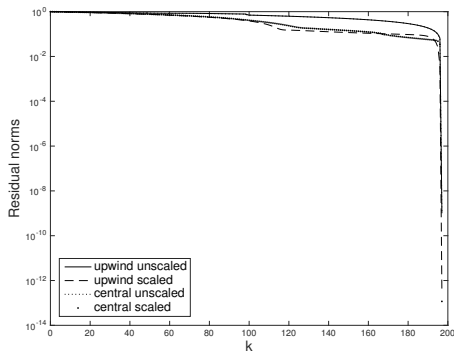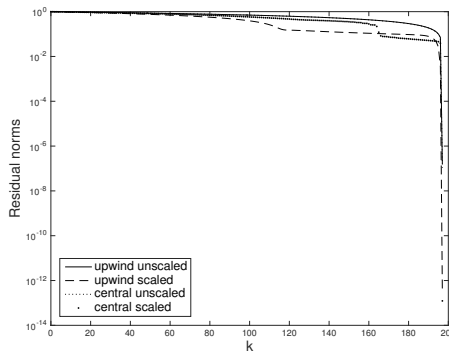
FIG. 2.2. *GMRES convergence for $\epsilon = 10^{-4}$.*

FIG. 2.3. *GMRES convergence for $\epsilon = 10^{-6}$.*

by Stynes [22, p. 470], the convergence proof for the central differences (originally due to Andreyev and Kopteva [1]) is complicated since the scheme does not satisfy a discrete maximum principle. We meet similar complications in our analysis in Section 4.2 below.

Both schemes lead to highly ill-conditioned matrices A. The main reason is the large difference between the mesh sizes H and h, which implies large differences between the moduli of the nonzero entries of A corresponding to each subdomain. Thus, A is poorly scaled. As shown by Roos [21], a simple diagonal scaling reduces the order of the condition number for the matrix from the upwind scheme from $\mathcal{O}(\epsilon^{-1}(N/\ln N)^2)$ to $\mathcal{O}(N^2/\ln N)$. Although not shown by Roos, an analogous diagonal scaling appears to work well also for the central difference scheme; see Section 4.3 for the exact form of the scaling matrices. The first row in the following table provides a numerical illustration for $\epsilon = 10^{-8}$, $\alpha = 1$, $\beta = 0$ in (2.1), and $N = 198$.

	upwind	upwind scaled	central	central scaled
matrix cond.	4.0500×10^{10}	2.9569×10^3	6.2323×10^{10}	2.9514×10^3
eigenvector cond.	1.5143×10^{17}	1.2297×10^{19}	4.1070×10^3	1.8682×10^2

The second row of the table displays the condition numbers of the eigenvector matrices from the decomposition $A = VDV^{-1}$ computed by `[V,D]=eig(A)` in MATLAB. We observe that the upwind scheme yields matrices with very ill-conditioned eigenvectors, i.e., highly nonnormal matrices. Apparently, the eigenvector conditioning is not much affected by the diagonal scaling.

As mentioned in the introduction, linear algebraic systems resulting from discretizations of convection-dominated convection-diffusion problems represent a challenge for iterative solvers. Figures 2.2–2.4 illustrate that this also holds for the Shishkin mesh discretization of the model problem (2.1). These figures show the relative true residual norms of the (unpreconditioned) GMRES method with zero initial vector applied to $Au^N = f^N$ from the Shishkin mesh discretization of (2.1) with $\alpha = 1$, $\beta = 0$, $f(x) \equiv 1$, $u_0 = u_1 = 0$, $N = 198$, and different values of ϵ. The GMRES convergence is virtually the same for both discretizations (upwind and central differences). Neither the scaling nor the eigenvector conditioning appears to have a significant effect on the performance of the iterative solver.

3. The multiplicative Schwarz method. Any Shishkin mesh discretization naturally leads to a decomposition of the given domain into overlapping subdomains. In our model problem the domain is the interval $[0, 1]$, and the overlapping subdomains are the intervals
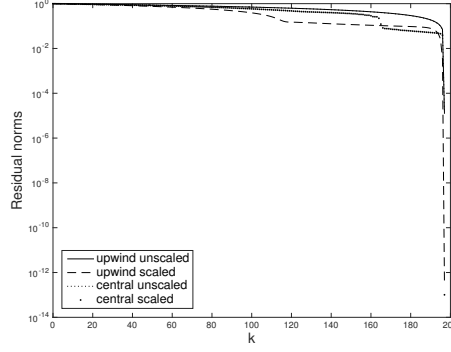
FIG. 2.4. *GMRES convergence for $\epsilon = 10^{-8}$.*

$[0, 1 - \tau + h]$ and $[1 - \tau - H, 1]$. The width of the overlap is $H + h = 2/N$, and the mesh transition point $x_n = 1 - \tau$ is the only mesh point in the overlap. We will now describe the multiplicative Schwarz method for solving the linear algebraic system $Au^N = f^N$.

In short, the method uses restriction operators for constructing a multiplicative iteration matrix in which each factor corresponds to a local solve in one of the subdomains. In the notation established above, the restriction operators can be written as

$$R_1 \equiv [I_n \quad 0] \quad \text{and} \quad R_2 \equiv [0 \quad I_n],$$

both of size $n \times (N - 1)$. The restrictions of the matrix A in (2.4) to the two subdomains are given by the two $n \times n$ matrices

$$A_1 \equiv R_1 A R_1^T \equiv \begin{bmatrix} A_H & b_H e_m \\ c e_m^T & a \end{bmatrix} \quad \text{and} \quad A_2 \equiv R_2 A R_2^T \equiv \begin{bmatrix} a & b e_1^T \\ c_h e_1 & A_h \end{bmatrix},$$

where $m \equiv n - 1$ and $e_1, e_m \in \mathbb{R}^m$. In the following, the unit basis vectors e_j are always considered to be of appropriate length, which for simplicity is sometimes not explicitly stated. Note that $A_H, A_h \in \mathbb{R}^{m \times m}$ are tridiagonal Toeplitz matrices. The matrices corresponding to the solves on the two domains then are given by

$$(3.1) \qquad\qquad P_i \equiv R_i^T A_i^{-1} R_i A, \quad i = 1, 2.$$

It is easy to see that $P_i^2 = P_i$, i.e., that these matrices are projections. Note also that since P_i is not symmetric, we have for the 2-norm, that $\|I - P_i\|_2 = \|P_i\|_2 > 1$; see, e.g., [23].

The multiplicative Schwarz method starting with the initial vector $x^{(0)} \in \mathbb{R}^{N-1}$ is defined by

$$(3.2) \qquad\qquad x^{(k+1)} = T x^{(k)} + v, \quad k = 0, 1, 2, \ldots,$$

where $T = (I - P_2)(I - P_1)$ or $T = (I - P_1)(I - P_2)$, and the vector $v \in \mathbb{R}^{N-1}$ is defined to make the method consistent. For example, for the iteration matrix $T = (I - P_2)(I - P_1)$ the consistency condition $u^N = T u^N + v$ yields

$$v = (I - T) u^N = (P_1 + P_2 - P_2 P_1) u^N,$$

which is (easily) computable since

$$P_i u^N = R_i^T A_i^{-1} R_i A u^N = R_i^T A_i^{-1} R_i f^N, \quad i = 1, 2.$$

The error of the multiplicative Schwarz iteration (3.2) is given by

$$e^{(k+1)} = u^N - x^{(k+1)} = (Tu^N + v) - (Tx^{(k)} + v) = Te^{(k)}, \quad k = 0, 1, 2, \ldots,$$

and hence $e^{(k+1)} = T^{k+1}e^{(0)}$ by induction. For any consistent norm $\| \cdot \|$, we therefore have

$$(3.3) \qquad \|e^{(k+1)}\| \leq \|T^{k+1}\| \, \|e^{(0)}\| \leq \|T\|^{k+1} \, \|e^{(0)}\|.$$

Our main goal in the following is the derivation of quantitative convergence bounds, where we consider both $T = (I - P_2)(I - P_1)$ and $T = (I - P_1)(I - P_2)$.

We point out that the analysis of the multiplicative Schwarz method in the following sections uses the unscaled linear algebraic system with A as in (2.4) having the entries (2.5) or (2.6). In Section 4.3 we explain why this analysis also applies to suitably scaled linear algebraic systems and in particular to the scaling suggested by Roos in [21].

4. Convergence bounds for the multiplicative Schwarz method. We start with a closer look at the structure of the iteration matrix T. Note that the matrices P_i from (3.1) satisfy

$$P_1 = R_1^T A_1^{-1} R_1 A = \begin{bmatrix} I_n \\ 0 \end{bmatrix} A_1^{-1} \begin{bmatrix} A_1 \mid be_n \mid 0 \end{bmatrix} = \left[\begin{array}{c|c|c} I_n & bA_1^{-1}e_n & 0 \\ \hline 0 & 0 & 0 \end{array} \right],$$

and

$$P_2 = R_2^T A_2^{-1} R_2 A = \begin{bmatrix} 0 \\ I_n \end{bmatrix} A_2^{-1} \begin{bmatrix} 0 \mid ce_1 \mid A_2 \end{bmatrix} = \left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & cA_2^{-1}e_1 & I_n \end{array} \right],$$

where $e_1, e_n \in \mathbb{R}^n$. We now denote

$$(4.1) \qquad \left[\frac{p^{(1)}}{\pi^{(1)}} \right] \equiv bA_1^{-1}e_n \quad \text{and} \quad \left[\frac{\pi^{(2)}}{p^{(2)}} \right] \equiv cA_2^{-1}e_1,$$

where $p^{(i)} = [p_1^{(i)}, \ldots, p_m^{(i)}]^T \in \mathbb{R}^m$ for $i = 1, 2$, and $\pi^{(1)}$ and $\pi^{(2)}$ are scalars. Then

$$I - P_2 = \left[\begin{array}{c|c|c} I_{m-1} & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & -\left[\dfrac{\pi^{(2)}}{p^{(2)}}\right] & 0 \end{array} \right], \qquad I - P_1 = \left[\begin{array}{c|c|c} 0 & -\left[\dfrac{p^{(1)}}{\pi^{(1)}}\right] & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & I_{m-1} \end{array} \right],$$

which gives

$$(4.2) \qquad (I - P_2)(I - P_1) = \left[\begin{array}{c|c|c} 0 & -p^{(1)} & 0 \\ \hline 0 & p_m^{(1)}\pi^{(2)} & 0 \\ \hline 0 & p_m^{(1)}p^{(2)} & 0 \end{array} \right] = \begin{bmatrix} -p^{(1)} \\ p_m^{(1)}\pi^{(2)} \\ p_m^{(1)}p^{(2)} \end{bmatrix} e_{n+1}^T$$

and

$$(4.3) \qquad (I - P_1)(I - P_2) = \begin{bmatrix} p_1^{(2)}p^{(1)} \\ p_1^{(2)}\pi^{(1)} \\ -p^{(2)} \end{bmatrix} e_{n-1}^T,$$

where $e_{n+1}, e_{n-1} \in \mathbb{R}^{N-1}$. Thus, both iteration matrices have rank one, and we can apply the following observation.

PROPOSITION 4.1. *Let T be a square matrix of rank one, i.e., $T = uv^T$ for some vectors u, v. Then $T^2 = \rho T$, with $\rho = v^T u$, and consequently $T^{k+1} = \rho^k T$ for $k \geq 0$.*

Proof. The proof follows by direct computation. ☐

COROLLARY 4.2. *In the notation established above, let* $T = (I - P_2)(I - P_1)$ *or* $T = (I - P_1)(I - P_2)$. *Then for any* $k \geq 0$ *we have*

$$(4.4) \qquad T^{k+1} = \rho^k T, \quad where \quad \rho \equiv p_m^{(1)} p_1^{(2)}.$$

Proof. Applying Proposition 4.1 to either (4.2) or (4.3) produces the desired result. ☐

Equation (4.4) shows, in particular, that $\|T^{k+1}\| = |\rho|^k \|T\|$ holds for any matrix norm $\|\cdot\|$. In order to obtain a convergence bound for the multiplicative Schwarz method we will bound $|\rho|$ and $\|T\|_\infty$. The following lemma will be essential in our derivations.

LEMMA 4.3. *In the notation established above,*

$$\begin{bmatrix} p^{(1)} \\ \pi^{(1)} \end{bmatrix} = \pi^{(1)} \begin{bmatrix} -b_H A_H^{-1} e_m \\ 1 \end{bmatrix}, \qquad \pi^{(1)} = \frac{b}{a - c b_H \left(A_H^{-1} \right)_{m,m}},$$

$$\begin{bmatrix} \pi^{(2)} \\ p^{(2)} \end{bmatrix} = \pi^{(2)} \begin{bmatrix} 1 \\ -c_h A_h^{-1} e_1 \end{bmatrix}, \qquad \pi^{(2)} = \frac{c}{a - b c_h \left(A_h^{-1} \right)_{1,1}}.$$

Proof. From (4.1) we know that $p^{(1)}$, $p^{(2)}$, $\pi^{(1)}$, and $\pi^{(2)}$ solve the systems

$$\left[\begin{array}{c|c} A_H & b_H e_m \\ \hline c e_m^T & a \end{array} \right] \begin{bmatrix} p^{(1)} \\ \pi^{(1)} \end{bmatrix} = b e_n, \qquad \left[\begin{array}{c|c} a & b e_1^T \\ \hline c_h e_1 & A_h \end{array} \right] \begin{bmatrix} \pi^{(2)} \\ p^{(2)} \end{bmatrix} = c e_1.$$

Hence the expressions for $p^{(1)}$, $p^{(2)}$, $\pi^{(1)}$, and $\pi^{(2)}$ can be obtained using Schur complements. ☐

Combining (4.4) and Lemma 4.3 gives

$$(4.5) \qquad \rho = \frac{b\, b_H \left(A_H^{-1} \right)_{m,m}}{a - c\, b_H \left(A_H^{-1} \right)_{m,m}} \cdot \frac{c\, c_h \left(A_h^{-1} \right)_{1,1}}{a - b\, c_h \left(A_h^{-1} \right)_{1,1}}.$$

In order to bound $|\rho|$ we thus need to bound certain entries of inverses of the tridiagonal Toeplitz matrices A_H and A_h. In the following lemma we show that this is straightforward in the case of an M-matrix. As we will see later in Lemma 4.5 and Lemma 4.8, the matrix A_h is an M-matrix for both the upwind and the central difference scheme. However, while A_H is an M-matrix for the upwind scheme, it is not an M-matrix in the most common situation for the central difference scheme. We then have to use a different technique for bounding the entry $(A_H^{-1})_{1,1}$; see Section 4.2.

Recall that a nonsingular matrix $B = [b_{i,j}]$ is called an M-*matrix* when $b_{i,i} > 0$ for all i, $b_{i,j} \leq 0$ for all $i \neq j$, and $B^{-1} \geq 0$ (elementwise).

LEMMA 4.4. *Let B be an $\ell \times \ell$ tridiagonal Toeplitz matrix,*

$$B = \begin{bmatrix} \hat{a} & \hat{b} & & \\ \hat{c} & \ddots & \ddots & \\ & \ddots & \ddots & \hat{b} \\ & & \hat{c} & \hat{a} \end{bmatrix},$$

with $\hat{a} > 0$ and $\hat{b}, \hat{c} < 0$. Moreover, let B be diagonally dominant, i.e., $\hat{a} \geq |\hat{b}| + |\hat{c}|$. Then B is an M-matrix with $B^{-1} > 0$ (elementwise),

$$(4.6) \qquad \left(B^{-1} \right)_{\ell,\ell} = \left(B^{-1} \right)_{1,1} \leq \min \left\{ \frac{1}{|\hat{b}|}, \frac{1}{|\hat{c}|} \right\},$$

and the entries of B^{-1} decay along the columns away from the diagonal. In particular,

$$\left(B^{-1}\right)_{1,1} > \left(B^{-1}\right)_{i,1} \quad for \quad 1 < i \leq \ell,$$
$$\left(B^{-1}\right)_{\ell,\ell} > \left(B^{-1}\right)_{i,\ell} \quad for \quad 1 \leq i < \ell.$$

Proof. The matrix B is an M-matrix since its entries satisfy the sign condition and B is irreducibly diagonally dominant; see, e.g., [4, Theorem 6.2.3, Condition M35] or [10, Criterion 4.3.10]. The elementwise nonnegativity of the inverse, $B^{-1} > 0$, follows since the M-matrix B is irreducible; see, e.g., [4, Theorem 6.2.7] or [10, Theorem 4.3.11].

Since B is a tridiagonal Toeplitz matrix, its $(1,1)$ and (ℓ,ℓ) minors are equal. Therefore, the classical formula $B^{-1} = (\det(B))^{-1}\mathrm{adj}(B)$ implies that $\left(B^{-1}\right)_{1,1} = \left(B^{-1}\right)_{\ell,\ell}$. Moreover, since $\hat{a} \geq |\hat{b}| + |\hat{c}|$, we can apply [20, Lemma 2.1, equation (2.8)] to obtain

$$\left(B^{-1}\right)_{1,1} \leq \frac{1}{\hat{a} - |\hat{b}|} \leq \frac{1}{|\hat{c}|}, \quad \left(B^{-1}\right)_{\ell,\ell} \leq \frac{1}{\hat{a} - |\hat{c}|} \leq \frac{1}{|\hat{b}|}.$$

Finally, the bounds on the entries of B^{-1} are special cases of [20, Theorem 3.11], where it was shown that

$$\left(B^{-1}\right)_{i,j} \leq \omega^{i-j} \left(B^{-1}\right)_{j,j} \ \text{ for } \ i \geq j \quad \text{and} \quad \left(B^{-1}\right)_{i,j} \leq \tau^{j-i} \left(B^{-1}\right)_{j,j} \ \text{ for } \ i \leq j,$$

with some $\tau, \omega \in (0,1)$. (They can be expressed explicitly using the entries of B.) $\quad\square$

In the next two subsections we separately treat the upwind and the central difference schemes.

4.1. Bounds for the upwind scheme. Using Lemma 4.4 we can prove the following result for the upwind scheme.

LEMMA 4.5. *For the upwind scheme both matrices A_H and A_h satisfy the assumptions of Lemma 4.4, and the related quantities from Lemma 4.3 satisfy*

$$|\pi^{(1)}| \leq 1, \quad \|p^{(1)}\|_\infty = |p_m^{(1)}| \leq \frac{\epsilon}{\epsilon + \alpha H}, \quad |\pi^{(2)}| \leq 1, \quad \|p^{(2)}\|_\infty = |p_1^{(2)}| \leq 1.$$

Proof. It is easy to see from (2.5) that both matrices A_H and A_h resulting from the upwind scheme satisfy the assumptions of Lemma 4.4. Thus, from (4.6) we have

$$|b_H|\left(A_H^{-1}\right)_{m,m} \leq 1 \quad \text{and} \quad |c_h|\left(A_h^{-1}\right)_{1,1} \leq 1.$$

Moreover, $a > 0$ and $b, c < 0$ as well as $a + b + c = \beta \geq 0$, so that

$$|\pi^{(1)}| = \frac{|b|}{a + c|b_H|\left(A_H^{-1}\right)_{m,m}} \leq \frac{|b|}{a + c} \leq 1,$$
$$|\pi^{(2)}| = \frac{|c|}{a + b|c_h|\left(A_h^{-1}\right)_{1,1}} \leq \frac{|c|}{a + b} \leq 1.$$

Using these inequalities and the fact that the entries of A_h^{-1} decay along a column away from the diagonal yields

$$\|p^{(2)}\|_\infty = |p_1^{(2)}| = |\pi^{(2)}|\,|c_h|\left(A_h^{-1}\right)_{1,1} \leq 1.$$

Using the decay of the entries of A_H^{-1} and

$$|b_H|\left(A_H^{-1}\right)_{m,m} \leq \left|\frac{b_H}{c_H}\right|,$$

which follows from (4.6), as well as the definitions of the entries in (2.5), we obtain

$$\|p^{(1)}\|_\infty = |p_m^{(1)}| = |\pi^{(1)}| \, |b_H| \, \left(A_H^{-1}\right)_{m,m} \leq \left|\frac{b_H}{c_H}\right| = \frac{\epsilon}{\epsilon + \alpha H}. \qquad \square$$

We can now state our main result of this subsection.

THEOREM 4.6. *For the upwind scheme we have*

$$(4.7) \qquad\qquad |\rho| \leq \frac{\epsilon}{\epsilon + \alpha H}$$

and

$$\|(I - P_2)(I - P_1)\|_\infty \leq \frac{\epsilon}{\epsilon + \alpha H}, \qquad \|(I - P_1)(I - P_2)\|_\infty \leq 1.$$

Thus, the error of the multiplicative Schwarz method (3.2) *satisfies*

$$\frac{\|e^{(k+1)}\|_\infty}{\|e^{(0)}\|_\infty} \leq \begin{cases} \left(\frac{\epsilon}{\epsilon+\alpha H}\right)^{k+1}, & \text{if } T = (I - P_2)(I - P_1), \\ \left(\frac{\epsilon}{\epsilon+\alpha H}\right)^{k}, & \text{if } T = (I - P_1)(I - P_2). \end{cases}$$

Proof. For the bound on $|\rho|$ we apply Lemma 4.5 to the expression $\rho = p_m^{(1)} p_1^{(2)}$ from (4.4). From (4.2) and (4.3) we respectively see that

$$\|(I - P_2)(I - P_1)\|_\infty = \left\| \begin{bmatrix} -p^{(1)} \\ p_m^{(1)} \pi^{(2)} \\ p_m^{(1)} p^{(2)} \end{bmatrix} \right\|_\infty \qquad \text{and}$$

$$\|(I - P_1)(I - P_2)\|_\infty = \left\| \begin{bmatrix} p_1^{(2)} p^{(1)} \\ p_1^{(2)} \pi^{(1)} \\ -p^{(2)} \end{bmatrix} \right\|_\infty .$$

Thus, using Lemma 4.5,

$$\|(I - P_2)(I - P_1)\|_\infty = \max\left\{ |p_m^{(1)}|, \ |p_m^{(1)} \pi^{(2)}|, \ |p_m^{(1)} p_1^{(2)}| \right\} \leq |p_m^{(1)}| \leq \frac{\epsilon}{\epsilon + \alpha H},$$

$$\|(I - P_1)(I - P_2)\|_\infty = \max\left\{ |p_1^{(2)} p_m^{(1)}|, \ |p_1^{(2)} \pi^{(1)}|, \ |p_1^{(2)}| \right\} \leq |p_1^{(2)}| \leq 1.$$

Using these bounds and (4.4) in the first inequality in (3.3) yields the convergence bound for the multiplicative Schwarz method. $\square$

Suppose that $\epsilon < \alpha H$, which is a reasonable assumption in our context. Then

$$|\rho| \leq \frac{\epsilon}{\epsilon + \alpha H} = \frac{\epsilon}{\alpha H} + \mathcal{O}\left(\left(\frac{\epsilon}{\alpha H}\right)^2\right).$$

This expression shows that the convergence of the multiplicative Schwarz method will be very rapid in case of the upwind scheme and a strong convection-dominance. Numerical examples are shown in Section 5.

Note that since $\frac{2}{N} = H + h \leq 2H$, we have $\frac{1}{N} \leq H$ and hence

$$(4.8) \qquad\qquad |\rho| \leq \frac{\epsilon}{\epsilon + \alpha H} \leq \frac{\epsilon}{\epsilon + \frac{\alpha}{N}}.$$

Using the expression on the right-hand of (4.8) in Theorem 4.6 would give (slightly) weaker convergence bounds for the multiplicative Schwarz method. However, the right-hand side of (4.8) represents a more convenient bound on the convergence factor which directly depends on the parameters ϵ, α, and N of our problem.

4.2. Bounds for the central difference scheme. We will now consider the discretization by the central difference scheme, i.e., the matrix A with the entries given by (2.6). It turns out that the analysis for this scheme is more complicated than for the upwind scheme since, as mentioned earlier, the matrix A_H need not be an M-matrix. Moreover, as we will see below, the multiplicative Schwarz method may not converge when the parameter m is odd.

The following result about the entries a, b, and c of A will be frequently used below.

LEMMA 4.7. *For the central difference scheme we have*

$$(4.9) \qquad a > 0, \quad c, b < 0, \quad -(c+b) = |c| + |b| = a - \beta \le a, \quad and \quad \left| \frac{b}{a} \right| < 1.$$

Proof. The inequalities $a > 0$ and $c < 0$ are obvious from (2.6). From (2.2)–(2.3) and $N \ge 4$ we have that

$$(4.10) \qquad\qquad \alpha h = 2\epsilon \frac{2 \ln N}{N} < 2\epsilon,$$

and thus

$$b = \frac{\alpha h - 2\epsilon}{h(H + h)} < 0.$$

Moreover, $-(c+b) = a - \beta \le a$, which yields

$$\left| \frac{b}{a} \right| = \left| \frac{b}{\beta - (c+b)} \right| < 1. \qquad \square$$

We next consider the matrix A_h from the central difference scheme.

LEMMA 4.8. *The matrix A_h from the central difference scheme satisfies the assumptions of Lemma 4.4, and for the corresponding quantities from Lemma 4.3 we have*

$$|\pi^{(2)}| \le 1 \quad and \quad \|p^{(2)}\|_\infty = \left| p_1^{(2)} \right| \le 1.$$

Proof. The inequalities $a_h > 0$ and $c_h < 0$ are obvious from (2.6) and using (4.10) we get

$$b_h = \frac{\alpha h - 2\epsilon}{2h^2} < 0.$$

Since also

$$|c_h| + |b_h| = \frac{2\epsilon}{h^2} \le a_h,$$

the matrix A_h satisfies the assumptions of Lemma 4.4. Thus, in particular, $|c_h| \left(A_h^{-1} \right)_{1,1} \le 1$. Using also (4.9) gives

$$|\pi^{(2)}| = \frac{|c|}{a + b|c_h| \left(A_h^{-1} \right)_{1,1}} \le \frac{|c|}{a + b} = \frac{|c|}{|c| + \beta} \le 1.$$

Finally, since the entries of A_h^{-1} decay along a column away from the diagonal, we obtain $\|p^{(2)}\|_\infty = \left| p_1^{(2)} \right| = |\pi^{(2)}| |c_h| \left(A_h^{-1} \right)_{1,1} \le 1. \qquad \square$

We now concentrate on bounding the quantities from Lemma 4.3 related to the matrix A_H for the central difference scheme. We will distinguish the three cases $\alpha H < 2\epsilon$, $\alpha H = 2\epsilon$, and $\alpha H > 2\epsilon$ or, equivalently, the cases that the entry

$$b_H = \frac{\alpha H - 2\epsilon}{2H^2}$$

of A_H is negative, zero, or positive. It is clear from (4.5) that the sign of b_H is important for the value $|\rho|$.

A simple computation shows that $b_H \leq 0$ if and only if

$$\epsilon \geq \frac{\alpha}{N + 2 \ln N} \, .$$

If $\epsilon \ll \alpha \approx 1$, then this condition means that $\epsilon(N + 2 \ln N) = \mathcal{O}(1)$, which is an unrealistic assumption on the discretization parameter N. Nevertheless, we include the case $b_H \leq 0$ for completeness.

We first assume that

(4.11) $$\alpha H < 2\epsilon,$$

which means that $b_H < 0$.

LEMMA 4.9. *If $\alpha H < 2\epsilon$, then the matrix A_H from the central difference scheme satisfies the assumptions of Lemma 4.4, and we have*

$$|\pi^{(1)}| \leq 1 \quad and \quad \|p^{(1)}\|_\infty = |p_m^{(1)}| < \frac{\epsilon}{\epsilon + \frac{\alpha}{N}} \cdot$$

Proof. The inequalities $a_H > 0$ and $c_H < 0$ are obvious from (2.6), and $b_H < 0$ holds because of (4.11). Moreover,

$$|c_H| + |b_H| = \frac{\alpha}{2H} + \frac{\epsilon}{H^2} + \frac{\epsilon}{H^2} - \frac{\alpha}{2H} = \frac{2\epsilon}{H^2} \leq a_H,$$

so that the matrix A_H satisfies the assumptions of Lemma 4.4. In particular,

$$|b_H| \left(A_H^{-1}\right)_{m,m} \leq 1.$$

Using (4.9) we obtain

$$|\pi^{(1)}| = \frac{|b|}{a + c|b_H| \left(A_H^{-1}\right)_{m,m}} \leq \frac{|b|}{a + c} \leq 1.$$

Moreover, using that the entries of A_H^{-1} decay along a column away from the diagonal as well as

$$|b_H| \left(A_H^{-1}\right)_{m,m} \leq \frac{|b_H|}{|c_H|},$$

which follows from (4.6), we see that

$$\|p^{(1)}\|_\infty = |p_m^{(1)}| = |\pi^{(1)}| \, |b_H| \left(A_H^{-1}\right)_{m,m} \leq \frac{|c_H|}{|b_H|} = \frac{2\epsilon - \alpha H}{2\epsilon + \alpha H} < \frac{2\epsilon - \alpha H + \alpha h}{2\epsilon + \alpha H + \alpha h}$$

$$< \frac{2\epsilon}{2\epsilon + \alpha(H + h)} = \frac{\epsilon}{\epsilon + \frac{\alpha}{N}},$$

where we used $h < H$ and $h + H = \frac{2}{N}$. □

Next we consider the (very) special case

(4.12) $$\alpha H = 2\epsilon,$$

which means that $b_H = 0$.

LEMMA 4.10. *If $\alpha H = 2\epsilon$, then the matrix A_H from the central difference scheme is nonsingular, and we have $|\pi^{(1)}| < 1$ and $p^{(1)} = 0$.*

Proof. If $b_H = 0$, then A_H is lower triangular and nonsingular since $a_H > 0$. Using the definitions of $p^{(1)}$ and $\pi^{(1)}$ from Lemma 4.3 and the last inequality in (4.9) we obtain

$$p^{(1)} = -\frac{bb_H A_H^{-1} e_m}{a - b_H \left(A_H^{-1}\right)_{m,m} c} = 0, \qquad |\pi^{(1)}| = \left|\frac{b}{a - cb_H \left(A_H^{-1}\right)_{m,m}}\right| = \left|\frac{b}{a}\right| < 1. \qquad \square$$

The third case we consider is

$$(4.13) \qquad\qquad \alpha H > 2\epsilon,$$

which means that $b_H > 0$. This is the most common situation from a practical point of view, but now A_H does not satisfy the assumptions of Lemma 4.4. We therefore need a different approach for bounding the quantities from Lemma 4.3 and in particular the entries of the vector $A_H^{-1} e_m$. Note that because of (4.13) we have

$$-1 < \frac{2\epsilon - \alpha H}{2\epsilon + \alpha H} = \frac{b_H}{c_H} < 0.$$

LEMMA 4.11. *If $\alpha H > 2\epsilon$, then the matrix A_H from the central difference scheme is a nonsingular tridiagonal Toeplitz matrix with the entries $a_H, b_H > 0$, and $c_H < 0$. Moreover,*

$$(4.14) \qquad 0 < \left|(A_H^{-1})_{i,m}\right| \leq (A_H^{-1})_{m,m} \frac{1 - \left(\frac{b_H}{c_H}\right)^i}{1 - \left(\frac{b_H}{c_H}\right)^m} \cdot \left|\frac{b_H}{c_H}\right|^{m-i}, \qquad i = 1, \ldots, m,$$

where the second inequality in (4.14) is an equality if $\beta = 0$. If $m = N/2 - 1$ is even, then

$$(4.15) \qquad\qquad b_H \left|(A_H^{-1})_{i,m}\right| < 2, \qquad i = 1, \ldots, m,$$

and

$$(4.16) \qquad\qquad b_H (A_H^{-1})_{m,m} \leq \frac{1 - \left|\frac{b_H}{c_H}\right|^m}{\left|\frac{c_H}{b_H}\right| + \left|\frac{b_H}{c_H}\right|^m} < \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}}.$$

Proof. The inequalities $a_H > 0$ and $c_H < 0$ are obvious from (2.6), and $b_H > 0$ holds because of (4.13).

In order to see that A_H is nonsingular, note that eigenvalues of the tridiagonal Toeplitz matrix A_H are given by

$$\lambda_i = a_H + 2\sqrt{b_H c_H} \cos\left(\frac{i\pi}{m+1}\right), \qquad i = 1, \ldots, m.$$

Since $b_H c_H < 0$, the number $\sqrt{b_H c_H}$ is purely imaginary, and hence all eigenvalues are nonzero.

Adapting [25, Theorem 2] to our notation (and formulating this result in terms of columns instead of rows) shows that the entries of the vector $\xi \equiv [\xi_1, \ldots, \xi_m]^T \equiv A_H^{-1} e_m$ can be written as

$$\xi_i = (-1)^{m-i} b_H^{m-i} \frac{\theta_{i-1}}{\theta_m}, \qquad i = 1, \ldots, m,$$

where

$$(4.17) \qquad \theta_i \equiv a_H \theta_{i-1} - b_H c_H \theta_{i-2}, \quad \theta_0 \equiv 1, \ \theta_1 \equiv a_H.$$

Since $b_H c_H < 0$ and $a_H > 0$, we have $\theta_i > 0$ for all $i \geq 0$ and $\xi_i \neq 0$. Since $b_H > 0$, ξ_i changes sign like $(-1)^{m-i}$, and $\xi_m > 0$. Consequently, the first inequality in (4.14) holds.

If we define the sequence of positive numbers

$$\omega_i \equiv \frac{\theta_{i-1}}{\theta_i}, \quad i = 1, 2, \dots,$$

then

$$(4.18) \qquad \xi_i = (-1)^{m-i} b_H^{m-i} \prod_{j=i}^{m} \omega_j = \xi_m (-1)^{m-i} b_H^{m-i} \prod_{j=i}^{m-1} \omega_j, \quad i = 1, \dots, m.$$

We will prove by induction that

$$(4.19) \qquad \omega_i \leq -\frac{c_H^i - b_H^i}{c_H^{i+1} - b_H^{i+1}}$$

for all $i \geq 1$, with equality if $\beta = 0$. For $i = 1$ we have

$$-\frac{c_H - b_H}{c_H^2 - b_H^2} = \frac{1}{-(c_H + b_H)} = \frac{1}{a_H - \beta} \geq \frac{1}{a_H} = \omega_1,$$

with equality if $\beta = 0$. Using the recurrence (4.17), the inequality $a_H \geq -(c_H + b_H)$, which is an equality if $\beta = 0$, and the induction hypothesis, we obtain

$$\frac{1}{\omega_i} = a_H - \omega_{i-1} b_H c_H \geq -(c_H + b_H) + \frac{c_H^{i-1} - b_H^{i-1}}{c_H^i - b_H^i} b_H c_H = -\frac{c_H^{i+1} - b_H^{i+1}}{c_H^i - b_H^i},$$

again with equality if $\beta = 0$.

Combining (4.18) and (4.19) yields

$$(4.20) \qquad |\xi_i| \leq \xi_m b_H^{m-i} \left| \frac{c_H^i - b_H^i}{c_H^m - b_H^m} \right| = \xi_m \frac{1 - \left(\frac{b_H}{c_H}\right)^i}{1 - \left(\frac{b_H}{c_H}\right)^m} \cdot \left| \frac{b_H}{c_H} \right|^{m-i},$$

showing the second inequality in (4.14), which is an equality if $\beta = 0$.

Now let m be even. Using (4.19) we obtain

$$(4.21) \qquad b_H \xi_m \leq -b_H \frac{c_H^m - b_H^m}{c_H^{m+1} - b_H^{m+1}} = \frac{1 - \left|\frac{b_H}{c_H}\right|^m}{\left|\frac{c_H}{b_H}\right| + \left|\frac{b_H}{c_H}\right|^m} < 1 - \left|\frac{b_H}{c_H}\right|^m,$$

which contains the first inequality in (4.16). Using (4.20) and (4.21) we obtain

$$|\xi_i| < \xi_m \frac{2}{1 - \left|\frac{b_H}{c_H}\right|^m} < \xi_m \frac{2}{b_H \xi_m} = \frac{2}{b_H},$$

which shows (4.15). Let us write

$$\left| \frac{b_H}{c_H} \right| = \frac{\alpha H - 2\epsilon}{\alpha H + 2\epsilon} = 1 - \frac{2\epsilon}{\epsilon + \frac{\alpha H}{2}} \equiv 1 - \nu.$$

Using (4.13) we have $0 < \nu < 1$, and by induction it can be easily shown that the inequality $1 - (1 - \nu)^m < m\nu$ holds for every integer $m \geq 2$. Thus,

$$b_H \xi_m < 1 - \left| \frac{b_H}{c_H} \right|^m = 1 - (1 - \nu)^m < m\nu = \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}} \ ,$$

which proves the second inequality in (4.16).   □

Using Lemma 4.11 and the assumption that m is even, we can bound the quantities from Lemma 4.3 related to the matrix A_H from the central difference scheme as follows.

LEMMA 4.12. *If* (4.13) *holds and if* $m = N/2 - 1$ *is even, then*

$$|\pi^{(1)}| < 1, \quad |p_m^{(1)}| < \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}}, \quad \|p^{(1)}\|_\infty < 2.$$

Proof. From (4.9) we know that $c < 0$, and from Lemma 4.11 we know that $b_H > 0$ and $\left(A_H^{-1} \right)_{m,m} > 0$. Therefore,

$$\left| \pi^{(1)} \right| = \frac{|b|}{a + |c| b_H \left(A_H^{-1} \right)_{m,m}} < \left| \frac{b}{a} \right| < 1,$$

where we have used (4.9). Thus, using also (4.16), we obtain

$$|p_m^{(1)}| = |\pi^{(1)}| \, b_H \left(A_H^{-1} \right)_{m,m} < \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}} \ .$$

Finally, (4.15) implies that $\|p^{(1)}\|_\infty = |\pi^{(1)}| b_H \|A_H^{-1} e_m\|_\infty < 2$.   □

We can now formulate an analogue of Theorem 4.6 for the central difference scheme.

THEOREM 4.13. *For the central difference scheme we have*

(4.22)
$$|\rho| < \begin{cases} \frac{\epsilon}{\epsilon + \frac{\alpha}{N}} & \text{if } \alpha H \leq 2\epsilon, \\ \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}} & \text{if } \alpha H > 2\epsilon \text{ and } m = N/2 - 1 \text{ is even.} \end{cases}$$

If $\alpha H \leq 2\epsilon$, *we have*

$$\|(I - P_2)(I - P_1)\|_\infty \leq 1, \quad \|(I - P_1)(I - P_2)\|_\infty \leq 1,$$

and if $\alpha H > 2\epsilon$, *we have*

$$\|(I - P_2)(I - P_1)\|_\infty < 2, \quad \|(I - P_1)(I - P_2)\|_\infty < 2.$$

Thus, the error of the multiplicative Schwarz method (3.2) *for both iteration matrices satisfies*

$$\frac{\|e^{(k+1)}\|_\infty}{\|e^{(0)}\|_\infty} < \begin{cases} \left(\frac{\epsilon}{\epsilon + \frac{\alpha}{N}} \right)^k & \text{if } \alpha H \leq 2\epsilon, \\ 2 \left(\frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}} \right)^k & \text{if } \alpha H > 2\epsilon \text{ and } m = N/2 - 1 \text{ is even.} \end{cases}$$

Proof. From (4.4) we know that $\rho = p_m^{(1)} p_1^{(2)}$, and hence the bounds on $|\rho|$ follow from $|p_1^{(2)}| \leq 1$ (Lemma 4.8), and Lemmas 4.9–4.10 for the case $\alpha H \leq 2\epsilon$, as well as Lemma 4.12 for the case $\alpha H > 2\epsilon$.

For the first iteration matrix we have

$$\|(I - P_1)(I - P_2)\|_\infty = \left\| \begin{bmatrix} \dfrac{p_1^{(2)} p^{(1)}}{p_1^{(2)} \pi^{(1)}} \\ -p^{(2)} \end{bmatrix} \right\|_\infty$$
$$= \max\{|p_1^{(2)}| \, \|p^{(1)}\|_\infty, |p_1^{(2)} \pi^{(1)}|, \|p^{(2)}\|_\infty\},$$

and for the second iteration matrix we have

$$\|(I - P_2)(I - P_1)\|_\infty = \left\| \begin{bmatrix} -p^{(1)} \\ \dfrac{p_m^{(1)} \pi^{(2)}}{p_m^{(1)} p^{(2)}} \end{bmatrix} \right\|_\infty$$
$$= \max\{\|p^{(1)}\|_\infty, |p_m^{(1)} \pi^{(2)}|, |p_m^{(1)}| \, \|p^{(2)}\|_\infty\}.$$

The bounds on these matrices now follow from the Lemmas 4.8, 4.9, 4.10, 4.12, and the error bound for the multiplicative Schwarz method follows from (3.3) and (4.4). □

As in the discussion of Theorem 4.6 we could use $\frac{1}{N} \leq H$ and $m = \frac{N}{2} - 1$, and thus obtain

$$|\rho| \leq \frac{2m\epsilon}{\epsilon + \frac{\alpha H}{2}} < \frac{N\epsilon}{\epsilon + \frac{\alpha}{2N}},$$

where the right-hand side again represents a bound on the convergence factor that directly depends of the parameters of our problem.

Because of the factor $2m \approx N$, the error bound for the central differences discretization can be significantly larger than for the upwind scheme. Thus, we expect that the multiplicative Schwarz method for the central differences discretization converges slower than for the upwind scheme, at least when $\alpha H > 2\epsilon$. An example with $\epsilon = 10^{-4}$ and $N = 198$, leading to $|\rho| = 8.3 \times 10^{-1}$ and a very slow convergence of the multiplicative Schwarz method is shown in Section 5. In this case, the bound (4.22) is even greater than one. It should be noted, however, that in a strongly convection-dominated case the situation $\epsilon N^2 = \mathcal{O}(1)$ is rather unrealistic.

Finally, let us discuss the situation when (4.13) holds, so that $-1 < b_H/c_H < 0$ but m is odd. In this case the multiplicative Schwarz method does not converge for all parameter choices, and we will now give an intuitive explanation for this fact. For simplicity, let $\beta = 0$. Then (4.19) yields

$$b_H (A_H^{-1})_{m,m} = b_H \xi_m = -\frac{1 - \left(\frac{b_H}{c_H}\right)^m}{\frac{c_H}{b_H} - \left(\frac{b_H}{c_H}\right)^m} = \frac{1 + \left|\frac{b_H}{c_H}\right|^m}{\left|\frac{c_H}{b_H}\right| - \left|\frac{b_H}{c_H}\right|^m}.$$

The essential inequality in (4.21) then fails to hold, and we may have $b_H (A_H^{-1})_{m,m} > 1$. Using the exact expression for ρ in (4.5), it is then easy to find parameters for which $|\rho| > 1$, and for which the multiplicative Schwarz method indeed diverges. Note that even if m is odd, it holds for sufficiently large values of m that $|\rho| < 1$, and the multiplicative Schwarz method converges to the discrete solution. However, in this case we cannot expect the convergence to be as fast as in the case of even values of m.

These problems with odd m correspond to the situation when equation (2.1) is discretized using central differences on a uniform mesh. Consider, for example, the discrete solution of the problem (2.1) with $\alpha = 1$, $\beta = 0$, $f(x) \equiv 1$, and $u_0 = u_1 = 0$, which can be found in [22, Section 4]. If the number of interior points of the uniform mesh is even, then the discrete

solution oscillates but with an amplitude bounded by one, so that some important information about the analytic solution is still preserved in the discrete solution. If the number of inner points is odd, the discrete solution is highly oscillating (cf. [22, Figure 4.1]), and therefore it does not provide much useful information about the analytic solution.

In our case of the Shishkin mesh, the multiplicative Schwarz method solves discrete problems on the coarse mesh and the fine mesh in an alternating way, and it combines the solutions of the two subproblems. If m is odd and the discrete solution on the coarse mesh is highly oscillating, then this solution is essentially useless, and the multiplicative Schwarz may fail to improve the approximation to the discrete solution.

4.3. Remarks on diagonally scaled linear algebraic systems. As mentioned in Section 2, the large difference between the mesh sizes H and h leads to highly ill-conditioned matrices A, both for the upwind and the central difference scheme. As shown by Roos [21] (for the upwind scheme), the ill-conditioning can be avoided by a simple diagonal scaling; cf. the numerical example in Section 2. In our notation, the linear algebraic system $Au^N = f^N$ is multiplied from the left by the diagonal matrix

$$(4.23) \qquad D = \left[\begin{array}{c|c|c} d_H I_m & & \\ \hline & d & \\ \hline & & d_h I_m \end{array} \right],$$

where for the upwind scheme we choose

$$d_H = \frac{H}{\alpha}, \qquad d = \frac{hH}{2\epsilon}, \qquad d_h = \frac{h^2}{\epsilon},$$

and for the central difference scheme we choose

$$d_H = \frac{2H}{\alpha}, \qquad d = \frac{hH + h^2}{2\epsilon}, \qquad d_h = \frac{h^2}{\epsilon}.$$

We will now explain the effect of a diagonal scaling with a matrix D as in (4.23) on our convergence analysis of the multiplicative Schwarz method.

Multiplying A from the left with D preserves the Toeplitz structure of the matrix as well as the M-matrix property of the submatrices (if it holds for the unscaled matrix). The derivations of all bounds in Sections 4.1 and 4.2 depend on these properties and on ratios between elements in the same row such as $|b/a|$ and $|b_H/c_H|$; see in particular Lemma 4.5 for the upwind and Lemma 4.11 for the central difference scheme. The ratios, however, are invariant under any diagonal scaling of the form (4.23). Consequently, all convergence bounds in Theorems 4.6 and 4.13 hold for the multiplicative Schwarz method applied to $DAu^N = Df^N$ for a diagonal matrix D as in (4.23) with *any* positive diagonal entries d_H, d, and d_h.

Note that, using a convenient scaling, one can also see the structure of the (scaled) matrices A_H and A_h. In particular, let $\beta = 0$ and $\epsilon \ll \alpha H$. Then there is a scaling such that the scaled matrix A_h is close to $\mathrm{tridiag}(-1, 2, -1)$. Another scaling can be found such that the scaled A_H is close to $\mathrm{tridiag}(-1, 1, 0)$ for the upwind scheme and close to $\mathrm{tridiag}(-1, 0, 1)$ for the central difference scheme. The matrix $(1/\sqrt{2})\mathrm{tridiag}(-1, 0, 1) \in \mathbb{R}^{m \times m}$ is orthogonal when m is even, but it is *singular* when m is odd. This gives another intuitive explanation for the problems occurring when an odd value of m is chosen in the central difference scheme (also see our discussion of this point above).

In our experiments, the diagonal scaling had virtually no effect on the actual convergence of the Schwarz method (similar to the GMRES method; see Figures 2.2–2.4). In the following section we therefore present only results for the unscaled systems with A as in (2.4)–(2.6).
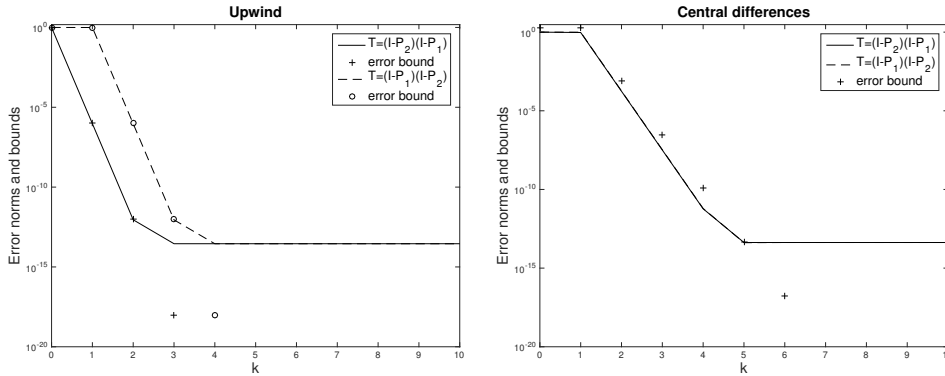
FIG. 5.1. *Convergence of multiplicative Schwarz and error bounds for $\epsilon = 10^{-8}$, $N = 198$.*
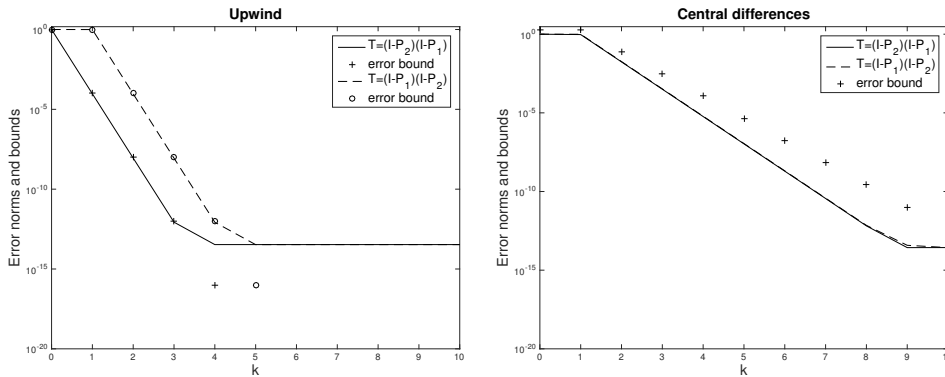

FIG. 5.2. *Convergence of multiplicative Schwarz and error bounds for $\epsilon = 10^{-6}$, $N = 198$.*

5. Numerical examples. We now illustrate the convergence behavior of the multiplicative Schwarz method applied to the Shishkin mesh discretization of the problem (2.1) with

$$\alpha = 1, \quad \beta = 0, \quad f(x) \equiv 1, \quad u_0 = u_1 = 0.$$

The analytic solution of this problem with $\epsilon = 0.01$ is shown in Figure 2.1.

We first consider $N = 198$, so that $m = N/2 - 1 = 98$ is even. Recall that the (unpreconditioned) GMRES method converges very slowly for both types of discretizations (upwind and central differences); see Figures 2.2–2.4.

For our experiments we computed $u^N = A^{-1} f^N$ using the backslash operator in MATLAB. (Applying iterative refinement in order to improve the numerical solution obtained in this way yields virtually the same results, so we do not consider iterative refinement here.) Using the solution obtained by MATLAB's backslash operator, we computed the error norms of the multiplicative Schwarz method by $\|e^{(k)}\|_\infty = \|x^{(k+1)} - u^N\|_\infty$ with $x^{(k+1)}$ as in (3.2) (rather than using the update formula $e^{(k)} = Te^{(k-1)}$). Consequently, the computed error norms stagnate on the level of the maximal attainable accuracy of the method. On the other hand, an error bound of the form $|\rho|^k$ for some $|\rho| < 1$ becomes arbitrarily small for $k \to \infty$.

We start with the upwind discretization. The left parts of Figures 5.1–5.5 show the error norms

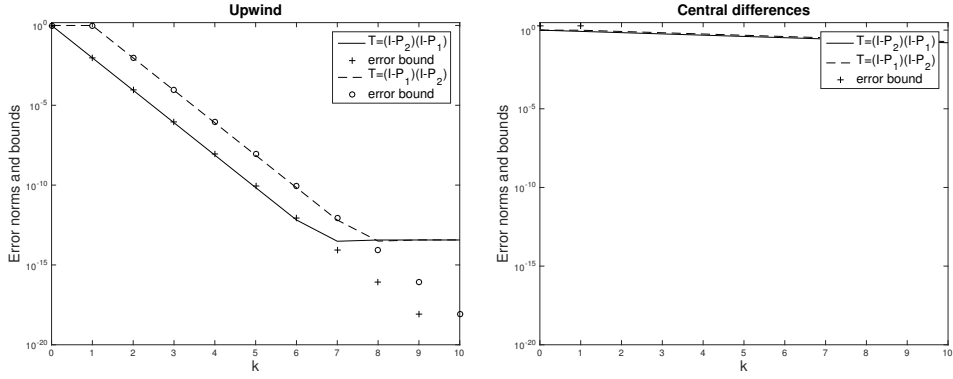$$\frac{\|e^{(k)}\|_\infty}{\|e^{(0)}\|_\infty}, \quad k = 0, 1, 2 \ldots,$$

FIG. 5.3. *Convergence of multiplicative Schwarz and error bounds for $\epsilon = 10^{-4}$, $N = 198$.*
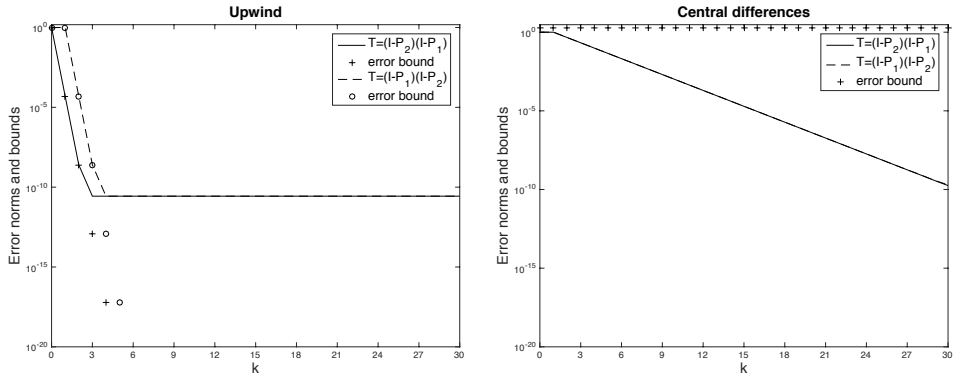

FIG. 5.4. *Convergence of multiplicative Schwarz and error bounds for $\epsilon = 10^{-8}$, $N = 10002$.*

for the iteration matrices $T = (I - P_1)(I - P_2)$ (solid) and $T = (I - P_1)(I - P_2)$ (dashed) as well as the corresponding upper bounds from Theorem 4.6, for increasing values of ϵ. We observe that the bounds are quite close to the actual errors. Moreover, in each case the error norm for the multiplicative Schwarz method with the iteration matrix $T = (I - P_1)(I - P_2)$ almost stagnates in the first step, as predicted by the bound in Theorem 4.6.

On the right parts of Figures 5.1–5.5 we display the error norms of the multiplicative Schwarz method and the corresponding convergence bounds from Theorem 4.13 for the central difference scheme. For our choice of parameters we have $\alpha H > 2\epsilon$. Note that the error norms are virtually the same for both iteration matrices. However, the bounds are not as tight as for the upwind scheme. For fixed N the bounds become weaker with increasing ϵ, i.e., decreasing convection-dominance. For our chosen parameters and $\epsilon = 10^{-4}$, giving $\epsilon N^2 = \mathcal{O}(1)$, the convergence of the multiplicative Schwarz method becomes very slow, and the bound (4.22) fails to predict convergence at all. The values of $|\rho|$ and the corresponding bounds from Theorems 4.6 and 4.13 are shown in Table 5.1 for the case $N = 198$. We observe that the bounds on $|\rho|$ for the upwind scheme are tighter than for the central difference scheme.

We also run the experiments for a larger value of N, namely $N = 10002$, to further illustrate our results. We consider the special cases $\epsilon N^2 \approx 1$ (Figure 5.4) and $\epsilon N \approx 1$ (Figure 5.5) which are mainly of theoretical interest. While the bound (4.7) for the upwind scheme is still tight and descriptive, the bound (4.22) for the central difference scheme does not predict convergence well. Note that the parameters used in Figure 5.5 yield $\alpha H \approx 1.9959 \times 10^{-4} < 2\epsilon$,
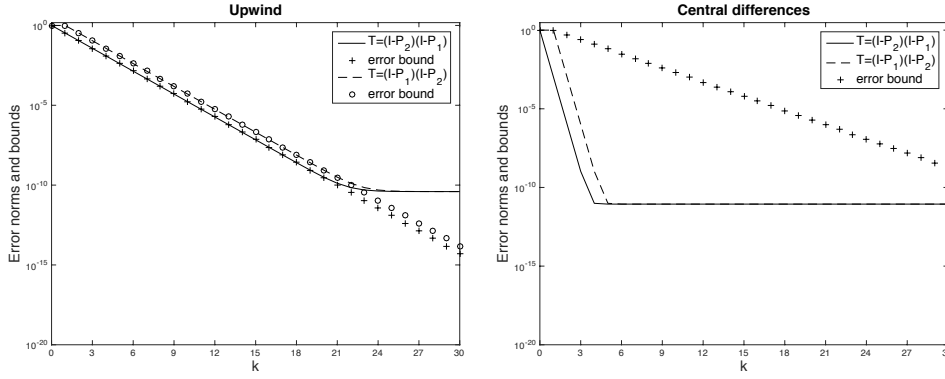
FIG. 5.5. *Convergence of multiplicative Schwarz and error bounds for $\epsilon = 10^{-4}$, $N = 10002$.*

TABLE 5.1
Values of $|\rho|$ computed using (4.5) and the corresponding bounds (4.7) and (4.22) for different ϵ and $N = 198$.

	upwind		central differences	
ϵ	$|\rho|$ (4.5)	bound (4.7)	$|\rho|$ (4.5)	bound (4.22)
10^{-8}	9.4×10^{-7}	9.9×10^{-7}	1.8×10^{-4}	3.9×10^{-4}
10^{-6}	9.4×10^{-5}	9.9×10^{-5}	1.8×10^{-2}	3.9×10^{-2}
10^{-4}	9.3×10^{-3}	9.8×10^{-3}	8.3×10^{-1}	3.8×10^{0}

and hence the right part of Figure 5.5 shows error norms and the convergence bound corresponding to the case $\alpha H \leq 2\epsilon$ in Theorem 4.13.

We conclude our numerical experiments by applying GMRES to the linear algebraic system *preconditioned with multiplicative Schwarz*, i.e., the linear algebraic system (6.1), in the case $N = 198$. The true (preconditioned) relative residual norms are shown in Figures 5.6–5.8. In all cases convergence is achieved in two iterations, which is explained in the next section. These figures are the counterparts to Figures 2.2–2.4, where GMRES makes little progress until iteration 198.

6. Concluding discussion. We studied the convergence of the multiplicative Schwarz method applied to upwind and central finite difference discretizations of one-dimensional singularly perturbed convection-diffusion model problems posed on a Shishkin mesh. The matrices that arise from the discretization are nonsymmetric and usually nonnormal as well as ill-conditioned, which leads to very slow convergence of standard iterative solvers like the (unpreconditioned) GMRES method.

In the simple one-dimensional case analyzed in this paper, the Shishkin mesh divides the discretized domain into two local subdomains where the solution presents a different characteristic nature. Therefore, a solution approach based on domain decomposition methods seemed only natural. For the upwind scheme, we proved rapid convergence of the multiplicative Schwarz method for all relevant problem parameters. The convergence for the central difference scheme is slower but still rapid when $N^2\epsilon < \alpha$ and if $N/2 - 1$ is even.

Based on (3.2), it is clear that the multiplicative Schwarz method can be seen as a *Richardson iteration* for the system
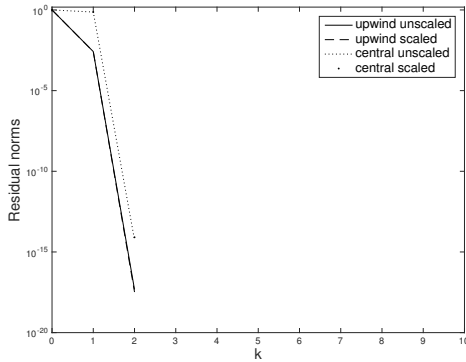
$$(6.1) \qquad\qquad (I - T)x = v.$$

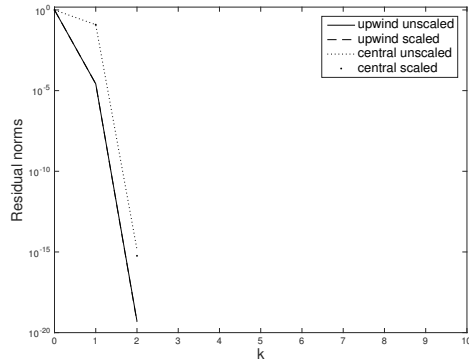FIG. 5.6. *GMRES convergence for $\epsilon = 10^{-4}$.*

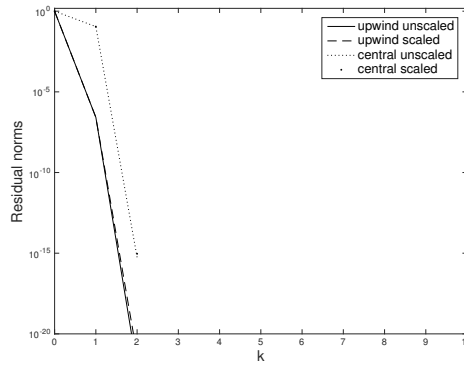FIG. 5.7. *GMRES convergence for $\epsilon = 10^{-6}$.*

FIG. 5.8. *GMRES convergence for $\epsilon = 10^{-8}$.*

Furthermore, the iteration scheme (3.2) can be written in the form

$$x^{(k+1)} = x^{(k)} + (I - T)(x - x^{(k)}),$$

so that the multiplicative Schwarz method as well as GMRES applied to (6.1) obtain their approximations from the Krylov subspace $\mathcal{K}_k(I - T, r_0)$. Consequently, in terms of the residual norm, GMRES applied to (6.1) will always converge faster than the multiplicative Schwarz method. Moreover, if one applies GMRES to (6.1), then the multiplicative Schwarz method can be seen as a *preconditioner* for the GMRES method; see, e.g., [11] where this approach is taken. The preconditioner M such that $M^{-1}Ax = M^{-1}b$ results in (6.1), can formally be written as $M = A(I - T)^{-1}$; see, e.g., [13, Lemma 2.3].

In general, if a matrix T satisfies $r = \operatorname{rank}(T)$, then for any initial residual r_0 we have

$$\dim\left(\mathcal{K}_k(I - T, r_0)\right) \leq r + 1,$$

so that GMRES applied to the system (6.1) converges to the solution in at most $r + 1$ steps (in exact arithmetic). In the one-dimensional model problem studied in this paper we have a matrix T with $r = 1$. Thus, GMRES applied to (6.1) converges in (at most) two steps (see Figures 5.6–5.8) even if the multiplicative Schwarz iteration itself converges slowly or diverges, which may happen for the central difference scheme and m odd. In a generalization of the approach presented in this paper to two- or three-dimensional problems and hence more

complicated Shishkin meshes with several transition points, one can possibly exploit a low rank structure of the iteration matrix as well.

It is important to point out that, typically, in practical implementations the local subdomain problems will not be solved exactly. In the case of inexact solves the bounds obtained in this paper and the exact termination of GMRES in $r+1$ steps will no longer hold. Nevertheless, the theory for the exact case presented here gives an indication of the behavior in the inexact case. This is a standard approach in the context of preconditioning. An example of this framework is given by the saddle-point preconditioners for which GMRES terminates in a few steps; see [3]. In the context of domain decomposition methods, in particular for Schwarz methods, the concept of inexact subdomain solves was investigated, e.g., in [2, Section 4]. See also [9], where a similar situation is described for algebraic optimized Schwarz methods.

REFERENCES

[1] V. B. ANDREEV AND N. V. KOPTEVA, *A study of difference schemes with the first derivative approximated by a central difference ratio*, Comput. Math. Math. Phys., 36 (1996), pp. 1065–1078.

[2] M. BENZI, A. FROMMER, R. NABBEN, AND D. B. SZYLD, *Algebraic theory of multiplicative Schwarz methods*, Numer. Math., 89 (2001), pp. 605–639.

[3] M. BENZI AND A. J. WATHEN, *Some preconditioning techniques for saddle point problems*, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., vol. 13 of Math. Ind., Springer, Berlin, 2008, pp. 195–211.

[4] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, 2nd ed., SIAM, Philadelphia, 1994.

[5] V. DOLEAN, P. JOLIVET, AND F. NATAF, *An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation*, SIAM, Philadelphia, 2015.

[6] P. A. FARRELL, A. F. HEGARTY, J. J. H. MILLER, E. O'RIORDAN, AND G. I. SHISHKIN, *Robust Computational Techniques for Boundary Layers*, Chapman & Hall, Boca Raton, 2000.

[7] P. A. FARRELL AND G. I. SHISHKIN, *Schwartz methods for singularly perturbed convection-diffusion problems*, in Analytical and Numerical Methods for Convection-Dominated and Singularly Perturbed Problems, L. G. Vulkov, J. J. H. Miller, and G. I. Shishkin, eds., Nova Science Publishers, Huntington, NY, 2000, pp. 33–42.

[8] M. J. GANDER, *Schwarz methods over the course of time*, Electron. Trans. Numer. Anal., 31 (2008), pp. 228–255.
  http://etna.ricam.oeaw.ac.at/vol.31.2008/pp228-255.dir/pp228-255.pdf

[9] M. J. GANDER, S. LOISEL, AND D. B. SZYLD, *An optimal block iterative method and preconditioner for banded matrices with applications to PDEs on irregular domains*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 653–680.

[10] W. HACKBUSCH, *Elliptic Differential Equations*, Springer, Berlin, 2010.

[11] G. A. A. KAHOU, E. KAMGNIA, AND B. PHILIPPE, *An explicit formulation of the multiplicative Schwarz preconditioner*, Appl. Numer. Math., 57 (2007), pp. 1197–1213.

[12] N. KOPTEVA AND E. O'RIORDAN, *Shishkin meshes in the numerical solution of singularly perturbed differential equations*, Int. J. Numer. Anal. Model., 7 (2010), pp. 393–415.

[13] P. J. LANZKRON, D. J. ROSE, AND D. B. SZYLD, *Convergence of nested classical iterative methods for linear systems*, Numer. Math., 58 (1991), pp. 685–702.

[14] T. LINSS AND M. STYNES, *Numerical methods on Shishkin meshes for linear convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 3527–3542.

[15] H. MACMULLEN, J. J. H. MILLER, E. O'RIORDAN, AND G. I. SHISHKIN, *Schwarz iterative method for convection-diffusion problems with boundary layers*, in Analytical and Numerical Methods for Convection-Dominated and Singularly Perturbed Problems, L. G. Vulkov, J. J. H. Miller, and G. I. Shishkin, eds., Nova Science Publishers, Huntington, NY, 2000, pp. 213–218.

[16] ———, *A second-order parameter-uniform overlapping Schwarz method for reaction-diffusion problems with boundary layers*, J. Comput. Appl. Math., 130 (2001), pp. 231–244.

[17] H. MACMULLEN, E. O'RIORDAN, AND G. I. SHISHKIN, *Schwarz methods for convection-diffusion problems*, in Numerical Analysis and its Applications (Rousse, 2000), L. Vulkov, J. Waśniewski, and P. Yalamov, eds., vol. 1988 of Lecture Notes in Comput. Sci., Springer, Berlin, 2001, pp. 544–551.

[18] ———, *The convergence of classical Schwarz methods applied to convection-diffusion problems with regular boundary layers*, Appl. Numer. Math., 43 (2002), pp. 297–313.

[19] J. J. H. MILLER, E. O'RIORDAN, AND G. I. SHISHKIN, *Fitted Numerical Methods for Singular Perturbation Problems: Error Estimates in the Maximum Norm for Linear Problems in One and Two Dimensions*, Revised Edition, World Scientific Publishing, Singapore, 2012.

[20] R. NABBEN, *Two-sided bounds on the inverses of diagonally dominant tridiagonal matrices*, Linear Algebra Appl., 287 (1999), pp. 289–305.

[21] H.-G. ROOS, *A note on the conditioning of upwind schemes on Shishkin meshes*, IMA J. Numer. Anal., 16 (1996), pp. 529–538.

[22] M. STYNES, *Steady-state convection-diffusion problems*, Acta Numer., 14 (2005), pp. 445–508.

[23] D. B. SZYLD, *The many proofs of an identity on the norm of oblique projections*, Numer. Algorithms, 42 (2006), pp. 309–323.

[24] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods—Algorithms and Theory*, Springer, Berlin, 2005.

[25] R. A. USMANI, *Inversion of Jacobi's tridiagonal matrix*, Comput. Math. Appl., 27 (1994), pp. 59–66.