MULTIPLICATIVE SCHWARZ METHODS FOR PARABOLIC PROBLEMS

XIAO-CHUAN CAI *

Abstract. The class of multiplicative Schwarz methods originated from the classical Schwarz alternating method. It has been shown to be one of the most powerful methods for solving finite element or finite difference elliptic problems. In this paper, we extend these methods to a class of singularly perturbed equations, that are encountered when discretizing parabolic equations by implicit methods such as the backward Euler's or Crank-Nicolson's schemes. We discuss several algorithms, including one-level, two-level and multilevel overlapping methods and study how the convergence rates depend on the time and space discretization parameters, as well as subspace decomposition parameters such as the number of subregions and the number of levels to which the finite element space is decomposed. We show that in the presence of a fine enough coarse mesh space the algorithms are optimal for both symmetric and nonsymmetric problems, i.e. the convergence rates are independent of all these parameters in both two and three dimensions. If the coarse mesh space is dropped, the algorithms are still optimal but only if the time step and the coarse mesh size satisfy certain relationship.

Key words. overlapping domain decomposition, multilevel, optimal convergence, parabolic equation, finite elements

AMS(MOS) subject classifications. 65N30, 65F10, 65N55

1. Introduction. In a pioneer work of Lions [20] on domain decomposition methods, the classical Schwarz alternating algorithm [25] was extended successfully to a class of parabolic equations. The basic idea is to divide the region into several overlapping subregions and then to solve the parabolic problem in each subregion alternatively with boundary information from the neighboring subregions. In this paper, we further extend this idea to the case of many overlapping subregions, as well as to many levels of overlapping subregions. In this case, at each time step many independent subproblems are being created. Thus parallel computers can be used more efficiently.

The focus of this paper is the study of the convergence rates and their dependence on the time and space discretization parameters, on the subspace decomposition parameters, i.e. the number of subregions and the number of levels into which the finite element space is decomposed. Most of the techniques that we shall use in this paper are borrowed from the abstract theory of the additive and multiplicative Schwarz methods for elliptic equations; see e.g. [3, 7, 8, 12, 13] and references therein. The additive version of some of the algorithms of this paper have previously been considered by the author in [4]. With a coarse mesh space, we show that under basically the same assumptions as for the additive Schwarz methods [4], the multiplicative algorithms converge with optimal rates that are independent of the time and space discretization parameters, the number of subregions and the number of levels into which the space is decomposed. In contrast to the Schwarz methods for elliptic problems, in which the coarse mesh space plays an essential role for the optimality, we prove that under the assumption that τ/H^2 is reasonably small, the algorithms remain optimal even if the coarse mesh space is eliminated. Here τ is on the order of the time step size and H is the diameter of the largest substructure.

Some computational aspects of the algorithms have been studied extensively in the context of solving elliptic problems; see e.g. the recent paper of Cai, Gropp and Keyes [6]. Other domain decomposition based algorithms that deal with parabolic problems have recently been developed; see e.g. Bramble, Pasciak and Schatz [2], Dawson and Dupont [9], Dryja [10], Ewing, Lazarov, Pasciak,

* Current address: Department of Computer Science, University of Colorado, Boulder, CO 80309, cai@cs.colorado.edu. The research was supported in part by the National Science Foundation and the Kentucky EPSCoR Program under grant STI-9108764.

and Vassilevski [16], and Kuznetsov [19] and references therein.

The paper is organized as follows. In the remainder of this section, we present the continuous and discrete parabolic equations, and some of their basic properties. In section 2, we discuss four overlapping decompositions of the finite element space, including a one-level decomposition, a two-level decomposition and two multilevel decompositions, and prove the uniformly boundedness of these decompositions. Four algorithms are introduced in Section 3 and their convergence rates are also analyzed. Finally, in section 4, we apply these algorithms to some parabolic problems, including a time dependent convection-diffusion equation. Throughout this paper, c and C, with or without subscripts, denote generic, strictly positive constants. Their values may be different at different occurrences, but are independent of the time and space discretization parameters, as well as the subspace decomposition parameters, that will be introduced later as we move along.

Let $\Omega \subset R^d$ be a bounded polygonal region and $d = 2$, or 3. We are interested in the finite element solution of the following problem: Find $u \in H_0^1(\Omega)$, such that

$$d_\tau(u, \phi) \equiv \tau b(u, \phi) + (u, \phi) = (f, \phi), \forall \phi \in H_0^1(\Omega). \tag{1}$$

Here $\tau > 0$ is a small number, which will be specified in Section 4 of this paper. The bilinear form $b(\cdot, \cdot)$ is defined as

$$b(u, v) = \sum_{i,j=1}^{d} \int_\Omega a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx + \sum_{i=1}^{d} \int_\Omega b_i(x) \frac{\partial u}{\partial x_i} v dx \ + \int_\Omega c(x) u v dx \tag{2}$$

and $(\cdot, \cdot)$ is the usual $L^2(\Omega)$ inner product. We assume that all coefficients are sufficiently smooth and the matrix $\{a_{ij}(x)\}$ is symmetric and uniformly positive definite. We also assume that the bilinear form $b(\cdot, \cdot)$ is bounded and positive definite, though not necessarily self-adjoint, i.e.,
(a) $|b(u, v)| \leq C\|u\|_{H_0^1(\Omega)}\|v\|_{H_0^1(\Omega)}, \forall u, v \in H_0^1(\Omega)$.
(b) $b(u, u) \geq c\|u\|^2_{H_0^1(\Omega)}, \ \ \forall u \in H_0^1(\Omega)$.
We use an $H_0^1(\Omega)$ equivalent norm, denoted by $\|\cdot\|_a$, defined by $a(u, v) = 1/2(b(u, v) + b(v, u))$. In addition, we define the bilinear forms $a_\tau(u, v) = 1/2(d_\tau(u, v) + d_\tau(v, u))$, which is symmetric positive definite, and $n_\tau(u, v) = 1/2(d_\tau(u, v) - d_\tau(v, u))$, which is skewsymmetric. It is not difficult to see that the norm $\|\cdot\|_{a_\tau}$, defined by $a_\tau(\cdot, \cdot)$, is equivalent to the norm $(\|\cdot\|^2_{L^2(\Omega)} + \tau\|\cdot\|^2_{H_0^1(\Omega)})^{1/2}$. As an immediate consequence of the above assumptions, we have that the bilinear form $d_\tau(\cdot, \cdot)$ is bounded and positive definite in the $\|\cdot\|_{a_\tau}$ norm and that $n_\tau(\cdot, \cdot)$ is bounded: There exists a constant C, such that

$$| n_\tau(u, v) | \leq C\tau\|u\|_{H_0^1(\Omega)}\|v\|_{L^2(\Omega)}, \ \ \forall u, v \ \in H_0^1(\Omega).$$

It can be shown, cf Grisvard [17] and Nečas [21], that $b(u, v)$ is $H^{1+\alpha}(\Omega)$-regular: For any $g \in L^2(\Omega)$, there exists a unique $u \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ such that

$$b(v, u) \ = \ (g, v), \ \ \forall v \in H_0^1(\Omega)$$

and

$$| \ u \ |_{H^{1+\alpha}(\Omega)} \leq C\|g\|_{L^2(\Omega)},$$

where α is at least $1/2$. We note that the convergence rates of some algorithms that are developed in this paper depend on the value α.

Let V^h be the usual piecewise linear conforming finite element subspace of $H_0^1(\Omega)$. The standard Galerkin approximation of (1) can then be defined by the following problem: Find $u_h^* \in V^h$, such that

$$d_\tau(u_h^*, \phi_h) = (f, \phi_h), \ \ \forall \phi_h \in V^h. \tag{3}$$

In the next section, we shall formally introduce the space V^h and then decompose it into the sum of certain subspaces. Related multiplicative Schwarz methods will then be introduced to solve the equation (3). The main purpose of this paper is the study of the convergence rates of these algorithms.

2. Overlapping Decompositions and the Stability Analysis.

In this section, we describe some uniformly overlapping subspace decompositions previously introduced by Dryja and Widlund [11, 14] for solving elliptic problems. These are the one-level, two-level and multilevel overlapping decompositions. We will also show that these decompositions are uniformly, or nearly uniformly, bounded in the τ dependent norm $\|\cdot\|_{a_\tau}$. In the multilevel case, the uniformity is also with respect to the number of levels.

2.1. One- and Two-level Decompositions.

Both decompositions have been discussed in [4], and the boundedness of the one-level decomposition is discussed only for the case $d = 2$. Here, we cover both $d = 2$ and 3.

Let $\{\Omega_i\}_{i=1}^M$ be a shape regular finite element triangulation, or a coarse mesh, of Ω and Ω_i has diameter of order H. In our second step, we further divide each substructure Ω_i into smaller simplices with diameter of order h. We assume that the resulting elements form a shape regular finite element subdivision of Ω. We call this the fine mesh or the h-level subdivision of Ω with mesh parameter h. Let us denote by $V^H \subset H_0^1(\Omega)$ and $V^h \subset H_0^1(\Omega)$ the continuous, piecewise linear function spaces, with zero trace on $\partial\Omega$, over the H-level and h-level subdivisions of Ω, respectively.

To obtain an overlapping decomposition, we extend each subregion Ω_i to a larger region Ω_i^{ext} such that $\Omega_i \subset \Omega_i^{ext} \subset \Omega$. Moreover, we assume that distance$(\partial\Omega_i^{ext} \cap \Omega, \partial\Omega_i \cap \Omega) \geq cH, \quad \forall i$. We suppose that $\partial\Omega_i^{ext}$ does not cut through any h-level elements.

Associated with the decomposition $\{\Omega_i^{ext}\}_{i=1}^M$, we define a undirected graph in which the nodes represent the extended subdomains and the edges intersections of the extended subdomains. This graph can be colored, using colors $1, \cdots, J$, such that no adjacent nodes have the same color. We regard the union of all subdomains with the same color as one subdomain (which is of course not simply connected), and denote them by $\Omega_1^{'}, \cdots, \Omega_J^{'}$. We also denote $\Omega_0^{'} = \Omega$ which has color 0. It is important to note that $J+1$, the number of colors, can be made to be independent of M, the number of substructures.

For each $\Omega_j^{'}$, $1 \leq j \leq J$, we define $V_j = \{v \in V^h | \ v(x) = 0, \, x \bar{\in} \Omega_j^{'}\} \subset V^h$. We also use the subspace $V_0 = V^H$. It is easy to see that we have two decompositions of V^h, i.e.,

$$V^h \ = \ V_1 + \cdots + V_J \tag{4}$$

and

$$V^h \ = \ V_0 + V_1 + \cdots + V_J, \tag{5}$$

which shall be referred as the one-level and two-level uniformly overlapping decompositions of V^h, or as the decomposition without and with the coarse mesh space, respectively.

It was proved in [4] that the two-level decomposition (5) is uniformly bounded in the $\|\cdot\|_{a_\tau}$ norm in both two- and three-dimensional spaces.

Lemma 2.1. *There exists a constant C_0, such that for any $v \in V^h$, there exist $v_i \in V_i$, $v = v_0 + \cdots + v_J$ and*

$$\|v_0\|_{a_\tau}^2 + \sum_{i=1}^J \|v_i\|_{a_\tau}^2 \leq C_0^2 \|v\|_{a_\tau}^2, \ \ \forall v \in V^h. \tag{6}$$

Here the constant C_0 is independent of τ, h and H.

If we drop the coarse mesh space V_0 from (5) and v_0 from the decomposition, the following estimate holds for both $d=2$ and 3. The case $d=2$ was concerned in [4].

LEMMA 2.2. *There exists a constant C_τ, such that for any $v \in V^h$, there exist $v_i \in V_i$, $v = v_1 + \cdots + v_J$ and*

$$\sum_{i=1}^{J} \|v_i\|_{a_\tau}^2 \le C(1+\frac{\tau}{H^2})\|v\|_{a_\tau}^2 \equiv C_\tau^2 \|v\|_{a_\tau}^2, \quad \forall v \in V^h, \tag{7}$$

where $C>0$ is independent of τ, h and H.

Proof. Let $\{\theta_i\}_{i=1}^J$ be a partition of unity and θ_i belongs to $C_0^\infty(\Omega_i^{'} \cap \Omega)$. It can be chosen so that $|\nabla \theta_i|$ is bounded by C/H. Let I_h be the piecewise linear interpolation operator which uses the function values at the h-level nodes. For any $v \in V^h$, we let $v_i = I_h(\theta_i v) \in V_i$. For each subregion $\Omega_i^{'}$, it is well-known, in the sense of equivalent norms, that

$$|v_i|_{H_0^1(\Omega_i')}^2 = \sum ((\theta_i v)(x_l) - (\theta_i v)(x_m))^2 h^{d-2}, \tag{8}$$

where the sum is taken over all adjacent pairs of h-level nodal points in $\Omega_i^{'}$. Let $K \subset \Omega_i^{'}$ be a single h-level triangle and x_l, $x_m \in \bar{K}$ be two of its vertices. Let $\bar{\theta}_{ilm} = (\theta_i(x_l) + \theta_i(x_m))/2$. We then have

$$(\theta_i v)(x_l) - (\theta_i v)(x_m) = (\theta_i(x_l) - \bar{\theta}_{ilm}) v(x_l) - (\theta_i(x_m) - \bar{\theta}_{ilm}) v(x_m) + \bar{\theta}_{ilm}(v(x_l) - v(x_m)) \ ,$$

which can be bounded from above by

$$C\left(\frac{h}{H} \max_{x \in \bar{K}} \{|v(x)|\}\right) + |v(x_l) - v(x_m)|.$$

By squaring this estimate, using the inequality $ab \le 1/2(a^2+b^2)$, summing over all $K \subset \Omega_i^{'}$ and using (8), we obtain

$$\begin{aligned} |v_i|_{H_0^1(\Omega_i')}^2 &\le C\left(H^{-2} \sum_{K \subset \Omega_i'} \max_{x \in \bar{K}} \{|v(x)|^2\} \cdot h^d \ + \ |v|_{H^1(\Omega_i')}^2\right) \\ &\le C\left(H^{-2}\|v\|_{L^2(\Omega_i')}^2 + |v|_{H^1(\Omega_i')}^2\right). \end{aligned}$$

Therefore,

$$\sum_{i=1}^{J} \|v_i\|_{H_0^1(\Omega_i')}^2 \le C(H^{-2}\|v\|_{L^2(\Omega)}^2 + |v|_{H_0^1(\Omega)}^2) \le C H^{-2} \|v\|_{H_0^1(\Omega)}^2. \tag{9}$$

We refer to Cai [4] for the L^2 estimate, i.e.,

$$\sum_{i=1}^{J} \|v_i\|_{L^2(\Omega_i')}^2 \le C\|v\|_{L^2(\Omega)}^2. \tag{10}$$

The proof follows immediately from the estimates (9) and (10). □

REMARK 2.1. *It is known that in the elliptic case, which corresponds to the use of the $\|\cdot\|_a$ norm in the estimate, if the coarse mesh space is dropped, a factor of $1/H^2$ would appear in the estimate and this makes this decomposition not so useful. However, as shown above, in the parabolic case, only a factor of τ/H^2 appears in the estimate and τ is usually in the order of the time step size.*

2.2. Multilevel Decompositions. Following Dryja and Widlund [14] and Zhang [27], we let $\{\mathcal{T}^l\}_{l=0}^L$ be a nested sequence of triangulations of Ω, i.e. $\mathcal{T}^0 = \{\tau_i^0\}_{i=1}^{N_0}$ is the initial coarse triangulation and $\mathcal{T}^l = \{\tau_i^l\}(l = 1, \cdots, L)$ is defined by dividing each triangle of $\mathcal{T}^{l-1}$ into several triangles. We assume that all the triangulations are shape regular. Let $h_i^l = \text{diam}(\tau_i^l)$, $h_l = \max_i\{h_i^l\}$, $H = \max_i h_i^0$ and $h = h_L$. We also assume that there exists $0 < r < 1$, such that h_l is proportional to Hr^l, for $l = 0, \cdots, L$. Let V^l be the usual conforming finite element space of continuous piecewise linear functions associated with the triangulation $\mathcal{T}^l$.

We construct L sets of overlapping subdomains $\{\hat{\Omega}_i^l\}_{i=1}^{J_l}, l = 1, 2, \cdots, L$; i.e. for each $1 \leq l \leq L$, we have

$$\Omega = \cup_{i=1}^{J_l} \hat{\Omega}_i^l.$$

Subdomains $\hat{\Omega}_i^l$ are defined as follows. Each triangle $\tau_i^{l-1}, i = 1, \cdots, N_l, l = 1, \cdots, L$, is extended to a larger region $\hat{\tau}_i^{l-1}$ so that

$$ch_i^{l-1} \leq \text{dist}(\partial\hat{\tau}_i^{l-1} \cap \Omega, \partial\tau_i^{l-1} \cap \Omega) \leq Ch_i^{l-1},$$

aligning $\partial\hat{\tau}_i^{l-1}$ with the boundaries of level l triangles. For each l, we color the $\hat{\tau}_i^{l-1}$ by using J_l colors, such that all substructures of the same color are disjoint. Here J_l is a fixed constant depends only on the triangulations. On each level l, we group the extended subregions $\hat{\tau}_i^{j-1}$ by colors, and obtain J_l sets of subregions. We denote by $\hat{\Omega}_i^l$ as the union of l-level extended subregions that share the same color $1 \leq i \leq J_l$. We denote $J = \max_l J_l$.

We define $V_i^l = \{v \in V^l |\ v(x) = 0,\ x\bar{\in}\hat{\Omega}_i^l\} \subset V^l$. Let $J_0 = 1$ and $V^0 = V_1^0 = V^H$. Thus, our finite element space $V^h = V^L$ can be represented as

$$V^L = V^0 + \sum_{l=1}^{L} V^l = V^0 + \sum_{l=1}^{L}\sum_{i=1}^{J_l} V_i^l \tag{11}$$

or

$$V^L = \sum_{l=1}^{L} V^l = \sum_{l=1}^{L}\sum_{i=1}^{J_l} V_i^l. \tag{12}$$

These two decompositions are referred to as the multilevel decomposition with and without a coarse mesh space, respectively. We note that the main difference between these two decompositions is that in (11) the coarsest mesh space is not decomposed into local subspaces. This works well if the degrees of freedoms involved in the coarse mesh problem are few. However, this is not always satisfied, especially for nonsymmetric problems where the coarse mesh needs to be sufficiently fine. In the latter case, it is desirable to further decompose the coarse mesh problem and therefore (12) is sometimes more useful.

It is known that the decomposition with the coarse mesh (11) is uniformly bounded in the $\|\cdot\|_a$ norm, i.e., for any $v \in V^h$, there exist $v_i^l \in V_i^l$ such that

$$v \ = \ v^0 + \sum_{l=1}^{L}\sum_{i=1}^{J_l} v_i^l.$$

Moreover, there exists a constant $C > 0$, independent of the parameters h, H and L, such that

$$\|v^0\|_a^2 + \sum_{l=1}^{L}\sum_{i=1}^{J_l} \|v_i^l\|_a^2 \leq C\|v\|_a^2, \ \ \forall v \in V^h.$$

We refer to Bramble and Pasciak[1], Dryja and Widlund [14], Oswald [22] and Zhang [27] for the analysis and discussions of this bound.

In order to prove that the decomposition (11) is also uniformly bounded in the τ dependent norm $\|\cdot\|_{a_\tau}$, we need only to show that the same decomposition is uniformly bounded in the $L^2(\Omega)$ norm. Let us define Q^l as the usual L^2 projection

$$(Q^l v, \phi) = (v, \phi) \quad \forall v \in V^h,\ \forall \phi \in V^l$$

for $l = 0, \cdots, L$. For any given $v \in V^h$, let $v^0 = Q^0 v \in V^0$ and $v^l = Q^l v - Q^{l-1} v \in V^l$ for $l \geq 1$. It is easy to verify that

$$v = v^0 + v^1 + \cdots + v^l.$$

By using the fact that, for any $v \in V^h$ and $l \geq 1$,

$$\|(Q^l - Q^{l-1})v\|^2_{L^2(\Omega)} = (v, (Q^l - Q^{l-1})v)$$

we establish the identity

$$\|Q^0 v\|^2_{L^2(\Omega)} + \sum_{l=1}^{L} \|(Q^l - Q^{l-1})v\|^2_{L^2(\Omega)} = \|v\|^2_{L^2(\Omega)}. \tag{13}$$

For each level $l \geq 1$, we define a partition of unity $\{\theta_i^l\}_{i=1}^{J_l}$ with $\theta_i^l \in H_0^1(\hat{\Omega}_i^l) \cap C_0^1(\hat{\Omega}_i^l \cap \Omega)$ such that $\sum_i \theta_i^l = 1$, $0 \leq \theta_i^l \leq 1$ and $|\nabla \theta_i^l| \leq C/h_{l-1}$. Now, each $v^l = (Q^l - Q^{l-1})v$ can be further decomposed as

$$v^l = \sum_{i=1}^{J_l} v_i^l,$$

where $v_i^l = I_{h_l}(\theta_i^l v^l) \in V_i^l$ and I_{h_l} is the piecewise linear interpolation operator from V^h to V^l. By using the second part of the proof of Lemma 4 of [4], we have that

$$\sum_{i=1}^{J_l} \|v_i^l\|^2_{L^2(\Omega)} \leq C \|v^l\|^2_{L^2(\Omega)}.$$

Therefore, by combining the above results, we obtain a decomposition

$$v = v^0 + \sum_{l=1}^{L} v^l = v^0 + \sum_{l=1}^{L} \sum_{i=1}^{J_l} v_i^l,$$

which is uniformly bounded in the $L^2(\Omega)$ norm, i.e.,

$$\|v^0\|^2_{L^2(\Omega)} + \sum_{l=1}^{L} \sum_{i=1}^{J_l} \|v_i^l\|^2_{L^2(\Omega)} \leq C \|v\|^2_{L^2(\Omega)}, \quad \forall v \in V^h,$$

where the constant $C > 0$ is independent of h, H and L. As a consequence, we proved that

LEMMA 2.3. *For any $v \in V^h$, there exist $v_i^l \in V_i^l$ such that*

$$v \; = \; v^0 + \sum_{l=1}^{L} \sum_{i=1}^{J_l} v_i^l$$

and moreover, there exists a constant $C_0^{\rm M}$, independent of the parameters h, H, L and τ, such that

$$\|v^0\|_{a_\tau}^2 + \sum_{l=1}^{L}\sum_{i=1}^{J_l} \|v_i^l\|_{a_\tau}^2 \le (C_0^{\rm M})^2 \|v\|_{a_\tau}^2, \quad \forall v \in V^h.$$

We now discuss the case when the coarse mesh space V^0 is dropped from the decomposition. In the analysis, we shall use the well-known approximation and boundedness properties of the operators $\{Q^l\}$, see e.g. Xu [26], i.e., for any $v \in V^h$,

$$\|(Q^l - Q^{l-1})v\|_{L^2(\Omega)}^2 \le C h_l^2 a(v,v), \quad l = 1, \cdots, L \tag{14}$$

and

$$\|Q^l v\|_a \le C \|v\|_a, \quad l = 1, \cdots, L. \tag{15}$$

Lemma 2.4. *For any $v \in V^h$, there exist $v_i^l \in V_i^l$ such that $v = \sum_{l=1}^{L}\sum_{i=1}^{J_l} v_i^l$ and*

$$\sum_{l=1}^{L}\sum_{i=1}^{J_l} \|v_i^l\|_{a_\tau}^2 \le C(1 + \frac{\tau}{H^2})\|v\|_{a_\tau}^2 \equiv (C_\tau^{\rm M})^2 \|v\|_{a_\tau}^2 \quad \forall v \in V^h, \tag{16}$$

where $C > 0$ is independent of the parameters h, H, L and τ.

Proof. For any given $v \in V^h$, we construct the multilevel overlapping decomposition as follows. Let $v^1 = Q^1 v$, $v^l = (Q^l - Q^{l-1})v$ for $l \ge 2$ and $v_i^l = I_{h_l}(\theta_i^l v^l)$ for $l = 1, \cdots, L$, $i = 1, \cdots, J_l$. These operators are linear, therefore

$$v = \sum_{l=1}^{L}\sum_{i=1}^{J_l} v_i^l.$$

We do the $H^1(\Omega)$ norm and $L^2(\Omega)$ norm estimates separately. For the $L^2(\Omega)$ estimate, by using the same arguments made for (13), we obtain

$$\sum_{l=1}^{L} \|v^l\|_{L^2(\Omega)}^2 = \|v\|_{L^2(\Omega)}^2, \tag{17}$$

which, combined with the proof of Lemma 4 of Cai [4], implies that

$$\sum_{l=1}^{L}\sum_{i=1}^{J_l} \|v_i^l\|_{L^2(\Omega)}^2 \le C \|v\|_{L^2(\Omega)}^2. \tag{18}$$

We next examine the boundedness of the same decomposition described above in the $\|\cdot\|_a$ norm. We use the fact that

$$\sum_{l=1}^{L} \|v^l\|_a^2 \le C \|v\|_a^2, \quad \forall v \in V^h, \tag{19}$$

where the constant $C > 0$ is independent of the parameters h, H and L. We refer to Bramble and Pasciak [1], Oswald [22] and Zhang [27] for the proofs and discussions.

Because of the estimate (6) of Dryja and Widlund [11], at each level $l \ge 2$, we have

$$|v_i^l|_{H_0^1(\hat\Omega_i^l)}^2 \le C\left(|v^l|_{H_0^1(\hat\Omega_i^l)}^2 + h_{l-1}^{-2}\|v^l\|_{L^2(\hat\Omega_i^l)}^2\right).$$

Here we used the fact that $\hat{\Omega}_i^l$ is the union of substructures of diameters on the order of h_{l-1}. Thus, by summing over $i = 1, \cdots, J_l$, and by using the approximation property (14), i.e., $\|v^l\|^2_{L^2(\Omega)} \leq Ch^2_{l-1}\|v^l\|^2_a$, we obtain

$$\sum_{i=1}^{J_l} \|v_i^l\|_a^2 \leq C\left(\|v^l\|_a^2 + h_{l-1}^{-2}\|v^l\|^2_{L^2(\Omega)}\right) \leq C\|v^l\|_a^2. \tag{20}$$

In the case $l = 1$, the desired estimate is known, see Lemma 8 of Cai [4], i.e.,

$$\sum_{i=1}^{J_1} \|v_i^1\|_a^2 \leq CH^{-2}\|v^1\|_a^2. \tag{21}$$

Therefore, the $\|\cdot\|_a$ estimate is accomplished by first adding (20), for $l = 2, \cdots, L$, and (21), and then using (19),

$$\sum_{l=1}^{L}\sum_{i=1}^{J_l} \|v_i^l\|_a^2 \leq CH^{-2}\sum_{l=1}^{L}\|v^l\|_a^2 \leq CH^{-2}\|v\|_a^2. \tag{22}$$

The proof of this lemma is completed by combining the results of the $L^2(\Omega)$ estimate (18) and the $H^1(\Omega)$ estimate (22). □

3. Schwarz Algorithms and Convergence Rates Analysis. In this section, we briefly describe the multiplicative Schwarz algorithms based on the various decompositions of V^h discussed in the previous section. The convergence rate of each algorithm is analyzed with the general Schwarz theory recently developed by Cai and Widlund [8].

3.1. Multiplicative Schwarz Algorithms. Assuming that we have a set of triplets of $\{W_i, S_i, g_i \mid i = 1, \cdots, m\}$, where W_i are some subspaces of a normed linear space W, S_i some mappings from W to W_i and $g_i \in W_i$, the multiplicative Schwarz algorithm can be described as follows.

ALGORITHM 3.1 (MULTIPLICATIVE SCHWARZ $\{W_i, S_i, g_i\}$).

Given an initial guess $u^0 \in W$ *:*
For $n = 0, 1, \cdots, n_{max}$*;*
  For $i = 1, \cdots, m$*;*
  $u^{n+\frac{i}{m}} = u^{n+\frac{i-1}{m}} + (g_i - S_i u^{n+\frac{i-1}{m}})$.

It is not difficult to see that with the *error propagation* operator defined by

$$E = (I - S_m)\cdots(I - S_1),$$

and

$$g = (I - S_m)\cdots(I - S_2)g_1 + (I - S_M)\cdots(I - S_3)g_2 + \cdots + g_m,$$

then $u^{n+1} = Eu^n + g$. The convergence rate of the multiplicative Schwarz algorithm is thus determined by the spectral radus of the operator E. In an application, Algorithm 3.1 can also be used as a preconditioner for another CG-type iterative methods, such as CG for symmetric positive definite problems, and GMRES [24, 15] or FGMRES [23] for other cases. When it is used as a preconditioner, the actual systems solved is the *transformed system* $(I - E)u = g$. We refer to [5, 6, 7, 8] for discussion and comments on the numerical implementations.

In the rest of the paper, we take $W = V^h$ and W_i as one of the V_i or V_i^l defined in the previous section. Operators for subspaces V_i are defined as follows: For each $v \in V^h$, a unique $P_i v \in V_i$ is defined as the finite element solution of

$$d_\tau(P_i v, \phi) = d_\tau(v, \phi), \ \forall \phi \in V_i.$$

Similarly, we can define P_i^l on V_i^l. We denote $g_i = P_i u_h^*$ and $g_i^l = P_i^l u_h^*$. We note that the g_i can be computed, without knowing the function u_h^* itself, by solving the finite element problems

$$d_\tau(g_i, \phi) = (f, \phi), \quad \forall \phi \in V_i. \tag{23}$$

Similarly, g_i^l can also be computed without knowing u_h^*.

Algorithm 3.2 (One-level multiplicative Schwarz). *Run Algorithm 3.1 with an initial guess* $u^0 \in V^h$ *and* $\{V_i, P_i, g_i \mid i = 1, \cdots, J\}$.

Algorithm 3.3 (Two-level multiplicative Schwarz). *Run Algorithm 3.1 with an initial guess* $u^0 \in V^h$ *and* $\{V_0, P_0, g_0, \text{ and } V_i, P_i, g_i \mid i = 1, \cdots, J\}$.

Algorithm 3.4 (Multilevel multiplicative Schwarz without a coarse mesh space). *Run Algorithm 3.1 with an initial guess* $u^0 \in V^h$ *and*

$$\{V_i^l, P_i^l, g_i^l \mid l = 1, \cdots, L, \ i = 1, \cdots, J_l\}.$$

Algorithm 3.5 (Multilevel multiplicative Schwarz with a coarse mesh space). *Run Algorithm 3.1 with an initial guess* $u^0 \in V^h$ *and*

$$\{V^0, P_0, g_0, \text{ and } V_i^l, P_i^l, g_i^l \mid l = 1, \cdots, L, \ i = 1, \cdots, J_l\}.$$

The convergence of these algorithms will be analyzed in the next subsection. We note that the Algorithm 3.1 is a purely sequential algorithm, however, Algorithms 3.2 - 3.5 can be made highly parallel. This is due to the fact that each V_i (or V_i^l), except V_0 (or V^0), is the sum of several subspaces that are mutually orthogonal. Therefore the subproblem defined on V_i can be regarded as a set of independent sub-subproblems that can be solved in parallel. In Algorithms 3.2 and 3.4, the bottleneck step with V_0 (or V^0) is removed as compared with their counterpart Algorithms 3.3 and 3.5. We show, in the next subsection, that the removal of the coarse mesh space does not degenerate by much the convergence rates for the class of parabolic problems under certain circumstances, although this is known to be not true for elliptic problems.

3.2. Convergence Rates Analysis. Let $E_{A3.3}$ be the error propagation operator for Algorithm 3.3, etc., we estimate the norm of these operators by using a theorem of Cai and Widlund [8]. The proof of this theorem is technical; interested readers are referred to [8] for details. To apply this theorem we need only to verify certain properties of the subspaces related to the operator.

Theorem 3.1 (Cai and Widlund). *Let* W *be a Hilbert space with inner product* $\nu(\cdot,\cdot)$ *and norm* $\|\cdot\|_\nu = \nu(\cdot,\cdot)^{1/2}$. *Suppose that* W *has a decomposition*

$$W = W_1 + \cdots + W_m, \tag{24}$$

where $W_1, \cdots, W_m$ *are subspaces of* W. *Let the matrix* $\mathcal{E} = \{\varepsilon_{ij}\}$ *be defined by the strengthened Cauchy-Schwarz coefficients, where* ε_{ij} *is the smallest constant for which*

$$|\nu(w_i, w_j)| \leq \varepsilon_{ij} \|w_i\|_\nu \|w_j\|_\nu \ , \forall w_i \in W_i \ , \forall w_j \in W_j \ , i, j = 1, \cdots, m \ , \tag{25}$$

holds. We assume that there are operators $T_i : W \to W_i$ *that satisfy the following assumptions*
(i) There exists a constant $\gamma > 0$ *and parameters* $\delta_i \geq 0$, *such that* $\sum_{i=1}^{m} \delta_i$ *can be made sufficiently small, such that*

$$T_i + T_i^T - T_i^T T_i \geq \gamma T_i^T T_i - \delta_i I. \tag{26}$$

(ii) There exists a constant $\bar{C}_0 > 0$, *such that*

$$\sum_{i=1}^{m} T_i^T T_i \geq \bar{C}_0^{-2} I \ . \tag{27}$$

Here T_i^T *is the adjoint operator of* T_i *given by* $\nu(T_i^T u, v) = \nu(u, T_i v)$. *Then, we have*

$$\|(I - T_m) \cdots (I - T_1)\|_\nu \leq \sqrt{1 - \bar{\gamma} \bar{C}_0^{-2}},$$

where $\bar{\gamma} = C(4(1+\gamma)^{-2} |\mathcal{E}|_{l^1}^2 + (\sum_{i=1}^{m} \delta_i)^2 + 1)^{-1}$ *and* C *is a positive constant.*

In the remainder of this paper, we shall apply this theorem to the Hilbert space V^h equipped with the inner product $a_\tau(\cdot, \cdot)$, the decompositions (4), (5), (11) and (12) discussed in the previous section and the operators P_i or P_i^l. Our main tasks include the verification of Assumptions (i) and (ii), and the estimate of $|\mathcal{E}|_{l^1}$ for the four algorithms. The two assumptions required in Theorem 3.1 will be verified through the next two lemmas.

Lemma 3.1. *For* $i = 0, 1, \cdots, J$,

$$P_i + P_i^T - P_i^T P_i \geq P_i^T P_i - \delta_i I \tag{28}$$

with $\delta_i = O(H)$, *for* $i \geq 1$, *and* $\delta_0 = C H^\alpha \sqrt{H^2/\tau + 1}$. *Here* $C > 0$ *is independent of* h, H *and* τ. *For* $l = 1, \cdots, L$ *and* $i = 1, \cdots, J_l$,

$$P_i^l + (P_i^l)^T - (P_i^l)^T P_i^l \geq (P_i^l)^T P_i^l - \delta_i^l I \tag{29}$$

with $\delta_i^l = O(h_l)$.

The proof for P_0 can be obtained by applying Lemma 6 of [4] and the definition of P_0. A similar approach works for other mapping operators.

Remark 3.1. *If the bilinear form* $b(\cdot, \cdot)$ *is selfadjoint and positive definite, then* $\delta_i = 0$, *for all* $i = 0, \cdots, J$, *and* $\delta_i^l = 0$, *for all* $l = 1, \cdots, L$, $i = 1, \cdots, J_l$. $\gamma = 1.0$ *in all cases.*

Remark 3.2. *For the one- and two-level methods, both summations*

$$\sum_{i=1}^{J} \delta_i \leq CJH \quad \text{and} \quad \delta_0 + \sum_{i=1}^{J} \delta_i \leq C \left(H^\alpha \sqrt{\frac{H^2}{\tau} + 1} + JH \right)$$

can made sufficiently small, therefore Assumption (i) is verified. Note that the fact that J *is independent of* H *is used here.*

REMARK 3.3. *For the two multilevel algorithms, since h_l is proportional to Hr^l, the quantity*

$$\sum_{l=1}^{L}\sum_{i=1}^{J_l}\delta_i^l \le C\frac{r}{1-r}\ JH$$

can be made sufficiently small if H is small. The extra member for the method with coarse mesh $H^\alpha\sqrt{H^2/\tau+1}$ can also be made sufficiently small. Thus, Assumption (i) is verified for the multilevel algorithms.

We next examine the Assumption (ii) for the four algorithms.

LEMMA 3.2. *The following four inequalities hold*

$$\sum_{i=1}^{J} P_i^T P_i \ge C\left(1+\frac{\tau}{H^2}\right)^{-2} I = C_d^{-2}C_\tau^{-2} I, \tag{30}$$

$$P_0^T P_0 + \sum_{i=1}^{J} P_i^T P_i \ge CI = C_d^{-2}C_0^{-2} I, \tag{31}$$

$$\sum_{l=1}^{L}\sum_{i=1}^{J_l}(P_i^l)^T P_i^l \ge C\left(1+\frac{\tau}{H^2}\right)^{-2} I = C_d^{-2}(C_\tau^{\mathrm{M}})^{-2} I, \tag{32}$$

$$P_0^T P_0 + \sum_{l=1}^{L}\sum_{i=1}^{J_l}(P_i^l)^T P_i^l \ge CI = C_d^{-2}(C_0^{\mathrm{M}})^{-2} I, \tag{33}$$

where C_d satisfies $|d_\tau(u,v)| \le C_d\|u\|_{a_\tau}\|v\|_{a_\tau}$ for any $u, v \in V^h$.

Proof. The proofs for these four estimates are essentially the same, and we therefore only provide the proof of the second one. For any $v \in V^h$, by using the decomposition (4), the definition of mapping operators P_i, Cauchy-Schwarz's inequality, the uniformly boundedness of the decomposition (6) in the $\|\cdot\|_{a_\tau}$ norm and the boundedness of $d_\tau(\cdot,\cdot)$, we have

$$\begin{aligned}
\|v\|_{a_\tau}^2 &\le d_\tau(v,v) = \sum_{i=0}^{J} d_\tau(v,v_i) = \sum_{i=0}^{J} d_\tau(P_i v, v_i)\\
&\le C_d\sqrt{\sum_{i=0}^{J}\|P_i v\|_{a_\tau}^2}\sqrt{\sum_{i=0}^{J}\|v_i\|_{a_\tau}^2}\\
&\le C_d C_0\|v\|_{a_\tau}\sqrt{\sum_{i=0}^{J}\|P_i v\|_{a_\tau}^2}.
\end{aligned}$$

Therefore, removing the common term $\|v\|_{a_\tau}$ and squaring both sides, we obtain

$$C_d^{-2}C_0^{-2}\|v\|_{a_\tau}^2 \le \sum_{i=0}^{J}\|P_i v\|_{a_\tau}^2,$$

which is the desired proof. □

REMARK 3.4. *If $d_\tau(\cdot,\cdot)$ is symmetric positive definite, then $C_d = 1$.*

We now discuss the bounds for $|\mathcal{E}|_{l^1}$. The cases for one- and two-level methods are simple. $|\mathcal{E}|_{l^1}$ is bounded by the maximum number of subregions that any given point in Ω belongs to. For example, for the box-decomposition of a rectangular region in the two dimensional space, as in Fig. 1, the subregions can be colored such that $|\mathcal{E}|_{l^1} = 5$. In general, $|\mathcal{E}|_{l^1} \le J$ or $J+1$ for one- or two-level methods. In the multilevel cases, we need an estimate of Zhang [27],

$$a(v_i^{l_1}, v_j^{l_2}) \le C(r^d)^{|l_1-l_2|-1}\|v_i^{l_1}\|_a\|v_j^{l_2}\|_a,\ \text{ for } l_1 \ne l_2, \text{ and } \forall v_i^{l_1} \in V_i^{l_1}, v_j^{l_2} \in V_j^{l_2}$$

and the lemma

LEMMA 3.3. *For* $l_1 \neq l_2$, *and any* $v_i^{l_1} \in V_i^{l_1}, v_j^{l_2} \in V_j^{l_2}$,

$$a_\tau(v_i^{l_1}, v_j^{l_2}) \leq C(r^d)^{|l_1-l_2|-1} \|v_i^{l_1}\|_{a_\tau} \|v_j^{l_2}\|_{a_\tau},$$

where $C > 0$ *is independent of* h, H, L *and* τ.

Proof. It suffics to show that, for $l_2 > l_1$,

$$(v_i^{l_1}, v_j^{l_2}) \leq C(r^d)^{l_2-l_1-1} \|v_i^{l_1}\|_{L^2(\Omega)} \|v_j^{l_2}\|_{L^2(\Omega)}. \tag{34}$$

If the supporting regions $\hat{\Omega}_i^{l_1}$ and $\hat{\Omega}_i^{l_2}$ of $v_i^{l_1}$ and $v_j^{l_2}$ do not intersect, then (34) is trivial, otherwise, we have

$$(v_i^{l_1}, v_j^{l_2}) \leq \|v_i^{l_1}\|_{L^2(\hat{\Omega}_i^{l_2})} \|v_j^{l_2}\|_{L^2(\Omega)}.$$

Let the triangle $K^{l_1} \in \mathcal{T}^{l_1}$ have vertices $y_1, \cdots, y_{d+1}$, and let $x_1, \cdots, x_k$ be all the nodal points of $\mathcal{T}^{l_2}$ in $\hat{\Omega}_i^{l_2}$. We can show that $k \leq CJ_{l_2}(h_{l_2-1}/h_{l_2})^d$. Since $v_i^{l_1}$ is linear in $\hat{\Omega}_i^{l_2} \cap K^{l_1}$, we have

$$\begin{aligned}
\|v_i^{l_1}\|^2_{L^2(\hat{\Omega}_i^{l_2} \cap K^{l_1})} &\leq Ch_{l_2}^d \Big(\sum_{1 \leq \sigma \leq k, x_\sigma \in K^{l_1}} (v_i^{l_1}(x_\sigma))^2 \Big) \\
&\leq Ck\, h_{l_2}^d \max_{x_\sigma \in K^{l_1}} \{|v_i^{l_1}(x_\sigma)|^2\} \\
&\leq Ck\, h_{l_2}^d \sum_{\xi=1}^{d+1} (v_i^{l_1}(y_\xi))^2 \\
&\leq Ck \frac{h_{l_2}^d}{h_{l_1}^d} \|v_i^{l_1}\|^2_{L^2(K^{l_1})} \leq C \left(\frac{h_{l_2-1}}{h_{l_1}} \right)^d \|v_i^{l_1}\|^2_{L^2(K^{l_1})}.
\end{aligned}$$

By using the assumption that h_l is proportional to Hr^l, we obtain

$$\|v_i^{l_1}\|^2_{L^2(\hat{\Omega}_i^{l_2} \cap K^{l_1})} \leq C(r^d)^{l_2-l_1-1} \|v_i^{l_1}\|^2_{L^2(K^{l_1})}.$$

Summing over all $K^{l_1} \in \mathcal{T}^{l_1}$, we have

$$\|v_i^{l_1}\|^2_{L^2(\hat{\Omega}_i^{l_2})} \leq C(r^d)^{l_2-l_1-1} \|v_i^{l_1}\|^2_{L^2(\Omega)},$$

which proves the estimate (34). ◻

Applying Lemma 3.3 and using the fact that $1 + r^d + r^{2d} + \cdots \leq 1/(1-r^d)$, we see that $|\mathcal{E}|_{l^1}$ is uniformly bounded for both decompositions (11) and (12), independent of the number of levels L.

We now summerize the main convergence results for the four algorithms in the next theorem. The proof is a simple consequence of Theorem 3.1 and the lemmas of this section. Let $J = \max_l J_l$.

THEOREM 3.2. *(a) For Algorithm 3.2 , there exists constants* $\tilde{c} > 0$ *and* $c = c(\tilde{c}) > 0$, *such that if* $\max\{H, H^\alpha \sqrt{H^2/\tau + 1}\} \leq \tilde{c}$ *then*

$$\|E_{\text{A}3.2}\|^2_{a_\tau} \leq 1 - \frac{c}{(1+J)^2 C_0^2};$$

(b) For Algorithm 3.3, there exist constants $H_0 > 0$ *and* $c(H_0) > 0$, *such that if* $H \leq H_0$, *then*

$$\|E_{\text{A}3.3}\|^2_{a_\tau} \leq 1 - \frac{c}{(1+J)^2 C_\tau^2}.$$

Recall that $C_\tau^2 = C(1+\tau/H^2)$;
(c) For Algorithm 3.4, there exist constants $H_0 > 0$ *and* $c(H_0) > 0$, *such that if* $H \le H_0$, *then*

$$\|E_{\mathrm{A3.4}}\|_{a_\tau}^2 \le 1 - \frac{c}{(1+J)^2 (C_\tau^{\mathrm{M}})^2}.$$

Recall that $(C_\tau^{\mathrm{M}})^2 = C(1+\tau/H^2)$;
(d) For Algorithm 3.5, there exists constants $\tilde{c} > 0$ *and* $c = c(\tilde{c}) > 0$, *such that if* $\max\{H, H^\alpha \sqrt{H^2/\tau + 1}\} \le \tilde{c}$ *then*

$$\|E_{\mathrm{A3.5}}\|_{a_\tau}^2 \le 1 - \frac{c}{(1+\hat{c}^2)C} = 1 - \frac{c}{(1+\hat{c}^2)(C_0^{\mathrm{M}})^2},$$

where $\hat{c} = J + H^\alpha \sqrt{1 + \frac{\tau}{H^2}} + JH$.
Here the constants c *may be different in different occurrence.*

REMARK 3.5. *As previously remarked, if the bilinear form* $b(\cdot,\cdot)$ *is symmetric positive definite, then the requirement the* H *be sufficiently small is not necessary, i.e.,* $H_0 = +\infty$.

REMARK 3.6. *In the nonsymmetric cases, when the coarse mesh size* $H \le H_0$, *the convergence is insured, generally, the smaller* H *is, the faster the algorithm.*

REMARK 3.7. *These algorithms are often referred as algorithms with exact solvers, because all the subproblems defined in extended subregions, and the coarse mesh problem, are obtained by discretizing the original partial differential operator and then solved exactly by direct method. In practice, this is sometimes not necessary, a different differential operator can be used instead and the linear systems need not to be solved exactly; cf., discussions in [6, 7, 8].*

4. Applications to Parabolic Problems. In this section, we apply the algorithms developed in the previous sections to solve the following parabolic convection-diffusion problem: Find $u(x,t)$, such that

$$\begin{cases} \dfrac{\partial u}{\partial t} + Lu &= f \quad \text{in} \quad \Omega \times [0,T], \\ u(x,t) &= 0 \quad \text{on} \quad \partial\Omega \times [0,T], \\ u(x,0) &= u_0(x) \quad \text{in} \quad \Omega, \end{cases} \tag{35}$$

where $\Omega \subset R^2$. A weak formulation of equation (35) is: Find $u(x,t) \in H_0^1(\Omega)$, $u(x,0) = u_0(x)$ in Ω, such that

$$\left(\frac{\partial u}{\partial t}, \phi\right) + b(u,\phi) = (f,\phi), \ \forall \phi \in H_0^1(\Omega), \ \forall t \in [0,T],$$

where the bilinear form $b(\cdot,\cdot)$ is the same as in (2). The existence and uniqueness of the solution of the weak parabolic convection-diffusion equation present no problems; see e.g. [18]. Two time discretization schemes are considered, namely, a backward Euler scheme and a Crank-Nicolson scheme. The space variable is discretized by the piecewise linear finite element method. Let Δt_n be the n^{th} time step, M the number of steps and $\sum_{n=1}^{M} \Delta t_n = T$. For the backward-Euler-Galerkin scheme,

$$\left(\frac{u_h^n - u_h^{n-1}}{\Delta t_n}, \phi_h\right) + b(u_h^n, \phi_h) = (f, \phi_h), \quad \forall \phi_h \in V^h,$$

with $u_h^0(x,t) = u_0(x)$ and $n = 1, \cdots, M$. For the Crank-Nicolson-Galerkin scheme

$$\left(\frac{u_h^n - u_h^{n-1}}{\Delta t_n}, \phi_h\right) + b\left(\frac{u_h^n + u_h^{n-1}}{2}, \phi_h\right) = (f, \phi_h), \quad \forall \phi_h \in V^h,$$

with $u_h^0(x,t) = u_0(x)$ and $n = 1, \cdots, M$. Both schemes lead to the following problem: For a given function $g_{n-1} \in H^{-1}(\Omega)$, find $w_h \in V^h$, such that

$$d_\tau(w_h, \phi_h) \equiv (w_h, \phi_h) + \tau b(w_h, \phi_h) = (g_{n-1}, \phi_h), \ \forall \phi_h \in V^h, \tag{36}$$

where τ is the time step parameter. The stability of both schemes is well-understood; see e.g. [18]. The algorithms discussed in the previous sections can thus be applied to the solution of (36) at each time step. An obvious initial guess is the approximate solution obtained at the previous time step. For the backward-Euler-Galerkin scheme

$$\begin{aligned} w_h &= u_h^n - u_h^{n-1}, \\ \tau &= \Delta t_n, \\ (g_{n-1}, \phi_h) &= \tau((f, \phi_h) - b(u_h^{n-1}, \phi_h)), \end{aligned}$$

and it is known that the truncation error is $O(\tau + h^2)$, therefore it is reasonable to assume that τ is of order h^2 to maintain the balance of the time and space discretization errors. The factor τ/H^2 is thus $(h/H)^2$ bounded, and thus the algorithms without coarse mesh spaces are optimal. Similarly, for the Crank-Nicolson-Galerkin scheme,

$$\begin{aligned} w_h &= u_h^n - u_h^{n-1}, \\ \tau &= \tfrac{\Delta t_n}{2}, \\ (g_{n-1}, \phi_h) &= \tau(2(f, \phi_h) - b(u_h^{n-1}, \phi_h)), \end{aligned}$$

and since the truncation error is $O(\tau^2 + h^2)$, the factor τ/H^2 is approximately h/H^2. As long as this h/H^2 is kept reasonably small, the methods without coarse mesh spaces should also behave well.

In the rest of this section, we present some numerical results for model problems. We have only tested the one- and two-level methods; the numerical performance of the multilevel methods will be studied elsewhere. To specify our model problems, we give only the elliptic part of the parabolic operator. We consider the following linear second order elliptic operator defined on $\Omega = [0,1] \times [0,1]$,

$$Lu = -\frac{\partial}{\partial x}(\xi \frac{\partial u}{\partial x}) - \frac{\partial}{\partial y}(\eta \frac{\partial u}{\partial y}) + \alpha \frac{\partial u}{\partial x} + \beta \frac{\partial u}{\partial y} + \gamma u,$$

with the homogenous Dirichlet boundary condition. The right-hand side is chosen so that the exact solution is $u = e^{xy} sin(\pi x) sin(\pi y)$. The coefficients are specified as follows.

Example 1. $\xi = 1$, $\eta = 1$ and $\alpha = \beta = \gamma = 0$.

Example 2. $\xi = 1 + x^2 + y^2$, $\eta = e^{xy}$, $\alpha = 5(x+y)$, $\beta = 1/(1+x+y)$ and $\gamma = 0$.

The domain Ω is partitioned with two uniform triangular grids that have sizes h and H, respectively. The actual values of h and H are given in Tables 1 and 2. The overlapping subdomains are colored by using four colors as in Fig. 1 and the coarse grid has color 0. The differential operator is discretized by the usual 5-point center finite difference method at both the coarse and fine levels. We assume that the time step has the form

$$\tau = h^\epsilon,$$

where $\epsilon = 0.25, 0.5, 1.0, 1.25$ and 1.5. All algorithms are accelerated by the GMRES method without restarting. The subdomain problems, as well as the coarse mesh problem, are solved exactly by a direct method. We present the iteration counts for GMRES that is required to solve equation (36). The iteration is stopped when the $\|r_i\|_2/\|r_0\|_2 \le 10^{-5}$, where r_i is the preconditioned residual at the ith iteration. The initial guess is always zero.

For the two-level method, we test various coarse grid sizes and the iteration numbers are given in Table 1. As predicted in the theory, the numbers are independent of H and τ(or ϵ).

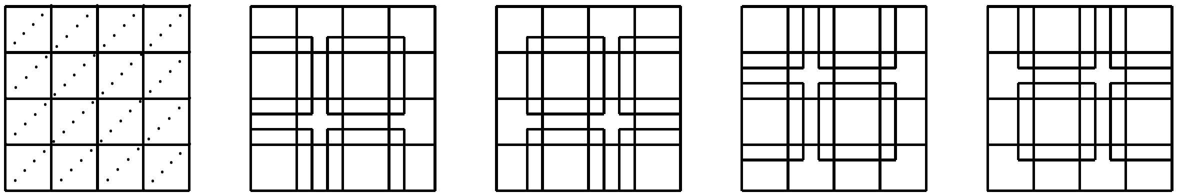

FIG. 1. *The coloring pattern of 16 fine grid overlapped subregions and a coarse grid region. Color "0" is for the global coarse grid. The extended subregions of the other colors are indicated by the dotted boundaries.*

TABLE 1
Iteration counts for solving the problems given in examples 1 and 2. The two-level method is used with the fine mesh size uniformly $h = 1/128$ and the overlap is $1/8$ of H.

	Example 1					Example 2				
$H\backslash\epsilon$	0.25	0.5	1.0	1.25	1.5	0.25	0.5	1.0	1.25	1.5
1/4	5	5	4	4	3	5	5	5	4	3
1/8	4	4	4	4	4	4	4	4	4	4
1/16	3	3	4	4	4	4	4	4	4	4

If the coarse solve is dropped as in the one-level method, the iteration numbers become more sensitive to the ratio τ/H^2 as seen in Table 2, especially for the nonsymmetric problem. We see that if the ratio τ/H^2 is relatively small, the results are acceptable as compared with the cases that use a coarse mesh solve.

Acknowledgement: I would like to thank Professor Olof Widlund for many valuable discussions and comments on this paper.

REFERENCES

[1] H. Bramble and J. E. Pasciak, *New estimates for multilevel algorithms including the V-cycle*, Brookhaven Nat. Lab. Rep. # BNL-46730, 1991.

[2] J. H. Bramble, J. E. Pasciak, and A. H. Schatz, *The construction of preconditioners for elliptic problems by substructuring, II*, Math. Comp., Vol. 49, pp. 1 - 16, 1987.

[3] J. H. Bramble, J. E. Pasciak, J. Wang and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition and multigrid,* Math. Comp., 57 (1991), pp. 1-22.

[4] X.- C. Cai, *Additive Schwarz algorithms for parabolic convection-diffusion equations,* Numer. Math. 60 (1991), pp. 41-62.

[5] X.- C. Cai, *An optimal two-level overlapping domain decomposition method for elliptic problems in two and three dimensions*, SIAM J. Sci. Comp., 14 (1993), pp. 239-247.

[6] X.- C. Cai, W. D. Gropp and D. E. Keyes, *A comparison of some domain decomposition and ILU preconditioned iterative methods for nonsymmetric elliptic problems,* J. Numer. Lin. Alg. Appl., June 1993.

[7] X.- C. Cai and O. B. Widlund, *Domain decomposition algorithms for indefinite elliptic problems,* SIAM J. Sci. Stat. Comp., 13 (1992), pp. 243-258.

[8] X. -C. Cai and O. B. Widlund, *Multiplicative Schwarz algorithms for nonsymmetric and indefinite elliptic problems,* SIAM J. Numer Anal., 30, 1993.

[9] C. Dawson and T. F. Dupont, *Explicit/implicit, conservative, Galerkin domain decomposition procedures for parabolic problems,* Department of Mathematical Science, Rice University, TR90-26.

[10] M. Dryja, *Substructuring methods for parabolic problems,* Fourth International Symposium

Table 2

Iteration counts for solving the problems given in examples 1 and 2. The one-level method is used with a uniform fine mesh $h = 1/128$ and the overlap is $1/8$ of H. The approximate ratio τ/H^2 is given as the subscript of the iteration number for Example 1. The ratios are the same for Example 2 and are therefore omitted.

	Example 1					Example 2				
$H\backslash\epsilon$	0.25	0.5	1.0	1.25	1.5	0.25	0.5	1.0	1.25	1.5
1/4	$9_{(4.76)}$	$9_{(1.41)}$	$6_{(0.13)}$	$4_{(0.04)}$	$3_{(0.01)}$	11	10	7	5	4
1/8	$17_{(19.03)}$	$16_{(5.66)}$	$10_{(0.50)}$	$6_{(0.15)}$	$4_{(0.04)}$	20	19	12	8	6
1/16	$33_{(76.11)}$	$30_{(22.63)}$	$17_{(2.00)}$	$10_{(0.59)}$	$7_{(0.18)}$	39	38	22	14	10

on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, Y. Kuznetsov, G. Meurant, J. Périaux and O. Widlund, edt., SIAM, Philadelphia, 1991.

[11] M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions,* Tech. Rep. 339, Dept. of Comp. Sci., Courant Institute, 1987.

[12] M. Dryja and O. B. Widlund, *Some domain decomposition algorithms for elliptic problems,* Iterative Methods for Large Linear Systems, L. Hayes and D. Kincaid, eds., Academic Press, San Diego, California, 1989.

[13] M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems,* Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., SIAM, Philadelphia 1990.

[14] M. Dryja and O. B. Widlund, *Multilevel additive methods for elliptic finite element problems,* Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Edt. Wolfgang Hackbusch, Kiel, January 19–21, 1990

[15] S. C. Eisenstat, H. C. Elman and M. H. Schultz, *Variational iterative methods for nonsymmetric system of linear equations,* SIAM J. Numer. Anal. 20, (1983), pp. 345-357.

[16] R. E. Ewing, R. D. Lazarov, J. E. Pasciak, and P. S. Vassilevski, *Finite element methods for parabolic problems with time steps variable in space,* Tech. Rep. # 1989-05, Inst. for Sci. Comp. at Univ. of Wyoming.

[17] P. Grisvard, *Elliptic Problems in Nonsmooth Domains,* Pitman, Boston, MA, 1985.

[18] C. Johnson, *Numerical Solution of Partial Differential Equation by the Finite Element Method,* Cambridge University Press, 1987.

[19] Y. A. Kuznetsov, *Domain decomposition methods for unsteady convection-diffusion problems,* IXth Int. Conf. on Comp. Meth. in Appl. Sci. and Eng., Edt. R. Glowinski and J. L. Lions, Paris, 1990

[20] P. L. Lions, *On the Schwarz alternating method. I,* Domain Decomposition Methods for Partial Differential Equations R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, edt., SIAM, Philadelphia, 1988.

[21] J. Nečas, *Sur la coercivité des formes sesquilinéaires, elliptiques,* Rev. Roumaine Math. Pures Appl., Vol. 9, pp. 47 – 69, 1964.

[22] P. Oswald, *On discrete morm estimates related to multilevel preconditioners in the finite element method,* Proc. Int. Conf. Theory of Functions, Varna 91, 1991.

[23] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm,* TR 91-279, Monnesota Supercomputer Institute, Univ. of Minnesota, 1991.

[24] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems,* SIAM J. Sci. Stat. Comp., 7, (1986). pp. 865-869.

[25] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, Vol. 2, Springer, Berlin, 1890, pp. 133-143. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, volume 15, 1870, pp. 272–286.

[26] J. Xu, *Theory of multilevel methods*, Ph.D. Thesis, Cornell 1989.

[27] X. J. Zhang, *Multilevel Schwarz methods,* Numer. Math., 63 (1992). pp. 521-539.