OPTIMAL ORDER MULTILEVEL PRECONDITIONERS FOR REGULARIZED ILL-POSED PROBLEMS

ANDREI DRĂGANESCU AND TODD F. DUPONT

Abstract. In this article we design and analyze multilevel preconditioners for linear systems arising from regularized inverse problems. Using a scale-independent distance function that measures spectral equivalence of operators, it is shown that these preconditioners approximate the inverse of the operator to optimal order with respect to the spatial discretization parameter $h$. As a consequence, the number of preconditioned conjugate gradient iterations needed for solving the system will decrease when increasing the number of levels, with the possibility of performing only one fine-level residual computation if $h$ is small enough. The results are based on previously known two-level preconditioners [20, 30], and on applying Newton-like methods to the operator equation $X^{-1} - A = 0$. We require that the associated forward problem has certain smoothing properties, however, only natural stability and approximation properties are assumed about the discrete operators. The algorithm is applied to a reverse-time parabolic equation, that is, the problem of finding the initial value leading to a given final state. We also present some results about constructing restriction operators with preassigned approximating properties that are of independent interest.

1. Introduction

We consider the equation

$$ Ku = f, $$

where $K : X \rightarrow Y$ is a compact linear operator between two Hilbert spaces. Since (1.1) is generally ill-posed, a generalized solution is given by $u = K^\dagger f$, where $K^\dagger : \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp \rightarrow X$ is the Moore-Penrose generalized inverse of $K$. For $f \in D(K^\dagger)$, the solution $K^\dagger f$ is defined as the element $u \in X$ of smallest norm that minimizes the expression $|Ku - f|^2$. If $\mathcal{R}(K)$ is not closed, then, in order to allow general perturbations in the right-hand side of (1.1), one can approximate $K^\dagger$ by a variety of regularization operators that are defined on all of $Y$ (see [16] for details). Perhaps the best known is the Tikhonov regularization, which amounts to solving the identity-perturbed normal equation

$$ H_\beta u = \beta^{-1}K^* f, $$

2000 Mathematics Subject Classification. Primary 65M32, 65M55; Secondary 65M60.

Key words and phrases. Multilevel methods, inverse problems, parabolic equations.

The work of the authors was supported by the ASCI Flash Center at the University of Chicago under DOE contract B532820, and by the MRSEC Program of the National Science Foundation under award DMR-0213745.

©1997 American Mathematical Society
where
\begin{equation}
H_\beta \overset{\text{def}}{=} I + \beta^{-1} H, \quad \text{and} \quad H \overset{\text{def}}{=} K^* K,
\end{equation}
with $K^*$ being the adjoint of $K$. For convenience we have multiplied the normal equation by $\beta^{-1}$. It should be noted that solving (1.2) is equivalent to applying one step of Newton’s method to the minimization of the quadratic functional
\begin{equation}
J_\beta(u) = \frac{\beta^{-1}}{2} |Ku - f|^2 + \frac{1}{2} |u|^2
\end{equation}
starting from the initial guess $u = 0$, and that $H_\beta$ is the Hessian operator associated to the quadratic $J_\beta$ via the $\mathcal{X}$-inner product. We add the subscript (resp. superscript) $h$ to a vector (resp. operator) to denote its discrete version, $h$ representing the discretization parameter. In this work, we are interested in fast solvers for computing the Tikhonov regularization of $K^T$ for operators $K$ with certain smoothing properties.

If $K^*K$ is smoothing, and $\beta$ is small enough, then the spectrum $\sigma(H_\beta)$ has only a few eigenvalues that are significantly larger than 1 (and they correspond to smooth eigenvectors), while the remainder of the spectrum is clustered around the value 1. Therefore unpreconditioned Krylov-type methods are expected to be efficient in inverting $H_\beta^h$. For example, we can show that, if $K$ is the solution operator for the heat equation mapping the initial value onto the final-time state, then the number of conjugate gradient (CG) iterations required to solve the regularized inverse problem down to machine precision is mesh-independent; moreover, the needed number of iterations grows only logarithmically as $\beta \to 0$ (see Chapter 7 in [16], and also [12, 1]). However, the number of iterations needed for convergence, even though independent of resolution, may be too large for practical use in case of large-scale problems, where the application of $K^h$ (i.e., the direct problem) requires, for example, solving a time-dependent three-dimensional partial differential equation. A standard strategy for solving large-scale optimization (or inverse) problems is to use reduced order models, and perform optimization (or inversion) on the model rather than the direct (large) problem itself; this strategy usually restricts the accuracy and resolution of the solution process. In this work we show that multilevel techniques can be successfully employed in efficiently solving certain types of inverse problems without resorting to reduced models, thus producing highly resolved solutions.

Multilevel methods, especially multigrid methods (MGMs), have been used extensively to efficiently solve linear systems related to partial differential equations [2, 3, 4, 6, 8, 19]. The efficiency of MGMs relies on the availability of good smoothers, that is, iterators that are inexpensive to apply, and that remove high-frequency components from the approximation error. In the classical multigrid theory, originally developed for inverting elliptic operators, standard iterators such as Jacobi and Gauß-Seidel are natural choices for smoothers, this being related to the fact that the operator to be inverted has roughening properties (being of differential type). However, for operators with smoothing properties such as $H_\beta^h$, the aforementioned iterators remove the low-frequency components from the error, thus leaving an error that cannot be represented accurately on a coarser mesh. In [25], frequently cited as one of the first studies of MGMs explicitly designed for ill-posed problems, a smoother was introduced based on the idea that that on high-frequency eigenvectors, the operator $H_\beta^h$ acts essentially as the identity, therefore the projector
on the high-frequency eigenspace acts as a smoother. A number of later studies [22, 23, 24] are dedicated to MGMs for unregularized inverse problems. ‘Unregularized’ is meant as ‘not explicitly regularized’, in fact regularization is achieved through discretization (by limiting the highest resolution that can be used) or by premature termination of the iterative process. The goal of MGMs for unregularized ill-posed problems is to design preconditioners that achieve a mesh-independent error reduction, in other words, whose approximation quality is mesh-independent. As mentioned before (see also [20, 30]), for the severely ill-posed problems under attention, this aim is already achieved by unpreconditioned CG. In this article we show that, under certain circumstances, these types of problems can actually be solved at a lower cost than just a mesh-independent number of iterations; in fact the number of iterations is shown to decrease with increasing number of levels.

We first construct a two-level preconditioner that is closely related to the additive Schwarz preconditioner introduced in [30]; we denote this by TLAS (Two-Level Additive Schwarz). The significant difference concerning TLAS between our work and [30, 20] is that the discrete operators we are considering here are not obtained by orthogonal projection from their continuous versions, but rather they are derived from natural discretizations. The only assumptions we make about the discretizations are related to their approximation and stability properties. Also, our two-level analysis is different from the one in [30, 20]; TLAS is shown here to approximate $(H_h^k)^{-1}$ to optimal order (see Section 4 for details). We use similar ideas to analyze a two-level preconditioner for an inverse semilinear parabolic problem in a forthcoming paper [13], the technical details therein being significantly more involved. By simply replacing the call to the coarse-level inverse operator in TLAS with a recursive call to TLAS, one could obtain a multilevel preconditioner which turns out to be of suboptimal quality. In particular, the number of preconditioned CG iterations would not decrease with increasing number of levels. However, the first Newton iterate of the operator-function $X \mapsto X^{-1} - H_h^k$ starting at this multilevel operator produces an optimal order multilevel preconditioner, denoted by MLAS (Multi-Level Additive Schwarz). MLAS has a W-cycle structure, and we should point out that it escapes the “usual” paradigm of pre-smoothing followed by error-correction and post-smoothing; here smoothing is intertwined with error-correction, thus introducing an inter-smoothing step.

We would like to point out that, as in [30], the main results in this article apply in the regime when $h^p \ll \beta$, where $p$ is the convergence order of the discretization $\mathcal{K}^h$ (see also (2.5)). We would like to argue that this is the case for a class of problems that motivated our research. The small parameter $\delta$ dictating the choice of $h$ and $\beta$ for an inverse problem is related to the noise level in the data $f$; that is, one typically assumes that the actual “measured” data $f$ used in (1.1) satisfies $\|f - f_0\| < \delta$ ,

where $f_0$ is the “correct” data. The natural assumption is that $h$ should be chosen so that $h^p \sim \delta$. A priori rules for choosing $\beta$ usually take into account source conditions on $u^0$ [16]: if a Hölder-type holds

$$u^0 = \mathcal{K}^1 f_0 \in \mathcal{R}(\mathcal{K}^* \mathcal{K})^\sigma$$ ,

then the choice $\beta \sim O(\delta^\frac{p}{p+1})$ gives the optimal convergence rate of

$$\|u_\beta - u\| = O(\delta^\frac{p}{p+1}) \quad \text{as } \delta \to 0 \ ,$$
where $u_\beta$ is the solution of (1.2). Therefore

$$\frac{h^p}{\beta} \sim \delta \frac{2^{m-1}}{\nu},$$

which implies that, for $\nu > \frac{1}{2}$,

$$h^p \ll \beta, \quad \text{as } \delta \rightarrow 0.$$

We refer the reader also to [32, 29] for discussions regarding the optimal choice of the regularization parameter $\beta$.

One of our motivating problems is the question of recovering an early stage of an air pollutant in an atmospheric model, given later time measurements [1]. In this case the direct operator is given by $K = S(T)$, where $t \mapsto S(t)$ is the solution operator of a linear advection-diffusion equation (see also Section 6), and $T > 0$ is a given time. If the coefficients of the parabolic equation are smooth, then $K$ is infinitely smoothing. In general, Hölder-type conditions are too restrictive, and should be replaced by logarithmic source conditions [21], case in which an a priori choice of the regularization parameter leads to $\beta \sim O(\delta)$. However, in a practical and plausible scenario, where concentration measurements of the pollutant are sparsely placed in space, there is a time-lag $\tau$ between the moment of the original spill and the time when sensors determine that the pollution event has taken place. Thus it is natural to assume that the “initial” state $u^I$ to be recovered has a history of length $\tau$, i.e., $u^I = S(\tau)u_0$. If $\tau > T$, that is, the “initial” state’s history is longer than the difference $T$ between measurements-collection time and “initial” time, then $u^I \in \mathcal{R}((K^*K)^{\nu})$ with $\nu > \frac{1}{2}$, under appropriate and reasonable assumptions on the parabolic equation (cf. Picard’s criterion, see [16]). A similar situation is encountered in a data assimilation setting for, say, weather prediction. Here $S(t)$ is the evolution operator for the velocity field of a fluid (air). Various direct or indirect measurements provide a part of the history of a process for which the entire state at a single time is unknown. Recovering a possible “initial state” at an artificially chosen time $t_0$ in the past enables “learning” the entire state at the current time, thus potentially improving predictions from this time on. Again, the initial state to be recovered naturally is $S(\tau)u$, where $u$ is a state at an even earlier time, which, again leads to a Hölder-type source condition with $\nu > \frac{1}{2}$, thus making the inverse problem amenable to the case when $h^p \ll \beta$.

This article follows closely [12], and is organized as follows: after describing the problem in Section 2, we define the spectral distance between operators with positive definite symmetric part in Section 3; this distance – a measure of spectral equivalence – introduces a framework that is convenient for analyzing the two-level (Section 4) and multilevel (Section 5) algorithms. Theorems 4.1 and 5.4 are the central results of this paper. We present an application of these methods to inverse problems of parabolic type in Section 6. The article concludes with a section on numerical results. Appendix B contains a brief notation summary.

2. Notation and Problem Formulation

Let $\mathcal{X} = \mathcal{Y} = L^2(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is an open set ($d \geq 1$). Throughout this paper we shall denote by $H^m(\Omega), H^m_0(\Omega)$ ($m \in \mathbb{N}$) the standard Sobolev spaces, and by $\|\cdot\| = \|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_m$ (resp. $\|\cdot\|_m$) the corresponding norms (resp. semi-norms); furthermore $\langle \cdot, \cdot \rangle$ (resp. $\langle \cdot, \cdot \rangle_m$) will be the standard inner product in $L^2(\Omega)$.
that have the nesting property

\[ H^{m}(\Omega). \]  

Let \( \bar{H}^{-m}(\Omega) \) be the dual (with respect to the \( L^2 \)-inner product) of \( H^{m}(\Omega) \cap H^1_0(\Omega) \) for \( m > 0 \). For \( T \in L(V_1, V_2) \) we denote the operator norm of \( T \) by

\[ \|T\|_{V_1, V_2} = \sup_{u \in V_1 \setminus \{0\}} \frac{|Tu|}{\|u\|}, \quad \text{and} \quad \|T\|_V \overset{\text{def}}{=} \|T\|_{V, V}. \]

In the absence of any subscript, \( \|T\| \) denotes \( \|T\|_{L^2(\Omega)} \). We consider a set of approximating spaces \((V_h)_{h \in I}\) with

\[ I = \{h_{\text{max}}/2^i : i \in \mathbb{N}\} \]

that have the nesting property

\[ V_{2h} \subset V_h \subset H^1_0(\Omega), \quad \forall h \in I \setminus \{h_{\text{max}}\}. \]

Furthermore, we denote by \( \pi_h \) the \( L^2 \)-orthogonal projection onto \( V_h \). Discretizations \( K^h \in L(V_h) \) of \( K \) give rise to the following discrete quadratic functional to be minimized:

\[ J^h_f(u_h) = \frac{\beta^{-1}}{2} |K^h u - f_h|^2 + \frac{1}{2} |u_h|^2, \]

where \( f_h = \pi_h(f) \in V_h \). Throughout this paper it is assumed that the operators \( K \) and \( K^h \) together with their adjoints satisfy

**Condition 2.1.** There exists a number \( p > 0 \) (the approximation order) and constants \( C_1 = C_1(p, \|K\|, \Omega) \) and \( C_2 = C_2(p, \Omega) \) such that for all \( h \in I \) the following hold:

- **[a]** stability:
  \[ \|K^h u\| \leq C_1 \|u\| \quad \text{and} \quad \|(K^h)^* u\| \leq C_1 \|u\|, \quad \forall u \in V_h ; \]

- **[b]** smoothed approximation:
  \[ \|K u - K^h u\| \leq C_1 h^p \|u\| \quad \text{and} \quad \|K^* u - (K^h)^* u\| \leq C_1 h^p \|u\|, \quad \forall u \in V_h ; \]

- **[c]** negative-index norm approximation of the identity by the projection:
  \[ \|(I - \pi_h)u\|_{\bar{H}^{-p}(\Omega)} \leq C_2 h^p \|u\|, \quad \forall u \in L^2(\Omega) ; \]

- **[d]** smoothing:
  \[ \|K u\| \leq C_1 \|u\|_{\bar{H}^{-p}} \quad \text{and} \quad \|K^* u\| \leq C_1 \|u\|_{\bar{H}^{-p}}, \quad \forall u \in L^2(\Omega) . \]

**Remark 2.2.** Condition 2.1 implies that there is a \( C = C(p, \|K\|, \Omega) \) such that

\[ \|Hu\|_{L^2} \leq C \|u\|_{\bar{H}^{-p}}, \quad \forall u \in L^2(\Omega) , \]

and

\[ \|K(I - \pi_h)u\| \leq Ch^p \|u\| \quad \text{and} \quad \|K^*(I - \pi_h)u\| \leq Ch^p \|u\|, \quad \forall u \in L^2(\Omega) . \]

In order to minimize \( J^h_f \) we need to invert the discrete version of \( H_{\beta} \), namely

\[ H_{\beta}^h \overset{\text{def}}{=} I + \beta^{-1} H^h, \quad \text{with} \quad H^h \overset{\text{def}}{=} (K^h)^* K^h . \]

**Lemma 2.3.** The following approximation property holds:

\[ \|\pi_h(H^h - H)u\| \leq Ch^p \|u\|, \quad \forall u \in V_h . \]

for some constant \( C = C(p, \|K\|, \Omega) \).
Proof. For $u \in \mathcal{V}_h$ we have

$$
|\langle \pi_h (H^h - H) u, u \rangle| = |\|K^h u\|^2 - \|K u\|^2| \\
\leq |K^h u - K u| \cdot (|K^h u| + |K u|) \\
\leq C (C + \|K\|) h^p \|u\|^2 ,
$$

and (2.11) follows from the symmetry of $\pi_h (H^h - H) \in L(\mathcal{V}_h)$. \qed

It follows from the definition of $H_{\beta}$ that for all $u \in L^2(\Omega)$

$$
\|u\|^2 \leq \langle H_{\beta} u, u \rangle \leq \left(1 + \frac{C}{\beta}\right) \|u\|^2 ,
$$

with $C = C(p, \|K\|, \Omega)$, and similar estimates hold for the discrete Hessian. This implies that the quadratic $\mathcal{J}_{\beta}$ is positive definite and has a unique minimizer given by

$$
u_{\min}^{\beta} = \beta^{-1} (H_{\beta})^{-1} \mathcal{K}^* f .
$$

Similarly the minimizer of the discrete quadratic is

$$
u_{h, \min}^\beta = \beta^{-1} (H_{h, \beta}^\beta)^{-1} (K^h)^* f_h .
$$

We refer the reader to [15, 16] for results concerning convergence of of $\nu_{\beta}$ or $\nu_{h, \beta}$ to $K$. The next result shows that $\nu_{h, \min}$ approximates $\nu_{\min}$ to optimal order in the $L^2$-norm, and plays a role in the analysis of the multilevel method.

**Theorem 2.4.** Assume that Condition 2.1 holds. Then there exists a constant $C = C(p, \|K\|, \Omega)$ such that for $h \leq h_0(\beta, p, \|K\|, \Omega)$ we have the following stability and error estimates:

$$
\|\nu_{h, \min}^\beta - \nu_{\min}^\beta\| \leq C \{ \|\nu_{\min}^\beta\| + \beta^{-1} h^p \|f\| \} ,
$$

$$
|\nu_{h, \min}^\beta - \nu_{\min}^\beta| \leq C \frac{h^p}{\beta} \{ \|f\| + \|\nu_{\min}^\beta\| \} .
$$

Proof. Denote by $e_h = \nu_{h, \min}^\beta - \nu_{\min}^\beta$. We have

$$
H_{\beta} e_h = \beta^{-1} \{ ((K^h)^* - \mathcal{K}^*) f_h - \mathcal{K}^*(I - \pi_h) f + (H - H^h) \nu_{h, \min}^\beta \} .
$$

Therefore

$$
\beta \|e_h\|^2 \leq \langle H_{\beta} e_h, e_h \rangle
$$

$$
\leq \langle ((K^h)^* - \mathcal{K}^*) f_h - \mathcal{K}^*(I - \pi_h) f + (H - H^h) \nu_{h, \min}^\beta, e_h \rangle
$$

$$
\leq C h^p \|f\| \|e_h\| + \langle (H - H^h) \nu_{h, \min}^\beta, e_h \rangle
$$

$$
\leq \langle Ch^p \|f\| + \|\nu_{h, \min}^\beta\| \rangle + \langle (H - H^h) \nu_{h, \min}^\beta, \mathcal{K}(I - \pi_h) e_h \rangle
$$

$$
\leq C h^p \|f\| \|\nu_{h, \min}^\beta\| + \|\nu_{h, \min}^\beta\| \mathcal{K}(I - \pi_h) e_h
$$

with $C = C(p, \|K\|, \Omega)$. Hence it follows that

$$
\|\nu_{h, \min}^\beta - \nu_{\min}^\beta\| \leq C \frac{h^p}{\beta} \{ \|f\| + \|\nu_{h, \min}^\beta\| \} .
$$
Since
\[ \|u^\min_h\| \leq \|u^\min\| + \|u^\min - u^\min_h\| \leq \|u^\min\| + C \frac{\|\hat{f}\| + \|u^\min_h\|}{\beta}, \]
we obtain (2.15) and (2.16), for h small enough.

3. The spectral distance

In this section we define a scale-independent distance between operators with positive definite symmetric part, and we study its relevant properties. This spectral distance is a measure of spectral equivalence between two operators, and introduces a convenient framework for the multilevel analysis in Section 5.

Throughout this section \((\mathcal{X}, \langle \cdot, \cdot \rangle)\) is a real, finite dimensional Hilbert space. As usual, \(\|u\| = \sqrt{\langle u, u \rangle}\) is the Hilbert-space norm of \(u \in \mathcal{X}\). Denote by
\[ L_+(\mathcal{X}) = \{ T \in L(\mathcal{X}, \mathcal{X}) : \langle Tu, u \rangle > 0, \quad \forall u \in \mathcal{X} \setminus \{0\} \} \]
the set of operators with positive definite symmetric part. All operators in this section are assumed to be in \(L_+(\mathcal{X})\). If \(A \in L_+(\mathcal{X})\) is symmetric we write \(\|u\|_A = \sqrt{\langle Au, u \rangle} = \|A^{1/2}u\|\). Our object of study is the preconditioned Richardson iteration
\[ x_{n+1} = x_n + M(b - Hx_n) \]
leading to the solution \(x^\ast\) of the equation
\[ Hx = b, \]
where \(b \in \mathcal{X}, H \in L_+(\mathcal{X})\). Later we will specialize to \(H\) being symmetric (we think of \(H\) being Hessian \(H_h^\beta\) from (2.10)), but we allow the preconditioner \(M\) to be non-symmetric, for reasons explained in Section 4.2. The results we prove for \(H\) symmetric (essentially Theorem 3.12 and Corollary 3.13) apply to the reverse situation as well, with the preconditioner \(M\) being symmetric and the operator \(H\) non-symmetric, a situation that has been studied beginning with [10, 14, 17]. Also, some of the techniques in this section are rooted in the aforementioned papers.

It is well known that the error \(e_n = x_n - x^\ast\) satisfies
\[ e_n = (I - MH)^n e_0. \]
If \(H\) is symmetric, then
\[ e_n = H^{-\frac{1}{2}} (I - H^{1/2}MH^{1/2})^n H^{1/2} e_0, \]
and
\[ \|e_n\|_H \leq \|I - H^{1/2}MH^{1/2}\|^n \cdot |e_0|_H. \]
Therefore the quantity \(\|I - MH\|\) (or \(\|I - H^{1/2}MH^{1/2}\|\), if \(H\) is symmetric) is an upper bound for the convergence rate of (3.1). Another way of measuring the quality of the preconditioner \(M\), especially useful when neither \(M\) nor \(H\) are symmetric, is the spectral radius \(\rho(I - MH)\). Although all the above quantities measure, in spirit, how far \(M^{-1}\) is from \(H\), none of them is a distance function in a strict mathematical sense. For technical reasons that will become clear in Section 5 we prefer to assess the quality of \(M\) by using an actual (scale-free) distance function to measure how far \(M^{-1}\) and \(H\) are from each other.

We denote the complexification of \(\mathcal{X}\) by
\[ \mathcal{X}^C = \mathcal{X} \otimes_{\mathbb{R}} \mathbb{C} = \{ u + iv : u, v \in \mathcal{X}\}, \]
and we consider the natural extension of the inner product on \( \mathcal{X} \) to a Hermitian product on the complex vector space \( \mathcal{X}^\mathbb{C} \):
\[
\langle u_1 + i v_1, u_2 + i v_2 \rangle = \langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle + i \left( \langle v_1, u_2 \rangle - \langle u_1, v_2 \rangle \right);
\]
for \( T \in L(\mathcal{X}) \) define the complexification of \( T \) to be \( T^\mathbb{C} \in L(\mathcal{X}^\mathbb{C}) \) (the space of \( \mathbb{C} \)-linear maps), where
\[
T^\mathbb{C}(u + iv) = T(u) + iT(v).
\]
We will drop the superscript \( \mathbb{C} \) whenever there is no potential for confusion. Furthermore, we denote by \( B_r(z_0) \) the open disc \( \{ z \in \mathbb{C} : |z - z_0| < r \} \).

**Definition 3.1.** Let \( T_1, T_2 \in L_+(\mathcal{X}) \). We define the *spectral distance* between \( T_1 \) and \( T_2 \) to be
\[
d_X(T_1, T_2) = \sup_{w \in \mathcal{X}^\mathbb{C} \setminus \{ 0 \}} \left| \ln \frac{\langle T^\mathbb{C}_1 w, w \rangle}{\langle T^\mathbb{C}_2 w, w \rangle} \right|
\]
\[
= \sup_{(u,v) \in \mathcal{X}^\mathbb{C} \times \mathcal{X}^\mathbb{C} \setminus \{(0,0)\}} \left| \ln \frac{\langle T_1 u, u \rangle + \langle T_1 v, v \rangle + i \left( \langle T_1 v, u \rangle - \langle T_1 u, v \rangle \right)}{\langle T_2 u, u \rangle + \langle T_2 v, v \rangle + i \left( \langle T_2 v, u \rangle - \langle T_2 u, v \rangle \right)} \right|,
\]
where \( \ln \) is the branch of the logarithm corresponding to \( \mathbb{C} \setminus (-\infty, 0] \).

We should point out that, for \( T_1, T_2 \in L_+(\mathcal{X}) \), the quotients \( \langle T^\mathbb{C}_1 w, w \rangle / \langle T^\mathbb{C}_2 w, w \rangle \) do not lie in \( (-\infty, 0] \). Also note that, if \( T = T_s + T_a \) with \( T_s \) (resp. \( T_a \)) being the symmetric (resp. antisymmetric) part of \( T \), then
\[
\langle T^\mathbb{C}(u + iv), u + iv \rangle = \langle T_s u, u \rangle + \langle T_s v, v \rangle + 2i \langle T_a v, u \rangle.
\]
The polarization identity
\[
\langle Tu, v \rangle = \frac{1}{4} \left( \langle T(u + v), u + v \rangle - \langle T(u - v), u - v \rangle + i \left( \langle T(u + iv), u + iv \rangle - \langle T(u - iv), u - iv \rangle \right) \right)
\]
implies that \( d_X(T_1, T_2) = 0 \) if and only if \( T_1 = T_2 \). We leave as an exercise to the reader the verification of the symmetry and triangle inequality (see also [12]).

The following elementary (but nontrivial) inequalities prove useful in evaluating the spectral distance.

**Lemma 3.2.** If \( \alpha \in (0, 1) \) and \( z \in B_\alpha(1) \), then
\[
\frac{\ln \left( 1 + \frac{1}{\alpha} \right)}{\alpha} |1 - z| \leq |\ln z| \leq \frac{\ln \left( 1 - \frac{1}{\alpha} \right)}{\alpha} |1 - z|.
\]
For \( |\ln z| \leq \delta \) we have
\[
\frac{1 - e^{-\delta}}{\delta} |\ln z| \leq |1 - z| \leq \frac{e^\delta - 1}{\delta} |\ln z|.
\]

**Proof.** The modulus of the analytic function \( f : B_\alpha(1) \to \mathbb{C} \) defined by \( f(z) = \ln z \) attains its extreme values on the boundary \( \partial B_\alpha(1) \), as \( f \) has no zeros in \( B_\alpha(1) \). On the circle \( C_\alpha(1) = \{ z : |1 - z| = \alpha \} \) we have \( |f(z)| = \alpha^{-1} |\ln z| \); hence the extreme values of \( |f(z)| \) are attained at the same points as the extremes of \( |\ln z|^2 \). The problem is thus reduced to showing that the maximum of \( |\ln z|^2 \) on the circle \( C_\alpha(1) \) is attained at \( \zeta = 1 - \alpha \), and the minimum at \( \zeta = 1 + \alpha \). We leave this calculus problem as an exercise (hint: Draw tangents from the origin to \( C_\alpha(1) \), and parametrize the two resulting arches using polar coordinates around 0. Use the
angle as the free variable. The functions to be analyzed are $g_{1, 2}(\theta) = \ln(\rho_{1, 2}(\theta)) + \theta^2$, with $\rho_{1, 2}(\theta) = \cos \theta \pm \sqrt{\alpha^2 - \sin^2 \theta}$. For (3.5) we proceed similarly. □

Remark 3.3. We should note the following short one-line proof of the fact that the global maximum of $|\ln z|$ on $C(1)$ is located at $z = 1 - \alpha$:

$$|\ln z| = \left| \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z - 1)^n}{n} \right| \leq \sum_{n=1}^{\infty} \frac{\alpha^n}{n} = |\ln(1 - \alpha)| .$$

However, we found no such simple argument for showing that the global minimum is at $z = 1 + \alpha$.

The following result shows that, even if neither $M$ nor $H$ is symmetric, the spectral radius of $(I - MH)$ is controlled by the spectral distance between $M^{-1}$ and $H$.

Lemma 3.4. Let $M, H \in L_+(\mathcal{X})$ such that $d_\mathcal{X}(M^{-1}, H) \leq \delta$. Then

$$\rho(I - MH) \leq \frac{e^\delta - 1}{\delta} d_\mathcal{X}(M^{-1}, H) .$$

In particular, if $d_\mathcal{X}(M^{-1}, H) < \ln 2$, then $\rho(I - MH) < 1$.

Proof. Let $\lambda \in \sigma(I - MH)$. Then there exists a unit vector $u \in \mathcal{X}^\mathbb{C}$ such that $(I - MH)u = \lambda u$, which further implies that

$$\lambda = 1 - \frac{\langle Hu, u \rangle}{\langle M^{-1}u, u \rangle} .$$

Hence

$$\rho(I - MH) \leq \sup \{ |1 - z| \ : \ z = \langle Hu, u \rangle / \langle M^{-1}u, u \rangle \text{ for some } u \in \mathcal{X}^\mathbb{C} \setminus \{0\} \} \overset{(3.5)}{\leq} \frac{e^\delta - 1}{\delta} d_\mathcal{X}(M^{-1}, H) .$$

□

Remark 3.5. In general, the converse of (3.6) does not hold. For example, let $H = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, and $M = I$. Then $\rho(I - MH) = 0$, but $d_\mathcal{X}(M^{-1}, H) \neq 0$.

A concept related to the spectral distance is the numerical range or field of values of an operator $T$, which is defined in [18] as

$$W(T) = \{ \langle T^c u, u \rangle : u \in \mathcal{X}^\mathbb{C}, \|u\| = 1 \} ;$$

the numerical radius is

$$w(T) = \sup \{ |\lambda| : \lambda \in W(T) \} .$$

We recall two results from [18]:

Theorem 3.6 (Theorem 1.3-1 in [18]). For any $T \in L(\mathcal{X})$

$$w(T) \leq \|T\| \leq 2w(T) .$$
This is basically saying that $w(\cdot)$, which is a norm, is equivalent to the operator norm $\| \cdot \|$. However, for symmetric operators we have

$$\| T \| = w(T),$$

since $W(T) = \text{co}(\sigma(T))$, (co(A) denotes the convex hull of the set A). The next result is about mapping of the numerical radius under a power function:

**Theorem 3.7** (Theorem 2.1-1 in [18]). For any $T \in L(X)$

$$w(T^n) \leq (w(T))^n, n = 1, 2, 3, \ldots.$$  

By analogy we define the joint numerical range to be

$$W(T_1, T_2) = \left\{ \frac{\langle T_1 w, w \rangle}{\langle T_2 w, w \rangle} : w \in X^C \setminus \{0\} \right\}.$$  

Note that $W(T, I) = W(T)$. With this definition the spectral distance becomes

$$d_X(T_1, T_2) = \sup \{ |\ln(\lambda)| : \lambda \in W(T_1, T_2) \}.$$  

**Remark 3.8.** If $T_2$ is symmetric then

$$W(T_1, T_2) = W(T_2^{-\frac{1}{2}}T_1T_2^{-\frac{1}{2}}).$$

This is easily seen from the change of variable $w = T_2^{-\frac{1}{2}}v$ in (3.10).

**Lemma 3.9.** If $T_1, T_2 \in L_+(X)$ are symmetric, then

$$d_X(T_1, T_2) = \sup_{u \in X \setminus \{0\}} \frac{\langle T_1 u, u \rangle}{\langle T_2 u, u \rangle},$$

and

$$\| I - T_1^{\frac{1}{2}}T_2^{-\frac{1}{2}}T_2^{\frac{1}{2}} \|_{d_X(T_1, T_2)} \to 1 \quad \text{as} \quad d_X(T_1, T_2) \to 0;$$

Proof. Let $T_{12} = T_2^{-\frac{1}{2}}T_1T_2^{-\frac{1}{2}}$. The symmetry of $T_{12}$ implies

$$W(T_1, T_2) \overset{(3.12)}{=} W(T_{12}) \overset{(3.13)}{=} \left\{ \frac{\langle T_{12} u, u \rangle + \langle T_{12} v, v \rangle}{\langle u, u \rangle + \langle v, v \rangle} : (u, v) \neq (0, 0) \right\}.$$  

Since $W(T_{12})$ is convex (Theorem 1.1-2 in [18]) we have

$$\left\{ \frac{\langle T_{12} u, u \rangle + \langle T_{12} v, v \rangle}{\langle u, u \rangle + \langle v, v \rangle} : (u, v) \neq (0, 0) \right\} \subseteq \left\{ \frac{\langle T_{12} u, u \rangle}{\langle u, u \rangle} : u \neq 0 \right\}.$$  

The reverse inclusion being evident, we get

$$W(T_1, T_2) = \left\{ \frac{\langle T_{12} u, u \rangle}{\langle u, u \rangle} : u \neq 0 \right\} \overset{(3.14)}{=} \left\{ \frac{\langle T_1 u', u' \rangle}{\langle T_2 u', u' \rangle} : u' \neq 0 \right\};$$

and (3.13) follows by (3.11). Moreover, since all operators involved are symmetric, $W(T_{12}) = \text{co}(\sigma(T_{12})) = [\lambda_{\text{min}}, \lambda_{\text{max}}]$, where $\lambda_{\text{min}}$ (resp. $\lambda_{\text{max}}$) is the smallest (resp. largest) eigenvalue of $T_{12}$. The fact that $d_X(T_1, T_2) \to 0$ translates into $\lambda_{\text{min}} \to \lambda_{\text{max}} \to 1$; hence

$$\| I - T_{12} \|_{d_X(T_1, T_2)} = \frac{\max(1 - \lambda_{\text{min}}, 1 - \lambda_{\text{max}})}{\max(\ln \lambda_{\text{min}}, \ln \lambda_{\text{max}})} \to 1.$$
The following lemma is a restating in terms of the spectral distance of equivalent results from [10, 14, 17].

**Lemma 3.10.** For symmetric operators \( T_1, T_2 \in L_+(X) \) we have

\[
\Delta(T_1, T_2) = \Delta((T_1)^{-1}, (T_2)^{-1})
\]

**Proof.** As in the proof of Lemma 3.9 we have

\[
W(T_1, T_2) = \left\{ \frac{\langle T_{12} u, u \rangle}{\langle u, u \rangle} : u \neq 0 \right\} \quad \text{with} \quad \frac{\langle v, v \rangle}{\langle T_{12}^{-1} v, v \rangle} : v \neq 0
\]

\[
= \left\{ \frac{\langle v, v \rangle}{\langle T_2^{-1} T_1^{-1} T_2 T_1^{-1} v, v \rangle} : v \neq 0 \right\}
\]

\[
= W(T_2^{-1}, T_1^{-1})
\]

and the conclusion follows from (3.11). \( \square \)

The remainder of this section is devoted to showing that if \( H \) is symmetric, and \( H^{1/2} M H^{1/2} \) is close to \( I \), then the spectral distance can replace \( \| I - H^{1/2} M H^{1/2} \| \) in the estimate (3.2), at the expense of a slightly increased constant on the right-hand side (see (3.27)); this fact is not surprising, since the values of \( \Delta(T_1, T_2) \) and \( \| I - T_1^{-1/2} T_2 T_1^{-1/2} \| \) become asymptotically close as \( \Delta(T_1, T_2) \to 0 \). It will be convenient to regard the Richardson iteration (3.1) as an iteration of preconditioners. More precisely, we think of (3.1) as a sequence of single iterations with “improved” preconditioners:

\[
x_n = x_0 + M_n (b - H x_0)
\]

with \( M_0 = 0 \) and \( M_n \) recursively defined by \( M_{n+1} = M + M_n - M H M_n \). A simple calculation shows that

\[
M_n = H^{-1} - (I - MH)^n H^{-1} = H^{-1/2} \left( I - \left( I - H^{1/2} M H^{1/2} \right)^n \right) H^{-1/2}
\]

In particular \( M_1 = M \); moreover, \( M_2 = 2M - M H M \) is the first Newton iterate with initial guess \( M_1 \) of the operator-function \( X \mapsto X^{-1} - H \), as shown below.

**Remark 3.11.** Define \( G : L_+(X) \to L(X) \) by \( G(X) = X^{-1} - H \). Since \( G'(X) U = -X^{-1} UX^{-1} \) (see [9]), it follows that \( (G'(X))^{-1} U = -UX \); the Newton iteration for solving \( G(X) = 0 \) is

\[
X_{i+1} = X_i - (G'(X_i))^{-1}(X_i^{-1} - H) = 2X_i - X_i H X_i
\]

Hence the iteration is \( X_{i+1} = N_H(X_i) \) with

\[
N_H(X) \triangleq 2X - XHX
\]

We will be using the operator \( N_H \) in designing a multilevel preconditioner in Section 5 in the following way: given an initial (and unacceptable) guess \( X \) at \( H^{-1} \), the next best guess is \( N_H(X) \). By (3.17) the first Richardson iterate with \( N_H(X) \) as a preconditioner instead of \( M \), namely the value \( x_0 + N_H(X)(b - H x_0) \), is equal to \( x_2 \) from (3.1). Therefore applying \( N_H(X) \) amounts to performing two Richardson iterations with \( M \) as a preconditioner.
Theorem 3.12. Let $M, H \in L_+(X)$ and $M_n$ be defined as in (3.17). If $H$ is symmetric and $d_X(M, H^{-1}) < \ln 2$, then $M_n \in L_+(X)$, and

\begin{equation}
\frac{d_X(M_n, H^{-1})}{d_X(M, H^{-1})} \leq g_n \left(\frac{d_X(M, H^{-1})}{H^{-1}}\right),
\end{equation}

where $g_n(x) = |\ln(1 - (e^x - 1)^n)| = x^n + O(x^{n+1})$; hence

\begin{equation}
\lim_{n \to \infty} M_n = H^{-1}
\end{equation}
in the metric $d_X$. 

Proof. Let $\alpha = d_X(M, H^{-1})$. By (3.11) and the fact that

\[
\max \{|e^w - 1| : w \in B_\alpha(0)\} = e^\alpha - 1,
\]

it follows that

\[
W(H^{1/2}MH^{1/2}) = W(M, H^{-1}) \subset B_{e^\alpha - 1}(1);
\]

therefore

\begin{equation}
w \left(I - H^{1/2}MH^{1/2}\right) \leq e^\alpha - 1.
\end{equation}

Equation (3.17) implies that

\begin{equation}
\left(I - H^{1/2}M_nH^{1/2}\right) = \left(I - H^{1/2}MH^{1/2}\right)^n;
\end{equation}

hence by (3.9) and (3.21) we get

\begin{equation}
w \left(I - H^{1/2}M_nH^{1/2}\right) \leq \gamma \overset{\text{def}}{=} (e^\alpha - 1)^n.
\end{equation}

Since $\max\{|\ln(z)| : z \in B_\gamma(1)\} = |\ln(1 - \gamma)|$ (here we used $\gamma < 1$ which follows from $\alpha < \ln 2$, and also implies that $M_n$ is positive definite), we get

\begin{equation}
d_X(M_n, H^{-1}) \leq |\ln(1 - (e^\alpha - 1)^n)|,
\end{equation}

which concludes the proof. □

Corollary 3.13. Under the conditions of Theorem 3.12 the sequence $x_n$ defined by the simple iteration (3.1) converges to the solution $x^*$, and we have the estimate

\begin{equation}
|e_n|_H \leq 2 g \left(\frac{d_X(M_n, H^{-1})}{H^{-1}}\right) \cdot |e_0|_H,
\end{equation}

where $g(x) = e^x - 1 = x + o(x)$.

Proof. The relations (3.2), (3.23) and (3.7) imply

\begin{equation}
|e_n|_H \leq 2 w \left(I - H^{1/2}M_nH^{1/2}\right) \cdot |e_0|_H.
\end{equation}

Similarly to the proof of Theorem 3.12 we have

\[
w \left(I - H^{1/2}M_nH^{1/2}\right) \leq g \left(d_X(M_n, H^{-1})\right),
\]

with $g$ as in the hypothesis. □

Remark 3.14. In light of the asymptotic behavior of the functions $g_n$ and $g$ above, we rewrite (3.25) in the following way: for every $C > 2$ there exists $\delta(C) > 0$ such that, if $d_X(M, H^{-1}) < \delta(C)$, then

\begin{equation}
|e_n|_H \leq C \left(d_X(M, H^{-1})\right)^n \cdot |e_0|_H.
\end{equation}
4. Two-level preconditioners for the discrete Hessian

In this section we construct and analyze two-level preconditioners for the discrete Hessian $H^h_\beta$ defined in (2.10). Given that $H^h_\beta$ is symmetric, it is natural to seek preconditioners that are symmetric as well (Section 4.1). However, the results in Section 3 show that symmetry of the preconditioner is not essential for the Richardson iteration to converge. In Section 4.2 we investigate the possibility of using slightly non-symmetric operators as preconditioners for $H^h_\beta$. Non-symmetry may occur by allowing restriction operators other than exact $L^2$-projections. The main results in this section are Theorems 4.1 and 4.3, which essentially state that the spectral distance between the constructed preconditioners and the inverse of the Hessian is $O(h^p/\beta)$.

We should point out that the results presented in this section, although similar to the ones in [30, 20] are obtained in a different context. While the discrete operators in the aforementioned papers are obtained by orthogonal projections of their continuous versions on the finite dimensional spaces under consideration, our discrete operators arise from natural finite element discretizations, and the only assumptions made are related to convergence and stability. Naturally the results in Section 4.1 have counterparts in [30, 20], but the techniques used here are different.

4.1. A symmetric preconditioner. For the remainder of this article we consider on $V_h$ the Hilbert-space structure inherited from $L^2(\Omega)$. Let $W_{2h}$ be the orthogonal complement of $V_{2h}$ in $V_h$, and let $\rho_{2h}$ be the orthogonal projection onto $W_{2h}$, so that $\pi_{2h} + \rho_{2h} = I$. We shall refer to $V_h$ as the fine space in relation to $V_{2h}$, which will be the coarse space. We consider the splitting

$$H^h_\beta = \pi_{2h} H^h_\beta \pi_{2h} + \rho_{2h} H^h_\beta \pi_{2h} + \pi_{2h} H^h_\beta \rho_{2h} + \rho_{2h} H^h_\beta \rho_{2h}.$$  

Standard Fourier analysis on uniform grids shows that $\pi_{2h}$ extracts the low-frequency components from a function in $V_h$, while $\rho_{2h}$ extracts the high-frequency components. For $h \ll 1$ it is expected that $H^h$ inherits the smoothing character of $H = K^*K$, therefore annihilates high-frequency components, i.e., $H^h \rho_{2h} \approx 0$. Similarly, $\rho_{2h} H^h \approx 0$, since this operation amounts to extracting high-frequency components from a “smoothened” function. This suggests that the cross terms $A_2, A_3$ are negligible (since $A_2 \approx \rho_{2h} \pi_{2h} = 0$, $A_3 \approx \pi_{2h} \rho_{2h} = 0$), and that $A_4 \approx \rho_{2h}$ (see [11] for more details). On “smooth” components both $H^h$ and $H^{2h}$ approximate well $H$, since all eigenvectors of $H$ corresponding to large eigenvalues are well represented on fine grids (the high-frequency eigenvectors of $H$ correspond to negligible eigenvalues); hence $A_1 \approx H^{2h}_\beta$. It remains that

$$H^h_\beta \approx M^h_\beta \overset{\text{def}}{=} H^{2h}_\beta \pi_{2h} + (I - \pi_{2h}).$$

Since $\pi_{2h}$ is a projection, the inverse $(M^h_\beta)^{-1}$ is given explicitly by

$$L^h_\beta = L^h_\beta(\pi_{2h}) \overset{\text{def}}{=} (M^h_\beta)^{-1} = (H^{2h}_\beta)^{-1} \pi_{2h} + (I - \pi_{2h}).$$

We propose $L^h_\beta \in L(V_h)$ as a two-level preconditioner for $H^h_\beta$. Due to the similarity with the additive Schwarz construction from domain decomposition, and in accordance with [30], we denote it here as the TLAS (Two-Level Additive Schwarz) preconditioner. We would like to remark that $M^h_\beta$ and $L^h_\beta$ are symmetric. It should be noted that the projector $I - \pi_{2h}$ removes high-frequency components from the
approximation error, therefore it acts as a smoother. However, unlike the case of classical multigrid, neither is the solution a fixed point for this smoother, nor does it make sense to apply the smoother more than once.

The remainder of this section is devoted to proving the following theorem, which gives an estimate of the distance between $L^h$ and $(H^h)^{-1}$. We write $d_h$ instead of $d_{V_h}$.

**Theorem 4.1.** Assume that the operators $K, K^h$ satisfy Condition 2.1. Then there exists a constant $C = C(p, \|K\|, \Omega)$ such that for $h \leq h_0$ with $h_0 = h_0(\beta, p, \|K\|, \Omega)$ we have $M^h_\beta \in L_+(V_h)$, and

$$
(4.2) \quad d_h(H^h_\beta, M^h_\beta) \leq C \frac{h^p}{\beta}.
$$

Before we proceed to the proof of the theorem we state the corollary which legitimizes the use of $L^h_\beta$ as a preconditioner for the Hessian.

**Corollary 4.2.** Under the hypotheses of Theorem 4.1 there exists a constant $C = C(p, \|K\|, \Omega)$ such that

$$
(4.3) \quad d_h \left(L^h_\beta, (H^h_\beta)^{-1}\right) \leq C \frac{h^p}{\beta},
$$

for $h \leq h_0$ with $h_0 = h_0(\beta, p, \|K\|, \Omega)$.

**Proof.** The result follows directly from Theorem 4.1 and Lemma 3.10.

We would like to remark that, for fixed $\beta$, the quality of the preconditioner $L^h_\beta$ increases with $h \to 0$. This is different from standard multigrid preconditioning for elliptic problems, where the goal would be to show that the right-hand side of (4.3) is independent of $h$. It should not be surprising that (4.3) is such an optimistic estimate. This is related to the adverse eigenstructure of the eigenvalue-frequency correlation for the operator $H^h_\beta$. By increasing resolution we only add high-frequency eigenfunctions whose eigenvalues are close to 1. On the other hand the low-frequency eigenvectors (with higher energy) are increasingly well approximating the continuous (smooth) eigenvectors of $H^2_\beta$, and therefore they are increasingly well approximated by corresponding eigenvectors of $H^2_\beta$. Hence the fine-space operator is increasingly well approximated by the identity on the high-frequency part, and by the coarse-space operator on the low-frequency part. We return to the proof of Theorem 4.1:

**Proof.** We have the following relation on $V_h$:

$$
\beta(M^h_\beta - H^h_\beta) = \underbrace{\pi_{2k}(H^{2h} - H)\pi_{2k}}_{A} + \underbrace{(\pi_{2k}H\pi_{2k} - \pi_kH\pi_k)}_{B} + \underbrace{(\pi_kH\pi_k - H^h)}_{C}.
$$

Note that all operators in the above sum are symmetric in $L(V_h)$. We analyze the products $\langle Au, u \rangle, \langle Bu, u \rangle, \langle Cu, u \rangle$. For $u \in V_h$

$$
|\langle Au, u \rangle| = |\langle (H^{2h} - \pi_{2k}H)\pi_{2k}u, \pi_{2k}u \rangle| \overset{(2.11)}{\leq} C(2h)^p \|\pi_{2k}u\|^2 \leq C(2h)^p \|u\|^2.
$$
Similarly we obtain $| (Cu, u) | \leq Ch^{p} \| u \|^2$. For the middle term we have

$$
| (Bu, u) | = | (\pi_{2h} H\pi_{2h} - \pi_{h} H\pi_{h})u, u | \leq \| S_{h} \| \| (H\pi_{2h} u, \pi_{2h} u) - (Hu, u) \| 
$$

$$
\leq \| K\pi_{2h} u \|^2 - \| K\pi_{h} u \|^2 = \| K(I - \pi_{2h})u \| \cdot (\| K\pi_{2h} u \| + \| K\pi_{h} u \|) 
$$

(2.9)

$$
\leq C h^{p} \| u \|^2 ,
$$

with $C = C(p, \| K \|, \Omega)$. Putting the above estimates together we get

$$
(4.4) \quad \frac{\langle M_{\beta}^{h} u, u \rangle}{\langle H_{\beta}^{h} u, u \rangle} - 1 \leq C \frac{h^{p}}{\beta} \cdot \frac{\| u \|^2}{\beta^{-1} \langle H^{h} u, u \rangle + \| u \|^2} \leq C \frac{h^{p}}{\beta} .
$$

Assuming $C\beta^{-1} h^{p} = \alpha < 1$ and $0 < h \leq h_{0}$ we obtain $M_{\beta}^{h} \in L_{+}(\mathcal{V}_{h})$, and

$$
\sup_{u \in \mathcal{V}_{h} \setminus \{0\}} \left| \frac{\langle M_{\beta}^{h} u, u \rangle}{\langle H_{\beta}^{h} u, u \rangle} \right| \leq \left| \frac{\ln(1 - \alpha)}{\alpha} \right| \sup_{u \in \mathcal{V}_{h} \setminus \{0\}} \frac{\langle M_{\beta}^{h} u, u \rangle}{\langle H_{\beta}^{h} u, u \rangle} - 1 \leq C \frac{\ln(1 - \alpha)}{\alpha} \cdot h^{p} .
$$

The conclusion follows from the symmetry of $M_{\beta}^{h}$ and $H_{\beta}^{h}$ and from Lemma 3.9. \hfill \Box

4.2. Non-symmetric preconditioners. The application of the preconditioner $L_{\beta}^{h}(\pi_{2h})$ from Section 4.1 requires computing exact (or very close to exact) $L^{2}$-projections. In this section we prove a result similar to Theorem 4.1 that applies to a larger class of preconditioners, obtained by replacing in (4.1) the orthogonal projection $\pi_{2h}$ with a more general restriction operator $R_{2h}^{h} \in L(\mathcal{V}_{h}, \mathcal{V}_{2h})$ that verifies Condition 2.1[c]. A local restriction operator that satisfies a similar, but slightly weaker condition is defined in [5]. In Section 6.3 we show how to construct local restriction operators with the required properties.

**Theorem 4.3.** Assume that the operators $K, K^{h}$ satisfy Condition 2.1, and that the restriction operator $R_{2h}^{h} \in L(\mathcal{V}_{h}, \mathcal{V}_{2h})$ verifies for all $\beta \in (0, 1]$ and $h \in I$

[f] stability:

$$
| R_{2h}^{h} u | \leq C \| u \| , \quad \forall u \in \mathcal{V}_{h} ;
$$

approximation of the identity in negative-index norm:

$$
\| (I - R_{2h}^{h}) u \|_{-\rho(\Omega)} \leq C h^{p} \| u \| , \quad \forall u \in \mathcal{V}_{h} ,
$$

with $C = C(p, \| K \|, \Omega)$. If we denote by

$$
L_{\beta}^{h} = L_{\beta}^{h}(R_{2h}^{h}) \quad \text{def} \quad (H_{\beta}^{h})^{-1} R_{2h}^{h} + (I - R_{2h}^{h}) ,
$$

then for $h \leq h_{0}$ with $h_{0} = h_{0}(\beta, p, \| K \|, \Omega)$ we have $L_{\beta}^{h} \in L_{+}(\mathcal{V}_{h})$, and

$$
d_{h} ((H_{\beta}^{h})^{-1}, L_{\beta}^{h}) \leq C h^{p} .
$$

for some constant $C = C(p, \| K \|, \Omega)$.

Note that this result, though applying to a larger class of restriction operators, is weaker than Theorem 4.1, due of the extra power of $\beta$ in the denominator of the right-hand side of (4.8). We first prove
Lemma 4.4. Under the hypotheses of Theorem 4.3 there exists a constant $C = C(p, \|K\|, \Omega)$ such that, for all $u \in \mathcal{V}_h$, the following hold:

(i) smoothing properties of the discrete Hessian:

\[
\max(\{H^h(I - R_{2h}^h)u, \|H^{2h}(\pi_{2h} - R_{2h}^h)u\|\} \leq C h^p \|u\| ;
\]

(ii) smoothing properties of the inverse discrete Hessian:

\[
\|((H^2h)\beta^{-1} - I)(\pi_{2h} - R_{2h}^h)u\| \leq C \beta^{-1} h^p \|u\| .
\]

Proof. Throughout this proof $C$ will denote a generic constant depending only on $p, \|K\|$ and $\Omega$.

(i) Conditions (2.6) and (4.6) imply that for $u \in \mathcal{V}_h$

\[
\|((\pi_{2h} - R_{2h}^h)u)\|_{\mu - p} \leq C h^p \|u\| .
\]

For $u \in \mathcal{V}_h$ we have

\[
\| (H^{2h} - \pi_{2h}H)(\pi_{2h} - R_{2h}^h)u \| \leq C(2h)^p \|((\pi_{2h} - R_{2h}^h)u \| \leq C h^p \|u\| ;
\]

hence

\[
\| H^{2h}(\pi_{2h} - R_{2h}^h)u \| \leq \|\pi_{2h}H(\pi_{2h} - R_{2h}^h)u\| + C(2h)^p \|u\|
\]

\[
\leq C \|((\pi_{2h} - R_{2h}^h)u\|_{\mu - p} + C h^p \|u\|
\]

\[
\leq C h^p \|u\| .
\]

The inequality

\[
\|H^h(I - R_{2h}^h)u\| \leq C h^p \|u\| , \quad u \in \mathcal{V}_h ,
\]

follows along the same lines, and (4.9) is proved.

(ii) Since $\langle H^{2h}u, u \rangle \geq 0$, it follows that $\langle H^{2h}_\beta u, u \rangle \geq \|u\|^2$, $\forall u \in \mathcal{V}_{2h}$. Hence

\[
\| (H^2h)\beta^{-1} u \| \leq \|u\| , \quad u \in \mathcal{V}_{2h} .
\]

For $u \in \mathcal{V}_h$

\[
\|((H^2h)\beta^{-1} - I)(\pi_{2h} - R_{2h}^h)u\| = \beta^{-1} \|((H^2h)\beta^{-1} - I)(H^2h)(\pi_{2h} - R_{2h}^h)u\|
\]

\[
\leq C \beta^{-1} h^p .
\]

We now proceed to the proof of Theorem 4.3.

Proof. Recall that in Corollary 4.2 it was shown that

\[
d_h \left( (M^h_\beta)^{-1}, (H^h_\beta)^{-1} \right) \leq C \beta^{-1} h^p ,
\]

where

\[
(M^h_\beta)^{-1} = (H^h_\beta)^{-1} \pi_{2h} + (I - \pi_{2h}) .
\]

Since $\beta \leq 1$, it suffices to show that

\[
d_h \left( (M^h_\beta)^{-1}, L^h_\beta \right) \leq C \beta^{-2} h^p .
\]

We rewrite (4.7) as

\[
L^h_\beta = I + ((H^2h)\beta^{-1} - I) R_{2h}^h .
\]
and we start by evaluating the antisymmetric part of $L^h_\beta$:

$$
\langle L^h_\beta u, v \rangle - \langle L^h_\beta v, u \rangle = \langle (H^2 - I) R^h_{2h} u, v \rangle - \langle (H^2 - I) R^h_{2h} v, u \rangle
$$

$$
= \langle (H^2 - I) R^h_{2h} u, \pi_{2h} v \rangle - \langle (H^2 - I) R^h_{2h} v, \pi_{2h} u \rangle
$$

$$
= \langle (H^2 - I) (R^h_{2h} - \pi_{2h}) u, \pi_{2h} v \rangle
$$

$$
- \langle (H^2 - I) (R^h_{2h} - \pi_{2h}) v, \pi_{2h} u \rangle.
$$

For the last equality we have added and subtracted $\langle (H^2 - I) \pi_{2h} u, \pi_{2h} v \rangle$ and used the symmetry of $(H^2 - I)$ on $V_{2h}$. By Lemma 4.4 (ii) we get

$$
\langle L^h_\beta u, v \rangle - \langle L^h_\beta v, u \rangle \leq C\beta^{-1} h^p \|u\| \cdot \|v\|.  \tag{4.14}
$$

Note that $M^h_\beta = H^2 + \pi_{2h} + (I - \pi_{2h}) = \beta^{-1} H^2 + I$; hence

$$
\langle u, u \rangle \leq \langle M^h_\beta u, u \rangle \leq C\beta^{-1} \langle u, u \rangle, \quad \text{for } u \in V_h,
$$

with $C = C(p, \|K\|, \Omega)$. This implies

$$
C^{-1} \beta \langle u, u \rangle \leq \langle M^h_\beta u, u \rangle \leq \langle u, u \rangle, \quad \text{for } u \in V_h. \tag{4.16}
$$

It follows that the projection of the set $W(L^h_\beta, (M^h_\beta)^{-1})$ onto the imaginary axis (see (3.3)) satisfies

$$
\frac{|\langle L^h_\beta u, v \rangle - \langle L^h_\beta v, u \rangle|}{\langle (M^h_\beta)^{-1} u, u \rangle + \langle (M^h_\beta)^{-1} v, v \rangle} \leq C\beta^p \frac{h^p}{\beta^2} \cdot \frac{|u| \cdot |v|}{|u|^2 + |v|^2} ;
$$

we obtain therefore

$$
\sup \{|\Im w| : w \in W(L^h_\beta, (M^h_\beta)^{-1})\} \leq C\beta^{-2} h^p,
$$

where $\Im w$ is the imaginary part of a complex number $w$. We now turn our attention to the projection of $W(L^h_\beta, (M^h_\beta)^{-1})$ onto the real axis. Since $(M^h_\beta)^{-1}$ is symmetric, this amounts to computing the joint numerical range of the symmetric part of $L^h_\beta$ and $(M^h_\beta)^{-1}$. By Lemma 3.9 we need to evaluate

$$
\frac{\langle L^h_\beta u, u \rangle}{\langle (M^h_\beta)^{-1} u, u \rangle} = 1 + \frac{\langle (L^h_\beta - (M^h_\beta)^{-1}) u, u \rangle}{\langle (M^h_\beta)^{-1} u, u \rangle} \tag{4.18}
$$

for $u \in V_h \setminus \{0\}$. A simple calculation shows that

$$
L^h_\beta - (M^h_\beta)^{-1} = (H^2 - I) (R^h_{2h} - \pi_{2h})
$$

therefore by from Lemma 4.4 (ii) it follows that

$$
|\langle (L^h_\beta - (M^h_\beta)^{-1}) u, u \rangle| \leq C\beta^{-1} h^p \|u\|^2 ;
$$

this, together with (4.18) and (4.16), implies

$$
\{ \Re w : w \in W(L^h_\beta, (M^h_\beta)^{-1}) \} \subseteq [1 - C\beta^{-2} h^p, 1 + C\beta^{-2} h^p], \tag{4.20}
$$
where $\Re w$ is the real part of a complex number $w$. Finally from (4.17) and (4.20) it follows that
\begin{equation}
W \left( L^h_\beta, (M^h_\beta)^{-1} \right) \subset B_{C^p_\beta \cdot 2^h \rho}(1).
\end{equation}
In particular this implies, possibly by further reducing $h_0$ by a factor of two, that $L^h_\beta \in L_+(V_h)$. Lemma 3.2 implies that (4.13) holds for $h \leq h_0(\beta, p, \|K\|, \Omega)$. □

5. A MULTILEVEL PRECONDITIONER

In this section we define a multilevel preconditioner $K^h_\beta$ that is of comparable quality with the two-level preconditioner $L^h_\beta$ defined in Section 4. By analogy we will denote the preconditioner by MLAS (Multi-Level Additive Schwarz).

5.1. Design and work estimates. Before constructing $K^h_\beta$ we want to point out that one natural way to extend $L^h_\beta$ to a multilevel preconditioner $G^h_\beta$ is by replacing $(H^{2h}_\beta)^{-1}$ in (4.1) with the coarse-space preconditioner; thus we define the preconditioner recursively:
\begin{equation}
G^h_\beta \overset{\text{def}}{=} G^h_\beta \pi_{2h} + (I - \pi_{2h}).
\end{equation}
At the coarsest level $h_0$ we use conjugate gradient as an “almost direct” solver. It is immediate that $G^h_\beta \in L_+(V_h)$. This strategy yields a multilevel preconditioner with a $V$-cycle structure, which is very similar to the one defined in [30]. In Theorem 5.2 we prove that the distance between $G^h_\beta$ and $(H^{2h}_\beta)^{-1}$ only depends on the coarsest level resolution, which is consistent with the results in [30]. This preconditioner is suboptimal, and this fact should not be surprizing: the multilevel $G^h_\beta$ is in fact only a two-level operator, in the sense that we would obtain the same $G^h_\beta$ if we had used only the finest and the coarsest levels. For example, if we use three levels corresponding to resolutions $h$, $2h$, and $4h$, then
\begin{align*}
G^h_\beta &= G^h_\beta \pi_{2h} + (I - \pi_{2h}) = (G^{4h}_\beta \pi_{4h} + (I - \pi_{4h})) \pi_{2h} + (I - \pi_{2h}) \\
&= G^{4h}_\beta \pi_{4h} + (I - \pi_{4h}), \text{ since } \pi_{4h} \pi_{2h} = \pi_{4h}.
\end{align*}
Also we should point out that, when applying $G^h_\beta$, residuals at intermediate levels are never computed. Thus the quality of the multilevel preconditioner $G^h_\beta$, while not degrading, also does not improve with increasing fine-level resolution. As a result, for example, the number of $G^h_\beta$-preconditioned conjugate gradient iterations will be constant if $h_0$ is kept fixed and $h \to 0$. Our goal is to design a multilevel preconditioner whose quality improves with increasing fine-level resolution. This would imply that asymptotically (as $h \to 0$) we may only need one preconditioned iteration at the finest level.

We adopt a different strategy in defining an improved preconditioner $K^h_\beta$: in addition to replacing $(H^{2h}_\beta)^{-1}$ with the coarse-space preconditioner, like before, we perform one Newton iteration for the operator-function $X \to X^{-1} - H^h_\beta$, as explained in Remark 3.11. More precisely, if we denote by $G^h : L(V_{2h}) \to L(V_h)$ the affine operator-function
\begin{equation}
G^h(T) = T\pi_{2h} + (I - \pi_{2h}),
\end{equation}
then we define $K^h_\beta$ recursively using (3.18) by
\begin{equation}
K^h_\beta \overset{\text{def}}{=} N^h_\beta (G^h(K^{2h}_\beta)) = 2G^h(K^{2h}_\beta) - G^h(K^{2h}_\beta) \cdot H^h_\beta \cdot G^h(K^{2h}_\beta).
\end{equation}
With this notation the symmetric preconditioner from Section 4.1 is written as

$$L^h = G^h((H^h)^{-1})$$.

We should note that $G^h$ is symmetry-preserving, therefore $K^h$ is symmetric if $K^h_0$ is so. The application of $K^h_0$ requires one residual computation at the finest level (see Algorithm MLAS); moreover, performing one $K^h$-preconditioned Richardson iteration simply means to complete two $G^h(K^h_2)$-preconditioned Richardson iterations, as explained in Section 3. This implies that there will be two calls to the coarser-space procedure, therefore $K^h$ has a W-cycle structure. We would like to remark that the calling sequence pre-smoothing - restriction - error correction - interpolation - post-smoothing from classical multigrid is not appropriate for Algorithm MLAS. Here restriction applies to the right-hand side as a whole as opposed to the smoothed residual, as in classical multigrid, and smoothing occurs again between the error correction steps.

Algorithm MLAS. ($b \mapsto K^h b$)

1. if $h = h_0$ then return $K^h_0 b$ // direct solve
2. else
3. $b_c \leftarrow \pi_{2h} b$ // restriction
4. $u_1 \leftarrow b - b_c + K^2 h b_c$ // smoothing and error correction
5. $r \leftarrow b - H^h b_1$ // residual computation
6. $r_c \leftarrow \pi_{2h} r$ // restriction
7. $u_2 \leftarrow u_1 + K^2 h r_c + r - r_c$ // smoothing and error correction
8. return $u_2$

Let $h_i = 2^{-i} h_0$, $i = 0, 1, \ldots, l$. If we denote by $W(i)$ (resp. $R(i)$) the work needed to apply $K^h_i$ (resp. $H^h_i$), then

$$W(i) \approx R(i) + 2W(i-1).$$

We have assumed that the cost of restriction and interpolation is negligible compared to that of a residual computation. Indeed this is the case when $H^h_i$ is represented by a dense matrix, and restriction by a sparse matrix, or when the direct problem (i.e. applying $K$) is a space-time process and restriction operates only on the spatial variables (see Section 6). If $R(i-1) \leq 2^{-g} R(i)$ for some $g > 0$, then (5.4) implies

$$W(l) \leq (1 - 2^{-g})^{-1} R(l) + 2^l W(0).$$

It is shown in [16] that unpreconditioned conjugate gradient takes a level-independent number of iterations $N_{cg}$ to solve the exponentially ill-posed problems ($N_{cg}$ still depends on $\beta$). This results in $W(0) = C N_{cg} R(0)$, with $C = O(1)$ being a universal constant. Hence we have the estimate

$$W(l) \leq \left( (1 - 2^{-g})^{-1} + C2^l(1-g) N_{cg} \right) R(l).$$

The number $g$ is related to the dimension of the problem to be solved. For example, if the forward operator $K$ is a three-dimensional space-time operator, and we use a time stepping procedure with the same convergence order as the spatial discretization, then $g = 4$. With $l = 3$ levels, (5.6) results in $W(3) \approx$
the distance
\[ e_h = d_h(K^h_\beta, (H^h_\beta)^{-1}) . \]
This quantity lies at the basis of the residual and error estimates, as pointed out in Section 3. We will also estimate \( d_h(G^h_\beta, (H^h_\beta)^{-1}) \).

**Lemma 5.1.** Let \( T_1, T_2 \in L_+(V_{2h}) \) be symmetric. Then
\[ \| T_1 \|_h(T_1, G^h(T_2)) \leq d_{2h}(T_1, T_2) . \]

**Proof.** A simple calculation shows that for any \( a, b > 0 \) the function \( g(x) = |\ln(a + x) - \ln(b + x)| \) is non-increasing on \([0, \infty)\). Therefore
\[ \ln \frac{a + x}{b + x} \leq \ln \frac{a}{b} . \]
Since \( (I - \pi_{2h}) \) is positive semidefinite, we obtain for \( u \in V_h \) with \( \pi_{2h} u \neq 0 \)
\[ \ln \frac{\langle (T_1 \pi_{2h} + I - \pi_{2h}) u, u \rangle}{\langle (T_2 \pi_{2h} + I - \pi_{2h}) u, u \rangle} \leq \ln \frac{\langle T_1 \pi_{2h} u, \pi_{2h} u \rangle + \langle (I - \pi_{2h}) u, u \rangle}{\langle T_2 \pi_{2h} u, \pi_{2h} u \rangle + \langle (I - \pi_{2h}) u, u \rangle} \leq d_{2h}(T_1, T_2). \]
The conclusion follows after passing to supremum over all \( u \in V_h \) with \( \pi_{2h} u \neq 0 \) in the inequality above. \( \square \)

**Theorem 5.2** (V-cycle estimates). Under the hypotheses and in the notation of Theorem 4.1 there exists a constant \( C = C(p, \|K\|, \Omega) \) such that
\[ d_h(G^h_\beta, (H^h_\beta)^{-1}) \leq C \frac{h^p}{\beta} . \]

**Proof.** Assume that \( h = h_0 2^{-i} \) with \( h_0 \) small enough so that Theorem 4.1 applies. Then
\[ d_h(G^h_\beta, (H^h_\beta)^{-1}) \leq \left[ d_h(G^h_\beta, L^h_\beta) + d_h(L^h_\beta, (H^h_\beta)^{-1}) \right] \leq d_{2h}(G^h_\beta, (H^h_\beta)^{-1}) + C \beta^{-1} h^p . \]
We apply the above recursively to obtain
\[ d_h(G^h_\beta, (H^h_\beta)^{-1}) \leq C \beta^{-1} h^p \sum_{i=0}^{l} 2^{-ip} < (1 - 2^{-p})^{-1} C \beta^{-1} h^p , \]
which proves (5.10). \( \square \)
Lemma 5.3. Let \((e_i)_{i \geq 0}\) and \((a_i)_{i \geq 0}\) be positive numbers satisfying the recursive inequality
\begin{equation}
    e_{i+1} \leq C(e_i + a_{i+1})^2,
\end{equation}
and
\begin{equation}
    a_{i+1} \leq a_i \leq f^{-1}a_{i+1}
\end{equation}
for some \(0 < f < 1\). If \(a_0 \leq \frac{f}{4C}\) and if \(e_0 \leq 4Ca_0^2\), then
\begin{equation}
    e_i \leq 4Ca_i^2, \quad \forall i > 0.
\end{equation}

Proof. We proceed by induction: the base case is in the hypothesis, and if we assume \(e_i \leq 4Ca_i^2\), then
\[
    e_{i+1} \leq C(e_i + a_{i+1})^2 \leq C(4Ca_i^2 + a_{i+1})^2 \leq C(4Cf^{-1}a_i + a_{i+1})^2 = C(4Cf^{-1}a_i + 1)^2a_{i+1}^2 \leq C(4Cf^{-1}a_0 + 1)^2a_{i+1}^2 \leq 4Ca_{i+1}^2,
\]
which concludes the proof. \(\square\)

Theorem 5.4 (MLAS error estimates). Assume that Condition 2.1 holds, and that
\begin{equation}
    h_0^p \leq \min \left(0.1, 2^{-(p+3)}\right) \cdot \beta
\end{equation}
where \(C\) is the constant from Theorem 4.1. If \(e_{h_0} \leq 8 \left(C\beta^{-1}h_0^p\right)^2\) (if we solve the system almost exactly on the coarsest grid this is automatically satisfied), then \(K^h_\beta \in L_+(V_h)\), and
\begin{equation}
    d_h(K^h_\beta, (H^h_\beta)^{-1}) \leq 8 C \frac{h_0^{2p}}{\beta^2}, \quad \forall h \in I.
\end{equation}

Proof. Theorem 3.12 implies that, given operators \(M, H \in L_+(V_h)\) with \(d_h(M, H^{-1}) < 0.4\), it follows that \(M \in L_+(V_h)\), and
\begin{equation}
    d_h(N_h(M), H^{-1}) \leq 2 d_h(M, H^{-1})^2,
\end{equation}
because \(|\ln(1 - (e^x - 1)^2)| \leq 2x^2\) for \(x \in [0,0.4]\). If \(e_{2h} \leq 0.2\) and \(C\beta^{-1}h^p \leq 0.1\), then
\[
    d_h(G^h(K^2h_\beta), (H^h_\beta)^{-1}) \leq d_h(G^h(K^2h_\beta), L^h_\beta) + d_h(L^h_\beta, (H^h_\beta)^{-1}) = e_{2h} + C\beta^{-1}h^p \leq 0.3;
\]
hence by (5.17) we obtain
\begin{equation}
    e_h \leq 2 \left[d_h(G^h(K^2h_\beta), (H^h_\beta)^{-1})\right]^2 \leq 2(e_{2h} + C\beta^{-1}h^p)^2 \leq 2 \cdot 0.3^2 < 0.2.
\end{equation}
Assume that \(h = 2^{-i}h_0\), and let \(h_i = 2^{-i}h_0\). Denote by \(e_i = e_{h_i}\), and by \(a_i = C\beta^{-1}h_i^p\). An inductive argument implies that \(e_i \leq 0.2\) for all \(i\), provided that it holds for \(e_0\), and that \(C\beta^{-1}h_0^p \leq 0.1\). It follows that \((a_i)_{i \geq 0}\) satisfies (5.12) and (5.13) with \(f = 2^{-p}\). The other hypotheses of the theorem are tailored to fit the corresponding hypotheses in Lemma 5.3, which implies that
\begin{equation}
    e_i \leq 8 \left(C\beta^{-1}h_i^p\right)^2, \quad \forall i \geq 0.
\end{equation}
\(\square\)
The previous result shows that the quality of the multilevel preconditioner improves with increasing resolution; moreover the estimate is optimal with respect to $h$. We should remark that a proper comparison of the multilevel preconditioner with the two-level preconditioner from Section 4 requires weighing $K_h^4 = \mathcal{N}_{L_h^3}(K_{2h}^2 \pi_{2h} + (I - \pi_{2h}))$ against $\mathcal{N}_{L_h^3}(L_h^3) = \mathcal{N}_{L_h^3}((H_0^1)^{-1} \pi_{2h} + (I - \pi_{2h}))$ (rather than $L_h^3$). Theorems 5.4, Corollary 4.2 and Theorem 3.12 show that, the distances from $K_h^4$ and $\mathcal{N}_{L_h^3}(L_h^3)$ to $(H_0^1)^{-1}$ are asymptotically comparable, an assertion that is confirmed by the numerical results in Section 7.

In the regime of its applicability, namely when $h_\beta \ll \beta$, the estimate (5.16) seems to provide a solution mechanism that is better than needed: if the iterative process for finding $u_{h_{\beta}}^{\text{min}}$ starts with an $O(1)$ close initial guess, then one $K_h^4$-preconditioned iteration will give an $O\left(\left(\beta^{-1}h_\beta^2\right)^2\right)$ approximation to $u_{h_{\beta}}^{\text{min}}$. This is better than the approximation of $u_{h_{\beta}}^{\text{min}}$ by $u_{h_{\beta}}^{\text{min}}$ as shown in (2.16), therefore unnecessary. As shown at the end of Section 5.1, $K_{h_{\beta}}^4$ should be implemented so that no finest-level residual computation should be performed inside the preconditioner. The multigrid preconditioner can be used in conjunction with a Krylov solver, and all finest-level Hessian-vector multiplications should be left to the iterations in the solver.

6. Application to the finite element Galerkin approximation for parabolic problems

In this section we identify a class of linear parabolic problems and their discretizations for which Condition 2.1 is satisfied; in particular it will follow that all results in Sections 4 and 5 apply. The verification amounts primarily to providing links to the corresponding results in the literature. In Section 6.2 we discuss a few results on regularization for inverse parabolic problems and we devote Section 6.3 to the construction of local restriction operators.

6.1. The inverse problem. Let $\Omega \subset \mathbb{R}^d$ be an open set, with $d \geq 1$ an integer. We consider the following initial value problem

$$\begin{aligned}
\frac{\partial}{\partial t}u + A(t)u &= 0, \quad \text{on } \Omega \times (0, \infty), \\
u(x, t) &= 0, \quad \text{on } \partial \Omega \times (0, \infty), \\
u(x, 0) &= u_0(x), \quad \text{for } x \in \Omega,
\end{aligned}$$

(6.1)

where

$$A(t)u = -\sum_i \partial_i \left( \sum_j a_{ij}(x, t) \partial_j u + b_i(x, t) u \right) + c(x, t) u$$

(6.2)

with $a_{ij}(x, t) = a_{ji}(x, t), b_i(x, t), c(x, t)$ being smooth functions with uniformly bounded derivatives of all orders on $\Omega \times [0, \infty)$. We define the time-dependent bilinear form $a : (0, \infty) \times H^1_0 \times H^1_0 \rightarrow \mathbb{R}$ by

$$a(t; \phi, \psi) = \sum_{i,j} \langle a_{ij} \partial_j \phi, \partial_i \psi \rangle + \sum_i \langle b_i \phi, \partial_i \psi \rangle + \langle c \phi, \psi \rangle, \quad \text{for } \phi, \psi \in H^1_0 .$$

(6.3)

It is assumed that $a$ is coercive, i.e., there exists a constant $c_1 > 0$ independent of $t$ such that

$$a(t; \phi, \phi) \geq c_1 \|\phi\|_{H^1_0}^2, \quad \text{for } \phi \in H^1_0 ,$$

(6.4)
and that the boundary $\partial \Omega$ is smooth enough for the following regularity condition to hold:

$$
\|\phi\|_2 \leq c_2 \|A(t)\phi\|, \quad \text{for } \phi \in H^1_0 \cap H^2.
$$

This is verified for example in case $\Omega \subset \mathbb{R}^2$ is a convex polygonal domain and $a_{ij}(x,t) = \alpha(x,t) \delta_{ij}$ with $\alpha(x,t)$ a positive smooth function ($\delta$ is the Kronecker symbol), or if $\partial \Omega \in C^1$. Cf. [26] and [27], for $u_0 \in L^2$ there exists a unique solution $u : (0, \infty) \rightarrow H^1_0$ to the weak formulation of (6.1), namely

$$
\begin{align*}
\begin{cases}
\langle \partial_t u, \phi \rangle + a(t; u, \phi) = 0, & \text{for all } \phi \in H^1_0, \ t > 0, \\
\lim_{t \to 0} |u(t) - u_0| = 0.
\end{cases}
\end{align*}
$$

We denote the time-$t$ solution operator by $S(t)u_0 \overset{\text{def}}{=} u(\cdot, t)$.

Given a fixed $T > 0$ we define the operator $\mathcal{K} \in L(\mathcal{X})$ by

$$
\mathcal{K} \overset{\text{def}}{=} S(T),
$$

with $\mathcal{X} = L^2(\Omega)$. We restate Lemmas 2.3, 2.6 and equation (2.12) in [26] as

**Lemma 6.1.** For $p = 0, 1, 2$ and $t > 0$

$$
\|S(t)u_0\|_{H^p} \leq Ct^{-p/2} \|u_0\|, \quad \forall u_0 \in L^2(\Omega),
$$

where $C$ is independent of $t$ and $u_0$.

For each fixed $t > 0$, the adjoint $(S(t))^*$ is the time-$t$ solution operator of the equation obtained from (6.1) by switching the signs of the $b_i$’s, thus resulting in an equation is of the same type as (6.1). Therefore Lemma 6.1 applies to $(S(t))^*$ as well.

**Corollary 6.2.** For $p = 0, 1, 2$ and $t > 0$

$$
\max\{|S(t)u_0|, \|(S(t))^*u_0\|\} \leq Ct^{-p/2} \|u_0\|_{\tilde{H}^{-p}}, \quad \forall u_0 \in L^2(\Omega),
$$

where $C$ is independent of $t$ and $u_0$.

**Proof.** For $u_0 \in L^2(\Omega)$ we have

$$
\begin{align*}
|S(t)u_0|^2 &= \langle u_0, (S(t))^*S(t)u_0 \rangle \\
&\leq |u_0|_{\tilde{H}^{-p}} : \|(S(t))^*S(t)u_0\|_{H^p} \\
&\leq C t^{-p/2} \|u_0\|_{\tilde{H}^{-p}} : \|S(t)u_0\|
\end{align*}
$$

which concludes the proof. \qed

We conclude that the forward and adjoint operators $\mathcal{K}$ and $\mathcal{K}^*$ satisfy Condition 2.1[d]. We discretize the forward problem via the Galerkin method using continuous piecewise linear functions in space, and backward Euler in time. In order to verify Conditions 2.1[a] and [b] we resort to the literature on error estimates for fully discrete parabolic problems with irregular data [26, 27, 28]. For an extensive presentation of finite element methods for parabolic problems see [31]. For simplicity we focus on the two-dimensional case ($d = 2$). Let $T_{h_0}$ be a triangulation of the domain $\Omega$, and let $T_{h/2}$ be defined inductively to be the Goursat refinement of $T_h$ for all $h \in I = \{h_0/2^i : i \in \mathbb{N}\}$ (each triangle in $T \in T_h$ is cut along the three lines obtained by joining the midpoints of its edges). Note that $(T_h)_{h \in I}$ is a sequence of quasi-uniform triangulations. For $h \in I$ define

$$
\mathcal{V}_h = \{ f \in C(\bar{\Omega}) : \forall T \in T_h \ f|_T \text{ linear, and } f|_{\partial \Omega} \equiv 0 \}.
$$
We have $V_{h/2} \subset V_h \subset H^1(\Omega)$. Cf. [7], the following estimate holds for $v \in H^1_0 \cap H^2$:

$$\inf_{\phi_h \in V_h} \|v - \phi_h\| \leq c_3 h^{2-j} \|v\|_2, \quad j \in \{0, 1\}.$$  \hfill (6.11)

Let $t_m = mk$ with $m = 0, 1, \ldots, M_h$ so that $t_{M_h} = T$, with $k = k(h)$. The backward difference approximation $U^m$ to $u(t_m)$ is computed successively by

$$\begin{cases}
    \langle d_t U^{m+1}, \phi \rangle + a(t_{m+1}; U^{m+1}, \phi) = 0, & \forall \phi \in \mathcal{V}_h, \\
    U^0 = \pi_h u_0,
\end{cases} \hfill (6.12)$$

where $d_t U^{m+1} = k^{-1}(U^{m+1} - U^m)$. We will write for $u_0 \in L^2(\Omega)$

$$S^h(t_m)u_0 \overset{\text{def}}{=} U^m, \quad m = 1, 2, \ldots, M_h. \hfill (6.13)$$

The following estimate holds with $C$ independent of $h, m$ (see [26, 27]):

$$\|U^m - u(t_m)\| \leq C t_m^{-1} (h^2 + k) \|u_0\|, \quad m = 1, 2, \ldots. \hfill (6.14)$$

We will choose $k(h) = k_0 h^2$, therefore the discrete operator

$$K^h \overset{\text{def}}{=} S^h(T) \in L(\mathcal{V}_h) \hfill (6.15)$$

verifies Conditions 2.1[a, b]. The adjoint $(S^h(t_m))^*$ is the backward Euler Galerkin solution of the parabolic equation whose solution operator is $(S(t))^*$ (under the assumption that at each time step the linear solve is performed exactly). Therefore $(K^h)^*$ and $K^*$ also satisfy Conditions 2.1[a, b]. Conditions 2.1[c] is an easy consequence of the Bramble-Hilbert Lemma. Due to the fact that the time stepping method is only first order accurate in time, we would use four times less time steps for the coarser level, therefore a coarse-level residual computation is four times less costly than a fine-level residual computation. (the exponent $g$ in Section 5.1 is $g = 4$). We have thus verified that all results in Sections 4 and 5 apply to the problem (1.1) with $K, K^h$ defined as in (6.7) and (6.15). In fact this particular model problem has been the driving force behind the development and analysis of the MLAS preconditioner.

### 6.2. Some results on regularization for inverse parabolic problems.

In this section we present a few calculations that explain some of the qualitative behavior observed in the numerical solution of regularized inverse parabolic problems from Section 6.1. In Lemma 6.4 we show why the solution of the $L^2$-regularized inverse problem (1.1) with $K$ defined by (6.7) has a square variation that is comparable to that of the “true” initial value, if $\beta$ is chosen appropriately (see also Figure 1). This fact may seem surprising, given that $L^2$-regularization explicitly controls only size, and not derivatives. Example 6.5 is devoted to explaining another observed fact, namely that, when controlling the initial value for matching final-time measurements, recovery of intermediate states is of significantly better quality than that of the initial condition (see Figure 2 in Section 7). The immediate consequence is that, even if the recovered initial value may not be of acceptable quality (e.g. for localizing certain features), it still can be used successfully for improving predictions.

**Remark 6.3.** Let $\mathcal{X}, \mathcal{X}'$ be two Hilbert spaces such that $\mathcal{X}' \subseteq \mathcal{X}$ ($\mathcal{X}'$ does not have to be a Hilbert-subspace of $\mathcal{X}$). Furthermore, let $K \in L(\mathcal{X})$ be such that
\( \mathcal{K}(\mathcal{X}) \subseteq \mathcal{X} \), and assume that \( \mathcal{K}|_{\mathcal{X}'} \in L(\mathcal{X}') \) is compact. Then \( \sigma(\mathcal{K}|_{\mathcal{X}'}) \setminus \{0\} \) only consists of point-spectrum, therefore

\[
\sigma(\mathcal{K}|_{\mathcal{X}'}) \subseteq \sigma(\mathcal{K}) \cup \{0\}.
\]

**Lemma 6.4.** Assume \( \mathcal{K} \) is given by (6.7), and that the “data” \( f \) in (1.1) takes the form \( f = K u_0 + \varphi \), with \( \varphi \in \mathcal{X} \); furthermore, assume that \( u_0 \in H^1_0(\Omega) \). Then \( H_\beta \) is continuously invertible in \( H^1_0(\Omega) \) for all \( \beta > 0 \), and

\[
\|u^{\min} - u_0\|_1 \leq \|H_\beta\|^{-1}_H \cdot (|u_0|_1 + CT^{-\frac{1}{2}} \beta^{-1}\|\varphi\|) ,
\]

where \( C \) is the constant from (6.8).

**Proof.** Let \( \mathcal{X}' = H^1_0(\Omega) \). By Lemma 6.1 the operator \( H = \mathcal{K}^* \cdot \mathcal{K} \) is compact both in \( L(\mathcal{X}) \) and \( L(\mathcal{X}') \). Moreover, it is symmetric and positive definite in \( \mathcal{X} \); hence \( \sigma(H) \subset [0,\infty) \). Remark 6.3 implies that \( \sigma(H|_{\mathcal{X}'}) \subset [0,\infty) \), therefore \( H_\beta \) is invertible in \( L(\mathcal{X}') \) for all \( \beta > 0 \). Since

\[
u^{\min} - u_0 = (H_\beta)^{-1} (-u_0 + \beta^{-1}\mathcal{K}^* \varphi),
\]

the conclusion follows by (6.8). \( \square \)

It should be noted that, if \( \mathcal{K} \) is the solution operator for the heat equation, then \( H \) is symmetric with respect to the \( H^1_0 \) inner product as well, therefore \( \|H_\beta\|^{-1}_H \leq 1 \).

**Example 6.5.** Let \( S(t) \) to be the solution operator for the one-dimensional heat equation on \( \Omega = [0,2\pi] \) with zero boundary conditions. We will show that the relative distance between the “exact” solution \( S(t)u_0 \) and the recovered solution \( S(t)u^{\min} \) is decaying in time. For simplicity we restrict our attention to the case of unperturbed measurements, i.e. \( f = S(T)u_0 \).

Let \( \chi_k(x) = (2\pi)^{-1} e^{ikx} \), \( k \in \mathbb{Z} \), be the standard Fourier basis of \( L^2(\Omega) \), and let \( \hat{u} = (u, \chi_k) \) be the Fourier coefficients of a function \( u \in L^2(\Omega) \). We formulate our inverse problem in the space \( \mathcal{X} = \{ u \in L^2(\Omega) : \hat{u}_k + \hat{u}_{-k} = 0, \forall k \in \mathbb{Z} \} \) that is invariant under \( S(t) \). We prefer to express all quantities in terms of the basis \( \{ \chi_k \}_{k \in \mathbb{Z}} \) as opposed to a orthonormal basis of \( \mathcal{X} \). With the formulation and notation from Lemma 6.4 we have

\[
(H_\beta)^{-1} u_0 = \sum_{k=-\infty}^{\infty} \frac{\hat{u}_k}{1 + \beta^{-1} e^{-2\pi^2 k^2 \beta}} \chi_k ,
\]

where \( u_0 = \sum_{k=-\infty}^{\infty} \hat{u}_k \chi_k \). For simplicity we assume that \( \hat{u}_k \neq 0 \) if \( k \neq 0 \) (by construction \( \hat{u}_0 = 0 \)), and that \( \beta \ll 1 \). By (6.18)

\[
\frac{\|S(t)(u_0 - u^{\min})\|^2}{\|S(t)u_0\|^2} = \frac{\sum_{k=-\infty}^{\infty} (1 + \beta^{-1} e^{-2\pi^2 k^2})^{-1} e^{-2\pi^2 k^2} |\hat{u}_k|^2}{\sum_{k=-\infty}^{\infty} e^{-2\pi^2 k^2} |\hat{u}_k|^2} .
\]

Since for moderate and large values of \( k \) the eigenvalues \( \mu_{k,\beta} = (1 + \beta^{-1} e^{-2\pi^2 k^2})^{-1} \) of \( (H_\beta)^{-1} \) are close to 1, the error formula (6.18) implies that the corresponding components of \( u_0 \) are not well approximated by those of \( u^{\min} \); this results in a fairly large relative difference between \( S(t)u^{\min} \) and \( S(t)u_0 \) for small \( t \). A slight increase in \( t \) strongly reduces the size of the high-frequency components both in \( S(t)u^{\min} \) and \( S(t)u_0 \), and the eigenvalues of \( (H_\beta)^{-1} \) associated to the remaining low-frequency components are of size \( O(\beta) \). As a result, the expression on the right-hand side of (6.20) will decrease over a short time interval from \( O(1) \) to \( O(\beta^2) \). A
rough estimate, obtained by separating the first nonzero components from the rest in the numerator of the expression in (6.20), gives

\[
\frac{|S(t)(u_0 - u_{\min})|^2}{|S(t)u_0|^2} \leq (\beta e^{2T})^2 + \sum_{|k| \geq 2} e^{-2t(k^2 - 1)} |\hat{u}_k|^2 \\
\leq (\beta e^{2T})^2 + \frac{e^{-6t} \|u_0\|^2}{|\hat{u}_1|^2 + |\hat{u}_{-1}|^2 + e^{-6t} \|u_0\|^2},
\]

which clearly show the decay to be exponentially in \( t \).

6.3. Construction of local restriction operators. In this section we show a way to construct local restriction operators that satisfy conditions \([e, f]\) in Theorem 4.3. Local operators, whether defined in an explicit or implicit way, are more attractive for computational purposes than orthogonal projections, which require inverting the mass matrix on the coarse level. The local restriction operator inherited from classical multigrid, defined as the adjoint of the natural interpolation with respect to a mesh-dependent inner product approximating \((\cdot, \cdot)_{L^2}\), typically satisfies the approximation property (4.6) only up to order \( p = 1 \) (see Example 6.7). A local, explicit, symmetric restriction operator that approximates the identity-operator in negative-index norm up to higher order is defined in [5]. More precisely, for continuous piecewise linear elements the inequality (4.6) is verified for any \( p \in [0, 3/2] \), and a similar result holds for quadratics. However, our estimates in Section 4 require, e.g. for linear elements, that (4.6) holds for \( p = 2 \). The main result of this section, Theorem 6.6, shows how to construct restriction operators with needed approximation properties.

For simplicity we restrict our exposition to the two-dimensional case. Consider the quasi-uniform triangulations \((T_h)_{h \in I}\) defined in Section 6.1. For a triangle \( T \in T_h \) denote by \( P^s(T) \) the space of polynomials of total degree \( \leq s - 1 \) restricted to \( T \), and let \( s \geq 2 \) be fixed. Let \( V_h = V_h^s \) be the Lagrange finite element spaces

\[
V_h = \{ u \in C^0(\Omega) : u|_T \in P^s(T) \text{ for each } T \in T_h, \ u|_{\partial \Omega} = 0 \}.
\]

Denote by \( (\Phi_i^h)_{1 \leq i \leq N_h} \) be the Lagrange nodal basis of \( V_h \).

Theorem 6.6. Let \( R_{2h}^h \in L(V_h, V_{2h}) \) be a restriction operator that satisfies the stability condition (4.5), and is local in the following sense: for each \( i = 1, \ldots, N_h \) there exists a set \( N_i^h \) which is a union of triangles in \( T_h \), such that

(a) \( \text{supp}(R_{2h}^h \Phi_i^h) \cup \text{supp}(\Phi_i^h) \subset N_i^h \);

(b) \( \text{no triangle } T \in T_h \text{ lies in more than } M \text{ of the } N_i^h \text{'s, with } M \text{ independent of } h \);

(c) \( (I - R_{2h}^h) \Phi_i^h \perp P^s(N_i^h) \);

(d) \( \text{diam}(N_i^h) \leq Kh \) with \( K \) independent of \( h \).

Then there exists a constant \( C = C(K, M, s) \), independent of \( h \), such that

\[
\|(I - R_{2h}^h)u\|_{H^{-s}(\Omega)} \leq Ch^s \|u\|, \text{ for } u \in V_h.
\]

Proof. Let \( u = \sum_{i=1}^{N_h} u_i \Phi_i^h \) be a function in \( V_h \) and \( v \in H_0^s(\Omega) \). Then

\[
\langle (I - R_{2h}^h)u, v \rangle = \sum_{i=1}^{N_h} u_i \int_{N_i^h} (I - R_{2h}^h)\Phi_i^h(x) v(x) dx.
\]
For any \( q \in \mathcal{P}^s(\mathcal{N}_h^b) \)
\[
\left| \int_{\mathcal{N}_h^b} (I - R_{2h}^b) \Phi_i^b(x) v(x) \, dx \right| \overset{(b)}{=} \left| \int_{\mathcal{N}_h^b} (I - R_{2h}^b) \Phi_i^b(x) (v(x) - q(x)) \, dx \right| 
\leq \| (I - R_{2h}^b) \Phi_i^b \|_{L^2(\mathcal{N}_h^b)} \cdot \| v - q \|_{L^2(\mathcal{N}_h^b)} .
\]

The Bramble-Hilbert Lemma and (d) imply
\[
\left| \langle (I - R_{2h}^b)u, v \rangle \right| \overset{(4.5)}{=} C \inf_{q \in \mathcal{P}^s(\mathcal{N}_h^b)} \| v - q \|_{L^2(\mathcal{N}_h^b)} \leq C(Kh)^s \| \Phi_i^b \| \cdot \| v \|_{H^s(\mathcal{N}_h^b)}.
\]

Therefore
\[
\left| \langle (I - R_{2h}^b)u, v \rangle \right| \leq C(Kh)^s \sum_{i=1}^{N_h} |u_i| \cdot \| \Phi_i^b \| \cdot \| v \|_{H^s(\mathcal{N}_h^b)}
\]
\[
\leq CK^s h^s \left( \sum_{i=1}^{N_h} u_i^2 \| \Phi_i^b \|^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{N_h} |v|^2_{H^s(\mathcal{N}_h^b)} \right)^{\frac{1}{2}}
\]
\[
\overset{(b)}{\leq} (CK^s M^{\frac{1}{2}}) h^s \left( \sum_{i=1}^{N_h} u_i^2 \| \Phi_i^b \|^2 \right)^{\frac{1}{2}} \cdot |v|_{H^s(\Omega)}
\]
\[
\leq C_1 h^s |u| \cdot |v|_{H^s(\Omega)} ;
\]

for the last inequality we used the quasi-uniformity of the triangulation. It follows that
\[
\| (I - R_{2h}^b)u \|_{H^s(\Omega)} = \sup_{v \in H^s_0(\Omega)} \frac{|\langle (I - R_{2h}^b)u, v \rangle|}{|v|_{H^s(\Omega)}} \leq C_1 h^s |u| .
\]

\[\Box\]

**Example 6.7** (standard restriction operator on a uniform grid). Let \( \Omega \) be the unit square \((0,1) \times (0,1) \subset \mathbb{R}^2 \) and \( (T_{2^{k-1}})_{k=0,1,...} \) the uniform three-line meshes\(^1\) with each of the small triangles in \( T_{2^{-k}} \) having side length \( h_k = 2^{-k} \). Obviously \( T_{2^{-(k+1)}} \) is the Goursat refinement of \( T_{2^{-k}} \). Let \( s = 2 \) in (6.21) and define our spaces to consist of continuous, piecewise linear, double-periodic functions:

(6.23) \( \mathcal{V}_{h_k}^{\text{per}} = \{ u \in \mathcal{V}_{h_k} : u(x_1,0) = u(x_1,1), u(0,x_2) = u(1,x_2) \} \).

We extend the functions in \( \mathcal{V}_{h_k}^{\text{per}} \) to \( \mathbb{R}^2 \) in an obvious way. Let \( \{ P_i^k : i = 1, \ldots, 2^k \} \) be the vertices of \( T_{2^{-k}} \) not lying on \( \{(x_1, x_2) : x_1 = 1 \text{ or } x_2 = 1\} \), and denote by \( \phi_i^{k,}(x) \) the corresponding nodal basis functions. We introduce, as in [7], the mesh-dependent inner products on \( \mathcal{V}_{h_k}^{\text{per}} \):

(6.24) \[ \langle u, v \rangle_{k} \overset{\text{def}}{=} \frac{h_k^2}{2} \sum_{i=1}^{2^k} u(P_i^k) v(P_i^k), \text{ for } u, v \in \mathcal{V}_{h_k}^{\text{per}} . \]

\(^1\)The **three-line mesh** is obtained by dividing the square into equally sized squares with sides parallel to the coordinate axes, and by further cutting each little square along its slope-one diagonal.
It should be noted that \( \langle \cdot, \cdot \rangle_k \) is obtained by applying the second order correct cubature rule

\[
\int_{\Delta P_1 P_2 P_3} u(x) \, dx \approx \text{area}(\Delta P_1 P_2 P_3) \sum_{i=1}^{3} u(P_i)
\]
to the \( L^2 \)-inner product, hence

\[
\langle u, 1 \rangle_k = \langle u, 1 \rangle = \int_{\Omega} u(x) \, dx, \quad u \in \mathcal{V}_{h_k}^{\text{per}}.
\]
The standard restriction operator \( R_{k-1}^k : \mathcal{V}_{h_k}^{\text{per}} \to \mathcal{V}_{h_{k-1}}^{\text{per}} \) is defined by the equation

\[
\langle R_{k-1}^k u, v \rangle_{k-1} = \langle u, v \rangle_k, \quad \forall u \in \mathcal{V}_{h_k}^{\text{per}}, \ v \in \mathcal{V}_{h_{k-1}}^{\text{per}}.
\]

A simple calculation shows that

\[
\langle R_{k-1}^k u, (P_i^k) \rangle = \frac{1}{4} u(P_i^k) + \frac{1}{8} \sum_{\alpha=1}^{6} u(P_{i_\alpha}^k),
\]

where \( P_{i_1}^k, \ldots, P_{i_6}^k \) are the vertices directly connected to \( P_i^k \) in the graph of \( T_{h_k} \) (periodicity assures that every vertex has exactly six neighbors in the graph). This shows that

\[
\text{supp}(R_{k-1}^k \Phi_i^k) = \bigcup_{\alpha=1}^{6} \text{supp}(\Phi_{i_\alpha}^k) \overset{\text{def}}{=} N_i^k.
\]

Obviously \( N_i^k \) verifies conditions (a), (b) and (d) from Theorem 6.6 (\( \text{diam}(N_i^k) \leq 8\sqrt{2} h_k \)). It remains to verify (c). By construction, if \( u \in \mathcal{V}_{h_k}^{\text{per}} \) then

\[
\langle (I - R_{k-1}^k)u, 1 \rangle = \langle (I - R_{k-1}^k)u, 1 \rangle_k = \langle u, 1 \rangle_k - \langle R_{k-1}^k u, 1 \rangle_k = 0.
\]

Orthogonality on linear functions follows by symmetry: fix \( i \) and define the reflection with respect to \( P_i^k \) to be \( x \mapsto r_i^k \circ P_i^k \). Then \( N_i^k \) is \( r_i^k \)-invariant \( (r_i^k(N_i^k) = N_i^k) \) and \( \Phi_i^k \) is \( r_i^k \)-symmetric (i.e. \( \Phi_i^k \circ r_i^k = \Phi_i^k \)), and so is \( (I - R_{k-1}^k) \Phi_i^k \). If we denote by \( p_1 \) and \( p_2 \) the projections on the two coordinate axes, then the functions

\[
x \mapsto (I - R_{k-1}^k) \Phi_i^k(x) \cdot p_\alpha(x - P_i^k), \quad \alpha = 1, 2,
\]

are \( r_i^k \)-antisymmetric (if \( T : A \to A \) is a bijection and \( f : A \to \mathbb{C} \), then \( f \) is called \( T \)-antisymmetric if \( f \circ T + f = 0 \); hence

\[
\int_{N_i^k} (I - R_{k-1}^k) \Phi_i^k(x) \cdot p_\alpha(x - P_i^k) \, dx = 0.
\]

The equalities (6.30) and (6.31) can be rewritten as

\[
(I - R_{k-1}^k) \Phi_i^k \perp \mathcal{P}^2(N_i^k).
\]

Theorem 6.6 implies that

\[
\| (I - R_{k-1}^k)u \|_{H^{-2}(\Omega)} \leq C h^2 \| u \| \quad \text{for } u \in \mathcal{V}_{h_k}^{\text{per}}.
\]

For unstructured triangular meshes the restriction \( R_{k-1}^k \) still satisfies

\[
(I - R_{k-1}^k) \Phi_i^k \perp 1,
\]

but typically \( (I - R_{k-1}^k) \Phi_i^k \) is not orthogonal on all linear functions; hence in general we can only conclude that

\[
\| (I - R_{k-1}^k)u \|_{H^{-1}(\Omega)} \leq C h \| u \| \quad \text{for } u \in \mathcal{V}_{h_k}^{\text{per}}.
\]
Remark 6.8. We note that a result similar to Theorem 6.6 holds for tensor-product finite elements, and that the same argument as in Example 6.7 can be used to show that the standard restriction operator used for continuous, piecewise linear tensor-product finite elements in $\mathbb{R}^d$ also satisfies the optimal negative-index norm estimates (6.33), provided the mesh is uniform and the functions are periodic.

Example 6.9. We return to the framework of Section 6.1, that is, the case of an unstructured quasi-uniform triangular grid on a polygonal domain $\Omega \subset \mathbb{R}^2$, with $V_h$ consisting of continuous piecewise linear functions ($s = 2$). We define a local restriction operator $R_{2h}^h$ by explicitly enforcing condition (c) in Theorem 6.6 in the following way: for each fine nodal basis function $\Phi_{M}^1$ associated to a coarse mesh vertex $A$, choose a coarse-mesh triangle $\Delta A_1 A_2 A_3$ with all vertices in the interior of $\Omega$, such that, in the fine mesh, $M$ is adjacent or equal to one of $A_1$, $A_2$ or $A_3$. (E.g., if $M$ is a coarse node, choose $A_1 = M$; if $M$ is the midpoint of an interior edge, let $A_1, A_2$ the vertices of that edge; if $M$ is the midpoint of an edge connecting a coarse-mesh vertex $A_1$ to the boundary, then again let $A_2, A_3$ two neighbors of $A_1$ that are adjacent to each other.) Let $(\Phi_{i}^{2h})_{i=1,2,3}$ be the nodal variables in $V_{2h}$ that are associated to $A_{1,2,3}$ respectively. We define

\begin{equation}
R_{2h}^h \Phi_M^h = \sum_{j=1}^{3} \alpha_j \Phi_{j}^{2h}, \tag{6.35}
\end{equation}

with $(\alpha_j)_{j=1,2,3}$ chosen so that

\begin{equation}
(I - R_{2h}^h) \Phi_M^h \perp \mathcal{L}, \quad \forall \mathcal{L} \in \mathcal{P}^2. \tag{6.36}
\end{equation}

Let $(L_i)_{i=1,2,3}$ be the basis of $\mathcal{P}^2$ obtained by linearly extending $\Phi_{i}^{2h}|_{\Delta A_1 A_2 A_3}$ to $\mathbb{R}^2$. The system (6.36) translates into

\begin{equation}
\sum_{j=1}^{3} \langle \Phi_{j}^{2h}, L_i \rangle \alpha_j = \langle \Phi_M^h, L_i \rangle, \quad i = 1, 2, 3. \tag{6.37}
\end{equation}

Under additional regularity assumptions on the meshes $(T_h)_{h \in I}$ we can show that the system (6.37) is diagonally dominant, therefore nonsingular, and that the resulting restriction operator is uniformly bounded, thus verifying condition (e) in Theorem 4.3. More precisely we can show that for every node $M$ in the fine mesh the system (6.37) has a solution $(\alpha_j)_{j=1,2,3}$ with $|\alpha_j| \leq C$, where $C$ is independent of $h$. This implies that the restriction operators are uniformly bounded. For details see Appendix A. The other hypotheses in Theorem 6.6 are satisfied by construction, therefore

\begin{equation}
\|(I - R_{2h}^h)u\|_{H^{-2}(\Omega)} \leq Ch^2 \|u\|, \quad \forall u \in V_h. \tag{6.38}
\end{equation}

7. A NUMERICAL EXAMPLE

We present a simple numerical example that illustrates the application of the multilevel preconditioner MLAS to the inverse problem defined in Section 6.1.

Data: In this example $\Omega = (0,1)$, and the coefficients are $a_{11}(x,t) = 4 \cdot 10^{-3}$, $b_1(x,t) = 0.4$, $c(x,t) = 0.05$. The exact initial value $u_0$ to be recovered is a Gaussian with $\mu = 0.75$ and $\sigma = 0.03$. The final time is $T = 1$. We consider six uniform grids on $(0,1)$ with the base grid having $h_0 = 1/200$, and for sequent grids $h_k = 2^{-k} h_0$, $i = 1, 2, 3, 4, 5$. We approximate the solution by its Galerkin projection
on the space of continuous piecewise linear functions, and we use backward Euler for
time discretization. At each level the time step is uniform, with the base level time
step of \( dt_0 = 1/100 \). Due to the fact that the approximation in time is only first
order, we refine the time step by \( dt_{k+1} = dt_k/4 \). For each of the levels \( k = 0, 1, 2, 3, 4 \)
we solve the inverse problem using CG preconditioned by MLAS with \( 1 \leq i \leq 6 - k \)
levels of refinement (one level of refinement simply means unpreconditioned CG);
that is, with each of the base base cases we refine the problem until the maximal
resolution of \( h_5 = 1/6400 \) is reached. For each of \( \beta = 10^{-3}, 10^{-4}, 10^{-6} \) we test
the V-cycle and the W-cycle preconditioner (MLAS) from Section 5. MLAS is
modified so that, at the finest level, only one recursive call to the coarse space
solver is performed at each iteration; therefore no residual is computed at the finest
level in the preconditioner. This way comparison between the W-cycle and the
V-cycle preconditioner is fair (in fact the two are identical if only two levels are
used). In Figure 1 we show the time-evolution of the system starting from the
"exact" initial value versus the recovered initial value at times 0, 0.5 and 1 for
\( \beta = 10^{-3} \).

**Targets:** Due to the fact that the Hessian is never formed, it is difficult to evaluate
directly the spectral distance between the inverse of the Hessian and the tested
preconditioners. We evaluate the quality of the preconditioners indirectly by looking
at the number of preconditioned CG iterations needed to obtain a residual whose
norm is \( rtol = 10^{-12} \) times smaller than the norm of the right-hand side. We use the
zero-function as the initial guess. Our goal is to verify the following consequences
of Theorems 5.2 and 5.4:
1. V-cycle: for a fixed base level the number of iterations \( N_{it} \) is independent of the number of refinement levels (provided the base level is sufficiently fine); \( N_{it} \) decreases with increasing base-level resolution;

2. W-cycle: given a sufficiently fine base level, \( N_{it} \) decreases with increasing number of refinement levels; in fact \( N_{it} \) depends only on the finest resolution (if at least two levels are used).

Furthermore we verify that unpreconditioned CG solves the problem in a number of iterations that is independent on mesh size, and increases slowly with \( \beta \to 0 \). We consider two ways of measuring the success of solution process. An absolute measuring unit is the number of flops needed for a direct (forward) solve. We call the cost of the inverse solve measured in forward solves the efficiency factor, denoted by \( I/F \):

\[
I/F = \frac{\text{cost of inverse solve}}{\text{cost of forward solve}}.
\]

For unpreconditioned CG each iteration requires a matrix-vector multiplication, which costs approximately two forward solves. Therefore the CG solve with \( N_{it} \) iterations costs about \( 2N_{it} \) forward solves; the whole solution process includes also a gradient computation, therefore the efficiency factor for the unpreconditioned CG is \( \approx 2(N_{it} + 1) \). In the tables 1 - 6 we show for each case the number of iterations, and in parentheses the cost of the solution process for the inverse problem computed in 'forward solves'. An alternative way of measuring success, that inherently takes into account the difficulty of the problem, is by using the work of unpreconditioned CG as a measuring unit (since it is quasi-proportional with the space-time size of the problem). E.g., if the measuring time \( T \) is large, then nearly all information about the initial value we are trying to recover is lost, and essentially we are inverting the identity operator. CG captures this fact by solving the problem in few iterations.

**Results and conclusions:** A visual inspection of the recovered solution in Figure 1 (dotted line at time 0), shows that the recovered \( u_{\text{min}} \) is a smooth curve, as predicted by Lemma 6.4. At the same time it shows that, despite the recovery of the initial value itself not being very accurate (\( u_{\text{min}} \) exhibits low frequency oscillations), \( S(0.5)u_{\text{min}} \) is very close to \( S(0.5)u_0 \); note that no measurements were taken at time \( t = 0.5 \). The plot in Figure 2 shows an exponential-like decay of the relative error \( \|S(t)(u_{\text{min}} - u_0)\|/\|S(t)u_0\| \), as expressed in Example 6.5.

The left column in the tables 1-6 shows the number \( N \) of intervals used for discretization; the column headers indicates the number of refinement levels used. E.g., for \( N = 400 \) and three levels of refinements the finest grid will have \( 2^{3-1}N = 1600 \) intervals. The entries in the first column support the assertion that unpreconditioned CG solves the inverse problem in a number of iterations which is mesh-independent. The two targets mentioned above are verified in the runs with \( \beta = 10^{-3} \) (Tables 1 and 2), and \( \beta = 10^{-4} \) (Tables 3 and 4). Mostly important, we notice the number of MLAS-preconditioned CG iterations to depend only on the finest-level resolution, independent of the number of levels. In particular, this confirms the conclusion of Theorem 5.4, thus showing MLAS to have the same approximation quality as TLAS. A stronger decrease in regularization parameter, namely \( \beta = 10^{-6} \), causes the base case with \( N = 200 \) to become unacceptably coarse. Specifically we see the V-cycle preconditioned CG taking a fairly large number of
iterations, thus making this solution process as expensive as the unpreconditioned CG. Also as a result of the base case being too coarse we see the W-cycle preconditioner losing its positive-definiteness. Another interesting fact is noticed when using $N = 400$ as a base case. Although the V-cycle preconditioner stagnates after $\approx 14$ iterations (Table 5), the W-cycle algorithm shows a certain ability of self-correction, thus overcoming the suboptimal choice of base level. The correct base level for $\beta = 10^{-6}$ seems to be the one with $N = 800$, point at which predicted behavior sets in. In terms of efficiency at a given resolution, the I/F factor readings suggest that MLAS is mostly efficient when maximizing the number of levels, constrained by having the base level acceptably coarse. Using this strategy we found MLAS-preconditioned CG to work up to four times faster than unpreconditioned CG (e.g. Table 6, $N = 800$ with 4 levels has an I/F of 17.7 compared to a predicted I/F of 72 for unpreconditioned CG).

**Figure 2.** Time evolution of relative error $||S(t)(u_{\text{min}} - u_0)|| / ||S(t)u_0||$.

**Table 1.** Iteration count (I/F) for the V-cycle; $\beta = 10^{-3}$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>15</td>
<td>11</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>400</td>
<td>16</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>800</td>
<td>16</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>1600</td>
<td>16</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>3200</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>
Table 2. Iteration count (I/F) for the W-cycle; $\beta = 10^{-3}$.

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>15 (32.3)</td>
<td>11 (61.1)</td>
<td>9 (29.6)</td>
<td>7 (19.4)</td>
<td>6 (16.2)</td>
<td>5 (13.7)</td>
</tr>
<tr>
<td>400</td>
<td>16 (34.1)</td>
<td>9 (48)</td>
<td>7 (22.8)</td>
<td>6 (16.8)</td>
<td>5 (14.4)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>16 (34)</td>
<td>7 (38)</td>
<td>6 (19.8)</td>
<td>5 (14.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>16 (34)</td>
<td>6 (32)</td>
<td>5 (16.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3200</td>
<td>17 (36)</td>
<td>5 (26.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Iteration count (I/F) for the V-cycle; $\beta = 10^{-4}$.

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>21 (44.4)</td>
<td>13 (82.2)</td>
<td>13 (35.1)</td>
<td>13 (28.9)</td>
<td>13 (28.1)</td>
<td>13 (28)</td>
</tr>
<tr>
<td>400</td>
<td>20 (42.1)</td>
<td>10 (61.3)</td>
<td>11 (29.5)</td>
<td>11 (24.7)</td>
<td>11 (24.1)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>21 (44)</td>
<td>8 (47.5)</td>
<td>8 (21.8)</td>
<td>8 (18.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>21 (44)</td>
<td>6 (36.2)</td>
<td>6 (16.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3200</td>
<td>21 (44)</td>
<td>5 (31.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Iteration count (I/F) for the W-cycle; $\beta = 10^{-4}$.

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>21 (44.4)</td>
<td>13 (82.2)</td>
<td>15 (49.6)</td>
<td>10 (27.5)</td>
<td>8 (21)</td>
<td>6 (16.1)</td>
</tr>
<tr>
<td>400</td>
<td>20 (42.1)</td>
<td>10 (61.3)</td>
<td>7 (25.5)</td>
<td>7 (19.5)</td>
<td>5 (13.9)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>21 (44)</td>
<td>8 (47.5)</td>
<td>6 (21)</td>
<td>6 (17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>21 (44)</td>
<td>6 (36.2)</td>
<td>5 (18)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3200</td>
<td>21 (44)</td>
<td>5 (31.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Iteration count (I/F) for the V-cycle; $\beta = 10^{-6}$.

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>27 (56.7)</td>
<td>19 (169)</td>
<td>23 (65.6)</td>
<td>25 (54.3)</td>
<td>22 (46.3)</td>
</tr>
<tr>
<td>400</td>
<td>32 (66.2)</td>
<td>15 (117.4)</td>
<td>14* (49.6)</td>
<td>13* (29.5)</td>
<td></td>
</tr>
<tr>
<td>800</td>
<td>34 (70)</td>
<td>9 (73.2)</td>
<td>10 (29.5)</td>
<td>10 (22.9)</td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>34 (70)</td>
<td>7 (57.2)</td>
<td>7 (21.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3200</td>
<td>35 (72)</td>
<td>6 (45.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
* stagnated at a low residual (corresponding to $rtol = 10^{-11}$) before converging.

Table 6. Iteration count (I/F) for the W-cycle; $\beta = 10^{-6}$.

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>27 (56.7)</td>
<td>19 (169)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>400</td>
<td>32 (66.2)</td>
<td>15 (117.4)</td>
<td>10 (39.3)</td>
<td>9 (25.6)</td>
<td>7 (18.8)</td>
</tr>
<tr>
<td>800</td>
<td>34 (70)</td>
<td>9 (73.2)</td>
<td>8 (30.6)</td>
<td>6 (17.7)</td>
<td>-</td>
</tr>
<tr>
<td>1600</td>
<td>34 (70)</td>
<td>7 (57.2)</td>
<td>6 (23.7)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3200</td>
<td>35 (72)</td>
<td>6 (45.7)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* multilevel preconditioner not positive definite.
The vicinity of each vertex is convex. The next condition states that the vicinity is somewhat balanced around the vertex.

**Condition A.1.** There are two constants $c_1, c_2 \geq 1$, independent of $h$, such that:

1. For any vertex $A$ of $T_h$, $D_h^A \leq c_1 d_h^A$;
2. For any neighboring vertices $A, B$ of $T_h$, we have $c_2^{-1} d_h^A \leq d_h^B \leq c_2 d_h^A$.

It is easy to verify that (i) implies (ii) in Condition A.1, with $c_2 = c_1$; however, in general, $c_2 \leq c_1$, and here it is advantageous to keep track of the individual constants. E.g., for a regular mesh (we call a mesh regular if all triangles are equilateral) $c_1 = 2/\sqrt{3}$, and $c_2 = 1$, and for the usual three-lines mesh (obtained from a rectangular grid with square elements by cutting each square along its slope-one diagonal) we have $c_1 = 2$, $c_2 = 1$. We wish to point out that, in case the triangulations are obtained by Goursat refinement, then the constants $c_1, c_2$ in Condition A.1 propagate from $T_h$ to finer meshes; in other words, finding the optimal constants in Condition A.1 only requires looking at the coarsest grid. For a point $A$ and $\rho > 0$ denote by $\psi_{A, \rho}$ the “cone-hat” function

$$\psi_{\rho}^A(x) = \begin{cases} 1 - \|x - A\|/\rho, & \text{for } x \in B_{\rho}(A) \\
0, & \text{otherwise} \end{cases}$$

It follows that a nodal basis function $\Phi_A$ satisfies

$$\psi_{\rho}^A \leq \Phi_A \leq \psi_{D_h^2}^A \quad (A.1)$$

The convexity of the $K_{2h}^A$ implies that $L_i \geq 0$ on supp($\Phi_{2h}^A$), therefore

$$a_{ii} = \langle \Phi_{2h}^A, L_i \rangle \geq \int_{B_{d_i}(A_i)} \psi_{d_i}^A L_i = \frac{\pi d_i^2}{3} \quad (A.2)$$

where $d_i = d_{2h}^A$, $D_i = D_{2h}^A$, $i = 1, 2, 3$. For $i \neq j$

$$a_{ij} = \langle \Phi_{2h}^A, L_i \rangle \leq \int_{L_i, \geq 0} \psi_{d_j}^A L_i + \int_{L_i, \leq 0} \psi_{d_j}^A L_i \quad (A.3)$$

Let $\delta_i$ be the distance from $A_i$ to the opposite side in $\Delta A_1 A_2 A_3$. Then $\delta_i \geq d_i \geq c_2^{-1} d_j$, and

$$\int_{L_i, \geq 0} \psi_{d_i}^A L_i = \int_{-\pi}^{\pi/2} \int_0^{D_j} (1 - \rho / D_j) \frac{\rho^2 \cos \theta}{d_i} d\rho = \frac{D_j^3}{6 \delta_i} \quad (A.4)$$

Similarly

$$\int_{L_i, \leq 0} \psi_{d_j}^A L_i = - \int_{L_i, \leq 0} \psi_{d_j}^A L_i = \frac{D_j^3}{6 \delta_i} \quad (A.5)$$
Hence

\[(A.6) \quad a_{ij} \leq \frac{D^j_i - D^j_j}{6 \delta_i} \leq c_2 \frac{c_1^3 - 1}{6} d^2_j.\]

Since \(-a_{ij}\) satisfies the same inequality we conclude that \((A.6)\) holds with \(|a_{ij}|\) in place of \(a_{ij}\). We should note that the estimate \((A.6)\) can be quite inaccurate even for “tame” meshes such as the three-lines mesh mentioned previously, where symmetry implies that \(a_{ij} = 0\) for \(i \neq j\), while \((A.6)\) is insufficient to prove that \(\sum_{j \neq i} a_{ij} \cdot c_2 c_3 \leq \sum_{j \neq i} d_{ij}^2\).

Since \(-a_{ij}\) satisfies the same inequality we conclude that \((A.6)\) holds with \(a_{ij}\) in place of \(a_{ij}\). We should note that the estimate \((A.6)\) can be quite inaccurate even for “tame” meshes such as the three-lines mesh mentioned previously, where symmetry implies that \(a_{ij} = 0\) for \(i \neq j\), while \((A.6)\) is insufficient to prove that \(\sum_{j \neq i} a_{ij} \cdot c_2 c_3 \leq \sum_{j \neq i} d_{ij}^2\).

If we divide the \(i\)th equation of \((6.37)\) by \(a_{ii}\), then the resulting equivalent system will be represented by a matrix of the form \(I - B\), \(B = (b_{ij})_{1 \leq i,j \leq 3}\), with \(b_{ij} = (\delta_{ij} - 1) a_{ij} / a_{ii}\). If

\[(A.7) \quad \max_{i=1}^{3} \sum_{j \neq i} |a_{ij}| \leq (1 - \gamma) a_{ii},\]

for some \(\gamma \in (0, 1)\), then \(|B|_{\infty} \leq 1 - \gamma\); hence

\[(A.8) \quad \|(I - B)^{-1}\|_{\infty} \leq \sum_{k=0}^{\infty} |B|_{\infty}^k = \gamma^{-1}.\]

We conclude our discussion with

**Lemma A.2.** If the meshes \((T_h)_{h \in T}\) satisfy Condition A.1 such that

\[(A.9) \quad \gamma \overset{\text{def}}{=} 1 - \frac{c_2 (c_1^3 - 1)}{\pi} > 0,\]

then there exists a constant \(C\) independent of \(h\) such that for each triangle-vertex of \(T_h\) and choice of coarse-mesh triangle \(\Delta A_1 A_2 A_3 \in T_{2h}\), the solution \((\alpha_j)_{j=1,2,3}\) of the system \((6.37)\) satisfies \(|\alpha_j| \leq C, j = 1, 2, 3.\)

**Proof.** Quasi-uniformity of the meshes imply that for each vertex \(M\) of \(T_h\)

\[c h \leq d^h_M \leq C h,\]

for some constants \(c, C\) independent of \(h\). Therefore, due to the proximity of \(M\) to the chosen triangle \(\Delta A_1 A_2 A_3 \in T_{2h}\), the right-hand side of the modified system \((6.37)\) satisfies \(|\langle \Phi^h_{M}, L_i \rangle | / a_{ii} \leq C'\). In light of the estimates \((A.2)\) and \((A.6)\) condition \((A.9)\) implies \((A.7)\). By \((A.8)\) we obtain \(|\alpha_j| \leq C' / \gamma, j = 1, 2, 3.\)

The condition \((A.9)\) is not necessary for the system \((6.37)\) to be nonsingular, as seen from the three-line mesh example. However, it is not very restrictive either, given that for a regular mesh the value of \(\gamma\) is fairly large (\(\gamma \approx 0.82\)). Hence any mesh that is locally “alike” to a regular mesh will likely satisfy \((A.9)\).

**Appendix B. Notation summary**

For quick reference we provide a list of notations that are used throughout the article, and the place where each is first used or defined, where applicable:
A compact operator from $L^2(\Omega)$ to $L^2(\Omega)$

the operator norm of $K$ (from $L^2(\Omega)$ to $L^2(\Omega)$)

Section 1

$\|K\|$ $K^*K$ (1.3)

$I + \beta^{-1}H$ (1.3)

$V_h$ approximation space Section 2

$h_0$ coarsest $h$ that can be chosen as base case

$K_h$ discretization of $K$ in $V_h$ Section 2

$\pi_h$ $L^2$-projection onto $V_h$ Section 2

$d_H(T_1, T_2)$ spectral distance between $T_1$ and $T_2$ Def. 3.1

$N_H(X)$ $2H - XHX$ (3.18)

$d_h$ $d_{V_h}$ Section 4

$R_{2h}$ restriction operator Section 4.2

$L_{\beta}^h$ two-level preconditioner (4.1), (4.7)

$G_h(T) = T\pi_{2h} + (I - \pi_{2h})$ (5.2)

$K_h^\beta$ MLAS preconditioner (5.3)

$e_h^\beta = d_h(K_h^\beta, (H_h^\beta)^{-1})$ (5.7)

$S(t)$ solution operator for a parabolic PDE Section 6

ACKNOWLEDGMENT

The authors thank Ridgway Scott, Volkan Akcèelik, George Biros, Omar Ghattas, Judith Hill, Bart van Bloemen Waanders, and Rouben Rostamian for helpful discussions, and the anonymous referee for valuable comments.

REFERENCES


Sandia National Laboratories, Albuquerque, New Mexico 87125

Current address: Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250
E-mail address: draga@math.umbc.edu

Department of Computer Science, University of Chicago, Chicago, Illinois 60637
E-mail address: t-dupont@uchicago.edu

\footnote{Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.}