V. Korneev[♯,†], *L. Xanthis*[†] *and I. Anoufriev*[♯]

Hierarchical and Lagrange *hp* discretizations and fast domain decomposition solvers for them [1] [2]

[♯]*St. Petersburg State Technical University, Russia*, and [†]*Harrow School of Computer Science, University of Westminster, UK*

Vadim.Korneev@pobox.spbu.ru, L.S.Xanthis@wmin.ac.uk, anufriev@amd.stu.neva.ru

We present fast *DD* (domain decomposition) algorithms for systems of algebraic equations, resulting from *hp* finite element discretizations of 2-nd order elliptic equations. Main attention is paid to the most time consuming component of such algorithms, which is solving local Dirichlet problems on subdomains of decomposition. Different types of *hp* discretizations, *i.e.,* by means of hierarchical and Lagrange (nodal) elements, including incomplete elements, are studied. We present *DD*-type and multilevel almost optimal iterative solvers for these problems.

Introduction

In the last decade, *DD* (domain decomposition) *hp* finite element solvers for second order elliptic equations have been in a focus of many studies. General schemes of such solvers for discretizations, based on square and triangular reference elements, have much in common and were analyzed in many papers, see the literature in [KFOF]. In this paper, we follow the approach of [IK,KJ2,K,KFOF], where considerable attention was paid to optimization of the basic components of *DD* algorithms. The interface component of *DD* solvers requires efficient preconditioners for the Schur complement matrix, resulting from the condensation of internal unknowns for each element. For designing such preconditioners, a number of techniques, similar for reference elements of the both geometries, but different for coordinate functions of different types (Lagrange, hierarchical etc.), were developed in [KJ2,KFOF]. Usually, the most time consuming and the most difficult for optimization component of *DD* algorithms are solvers for local Dirichlet problems on subdomains of decomposition, for which we use the abbreviation *LS*. Only recently, the first almost optimal *LS* was obtained in [K1]

[1]Research is supported by the grants from the Russian Foundation of Fundamental Research (project N 00-01-00772) and under the program "Universities of Russia" (project N 015.03.01.030).

[2]The first author has partially been supported by the Austrian Science Fund (FWF) through the special research programme (SFB) F013, project 16.

for one the most popular type of *hp* discretizations. Finite elements, in general curvilinear, of the discretization are images of the square reference element with the hierarchical coordinate functions, defined by the tensor products of the integrated Legendre polynomials. Some important properties of such discretizations were established in [IK]. There, for use in *LS*, a spectrally equivalent preconditioner in the form of finite-difference operator was suggested and analyzed, see also [KJ1,KJ2]. Although the use of the preconditioner provided a significant step in reducing the computational cost of *LS* and the *DD* algorithm in a whole, the next step, which is designing a fast[3] solution procedure for the preconditioner[4], still required an effort. This is for the reason of singularity of coefficients in the finite-difference-type preconditioner. As it was mentioned above, this step was completed in [K1], where an almost optimal solver was suggested, based on a combination of *DD* and FDFT (fast discrete Fourier transform) techniques. Later, multilevel solvers, based on the strengthened Cauchy inequality and other approaches, appeared in [Be] and in [K2].

The cited results allow us to obtain fast highly parallelizable *DD* solvers, which, to the best of the author's knowledge, until now have not been presented in the literature.

In this paper, we also suggest fast *LS*'s for discretizations with different from the pointed out above types of reference elements. Incomplete and transition elements are very important for adaptive computations. A remarkable fact, which we establish in the paper, is that almost optimal *LS*'s for the case of such elements may be not more complicated, than for complete elements. We present also fast algorithms for solving Dirichlet problems, arising on subdomains of decomposition in the *hp*-version with Lagrange finite elements. Since other components of *DD* algorithms have been thoroughly studied earlier, *e.g.,* see [KJ2,KFOF], and practically do not depend on the type of internal for elements coordinate functions, we are able to come to fast *DD* solvers for *hp* discretizations by Lagrange elements as well.

Let us note that we consider only the *DD* techniques, in which subdomains of decomposition are the domains, occupied by finite elements. The reason for this is the following. If subdomains with several elements were used, then, according

[3]By fast, we imply optimal and almost optimal, *i.e.,* optimal up to multiplier $(\log p)^k$, $p > 1$, with fixed k, solvers.

[4]Speaking about a solution procedure for some matrix, we mean the solution procedure for the system of equations with this matrix.

to our results, the most efficient solvers for such subdomains would be again *DD* solvers with smaller subdomains of decomposition. In this way we come to the algorithm, in which domains of finite elements taken for subdomains of decomposition. Note also that in our estimates of computational work, we do not count operations spent for multiplications of vectors by the finite element stiffness matrix, unavoidable in every iteration process.

The paper is arranged as follows. In Section 1, we describe the reference elements which generate the *hp* discretizations being considered in the paper. Some basic features of *DD* substructuring are discussed in Section 2. In Section 3, we formulate some results, concerning the component of *DD* solvers, usually termed as *prolongation*. The properties of prolongations as by means of iteration solvers for local Dirichlet problems, so by means of special prolongation operators in the polynomial spaces are discussed. Section 4 summarizes our earlier results on the Schur complement preconditioning. Section 5 presents fast *LS*'s for the discretizations with the hierearchical reference elements. We study the cases of complete and incomplete reference elements and *LS*'s based on the *DD* and multilevel approaches. Results of numerical experiments with some of these solvers are also described in this section. *LS* for the discretizations with the Lagrange (nodal) reference elements are described and analyzed in Section 6.

Let us describe some notations. Relations of the type $\mathbf{A} \le \mathbf{B}$ for symmetric positive matrices are understood in a sense of the scalar products, *i.e.,* as inequalities $\mathbf{v}^t\mathbf{A}\mathbf{v} \le \mathbf{v}^t\mathbf{B}\mathbf{v}$ with an arbitrary vector $\mathbf{v}$; $\mathbf{A}^+$ is the pseudo-inverse to the matrix $\mathbf{A}$; c with different indices denotes positive constants independent of p and the number of elements in a finite element assemblage; $\prec$ and $\asymp$ stand for the inequalities and equalities, held up to constants independent of p and the number of the finite elements; $|\cdot|_{k,\Omega}$, $||\cdot||_{k,\Omega}$ denote the seminorm and the norm in the Sobolev space $W_2^k(\Omega)$. We deal with the preconditioning of symmetric, positive or positive definite matrices. Therefore, the preconditioners possess the same properties, which as a rule we do not mention specially.

As a rule, we use thick or special fonts for matrices and vectors.

1 Two types of hp-discretizations

We consider Dirichlet problem for Poisson equation in 2-d domain. The problem is discretized into an assemblage of, in general, curvilinear finite elements

τ_r, $r = 1, 2, \ldots, \mathcal{R}$, specified by nondegenerate mappings $\mathbf{x} = \mathcal{X}^{(r)}(\mathbf{y}) : \bar{\tau}_0 \to \bar{\tau}_r$ with positive Jacobian's, where τ_0 is the reference square $\tau_0 = (-1, 1) \times (-1, 1)$. We assume that the conditions, called the *generalized conditions of the angular quasiuniformity*, are fulfilled for the mappings. If the mappings are bilinear, *i.e.,* elements have straight edges, these conditions are equivalent to shape regularity. In a more general case, they imply that the inequalities in the first line of (4.3) in [KJ1] are fulfilled with h specific for each element, but with $\alpha^{(1)}$ and θ independent of an element and h. The domain occupied by the assemblage is denoted Ω.

An *hp* discretization may be generated, *e.g.,* by one of the two types of reference elements, equipped with the space $\mathcal{Q}_{p,x}$ of all polynomials of the order not greater p, $p \geq 1$, in each variable. These reference elements are denoted as $\mathcal{E}_H$ and $\mathcal{E}_L$. For the bases in $\mathcal{Q}_{p,x}$, they have hierarchical and Lagrange interpolation polynomials, respectively. Let us introduce the set $\mathcal{M}_{1,p} = (\mathcal{L}_i(s)\,,\ i = 0, 1, \ldots, p)$ of polynomials of one variable

$$\mathcal{L}_0(s) = \tfrac{1}{2}(1+s)\,, \qquad \mathcal{L}_1(s) = \tfrac{1}{2}(1-s)\,,$$
$$\mathcal{L}_i(s) := \beta_i \int_{-1}^{s} P_{i-1}(t)\,dt = \gamma_i[P_i(s) - P_{i-2}(s)]\,, \quad i \geq 2\,,$$

where P_i are Legendre's polynomials and $\gamma_i = 0.5\sqrt{(2i-3)(2i+1)/(2i-1)}$. Thus, for $i \geq 2$, up to the multipliers β_i, $\mathcal{L}_i$ are the *integrated Legendre's polynomials.* These multipliers are chosen such that $\|\mathcal{L}_i(s)\|_{0,(-1,1)} = 1$, for $i \geq 2$. With $\omega := (\boldsymbol{\alpha} = (\alpha_1, \alpha_2) : 0 \leq \alpha_1, \alpha_2 \leq p)$, the hierarchical basis in $\mathcal{Q}_{p,x}$ is the set

$$\mathcal{M}_p = \left(L_{\boldsymbol{\alpha}}(\mathbf{x}) = \mathcal{L}_{\alpha_1}(x_1)\mathcal{L}_{\alpha_2}(x_2)\,,\ \boldsymbol{\alpha} \in \omega\right)\ .$$

For the description of the element $\mathcal{E}_L$, it is sufficient to specify the set of its nodes. One popular choice is the set of the so called GLL nodes, introduced by Gauss-Lobatto-Legendre quadrature. Another choice is the set of the GLC nodes, *i.e.,* having for coordinates the coordinates of Gauss-Lobatto-Chebyshev quadrature nodes plus the ends of the interval: $\mathbf{x} = \boldsymbol{\eta}^{(\alpha)} = (\eta_{\alpha_1}, \eta_{\alpha_2})$, $\eta_k = \cos(\pi(p-k)/p)$, $\boldsymbol{\alpha} \in \omega$.

In computations, incomplete, also called serendipity elements are widely used as well. The notations $\mathcal{E}_{\Im,H}, \mathcal{Q}_{\Im,p,x}, \mathcal{M}_{\Im,p}$ will stand for the incomplete reference element, the related polynomial space and the hierarchical basis in it, respectively. Let $\omega_o = (\boldsymbol{\alpha} : 2 \leq \alpha_1, \alpha_2 \leq p)$, $\omega_\Gamma = \omega \setminus \omega_o$, $\widehat{\omega}_o = (\boldsymbol{\alpha} : 2 \leq \alpha_1, \alpha_2,\ \alpha_1 + \alpha_2 \leq p)$, and $\omega_{\Im,o}$ is such that $\widehat{\omega}_o \in \omega_{\Im,o} \in \omega_o$. We define

$\mathcal{M}_{\Im,p} = (L_{\boldsymbol{\alpha}}(\mathbf{x}),\ \boldsymbol{\alpha} \in \omega_{\Im})$, where $\omega_{\Im} = \omega_{\Gamma} \cup \omega_{\Im,o}$, and $\mathcal{Q}_{\Im,p,x}$ as the space, spanned over polynomials from $\mathcal{M}_{\Im,p}$. *I.e.,* the set $\mathcal{M}_{\Im,p}$ of coordinate functions of the incomplete reference element is a subset of $\mathcal{M}_p$, which at least contains all coordinate functions $L_{\boldsymbol{\alpha}}$, which are not equal identically to zero on $\partial\tau_0$ and all internal coordinate functions of the total degree not greater p. In adaptive discretizations several incomplete as well as transition reference elements may be used and the basis $\mathcal{M}_{\Im,p}$ may be defined in a more complicated way.

If to use the notations $L_{\boldsymbol{\alpha}}(\mathbf{x})$ for coordinate functions of all these reference elements, then coordinate functions $p^{(r)}_{\boldsymbol{\alpha}}$ of any element τ_r are defined by the equalities $p^{(r)}_{\boldsymbol{\alpha}}(\mathcal{X}^{(r)}(\mathbf{y})) = L_{\boldsymbol{\alpha}}(\mathbf{y})$. We write the systems of finite element algebraic equations, resulting from the described discretizations in the form $\mathbf{K}u = \mathbf{f}$. For the reference element stiffness matrix we will use the notation $\mathbf{A}$.

2 Basic features of DD algorithms

Let $\mathcal{N}$ be the number of unknowns and $\mathcal{N}_I, \mathcal{N}_{II}, \mathcal{N}_{III}$ be the numbers of the *internal* for the elements, *edge* and *vertex* unknowns, respectively. The matrix $\mathbf{K}$ may be represented in either of the block forms

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}^{(I)} & \mathbf{K}^{(I,II)} \\ \mathbf{K}^{(II,I)} & \mathbf{K}^{(II)} \end{pmatrix} = \begin{pmatrix} \mathbf{K}_I & \mathbf{K}_{I,II} & \mathbf{K}_{I,III} \\ \mathbf{K}_{II,I} & \mathbf{K}_{II} & \mathbf{K}_{II,III} \\ \mathbf{K}_{III,I} & \mathbf{K}_{III,II} & \mathbf{K}_{III} \end{pmatrix},$$

where $\mathbf{K}^{(I)} = \mathbf{K}_I$ and $\mathbf{K}_L$ are $\mathcal{N}_L \times \mathcal{N}_L$ sub-matrices, related to the introduced groups of unknowns, $L = I, II, III$. Note that everywhere in the paper, we use similar indexation of different matrices with one distinction: for blocks of the matrices, related to reference elements, we use arabic numerals instead of roman.

Often, at construction of *DD* algorithms, especially for hierarchical *hp* discretizations, a formal splitting of the vertex unknowns from the rest is made. We outline such algorithms first.

Lemma 1. *Let $\mathbf{K}_v$ be obtained from $\mathbf{K}$ by putting zero blocks instead of $\mathbf{K}_{L,III}$, $L = I, II$, and their transposed. Then $c_v(1 + \log p)^{-1}\mathbf{K}_v \le \mathbf{K} \le 2\mathbf{K}_v$ with $c_v \ge 0$ depending only on the constants from the generalized angular quasi-uniformity conditions.*

The left inequality of lemma was proved in [IK], see also [KJ2]. In a slighter form and with additional assumption of orthogonality of some subsets of coordinate functions it was proved in [B]. Let $\mathbf{K}_\wedge$ be the block of $\mathbf{K}$ generated by the internal and edge coordinate functions. Further steps are related to the factorization

$$\mathbf{K}_\wedge = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{K}_{II,I}\mathbf{K}_I^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{K}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{K}_I^{-1}\mathbf{K}_{I,II} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \mathbf{S} = \mathbf{K}_{II} - \mathbf{K}_{II,I}\mathbf{K}_I^{-1}\mathbf{K}_{I,II}, \tag{2.1}$$

with $\mathbf{I}$ denoting unity matrices. In general, taking into account (2.1), we may define a *DD* preconditioner $\boldsymbol{\mathcal{K}}_\wedge$ for $\mathbf{K}_\wedge$ through its inverse by setting $\boldsymbol{\mathcal{K}}_\wedge^{-1} = \boldsymbol{\mathcal{K}}_I^{+} + \mathbb{P}\boldsymbol{\mathcal{S}}^{-1}\mathbb{P}^t$. Here, $\boldsymbol{\mathcal{K}}_I$ and $\boldsymbol{\mathcal{S}}$ are preconditioners for $\mathbf{K}_I$ and $\mathbf{S}$, and the $(\mathcal{N}_I + \mathcal{N}_{II}) \times \mathcal{N}_{II}$ matrix $\mathbb{P}$ is called the *prolongation matrix.* It will be discussed in Section 3.

Evidently, $\mathbf{K}_I = \mathrm{d}iag\,[\mathbf{K}_{I,1}, \mathbf{K}_{I,2}, .., \mathbf{K}_{I,\boldsymbol{\mathcal{R}}}]$, where $\mathbf{K}_{I,r}$ are the internal stiffness matrices of elements, and $\boldsymbol{\mathcal{K}}_I$ has the same block-diagonal structure. More over, under the conditions of generalized angular quasyuniformity, we may take all the blocks in $\boldsymbol{\mathcal{K}}_I$ the same, *i.e.,* set $\boldsymbol{\mathcal{K}}_I = \mathrm{d}iag\,[\boldsymbol{\mathcal{K}}_{I,o}, \boldsymbol{\mathcal{K}}_{I,o}, .., \boldsymbol{\mathcal{K}}_{I,o}]$, where $\boldsymbol{\mathcal{K}}_{I,o}$ is any good preconditioner for the reference element internal stiffness matrix. Suppose, there are m different edges of elements inside Ω. Although the structure of $\mathbf{S}$ is not block-diagonal, with some lost in condition, see Section 4, we may accept $\boldsymbol{\mathcal{S}} = \mathrm{d}iag\,[\boldsymbol{\mathcal{S}}_o, \boldsymbol{\mathcal{S}}_o, .., \boldsymbol{\mathcal{S}}_o]$, where the same matrix $\boldsymbol{\mathcal{S}}_o$ stands for each of m edges. The prolongation $\mathbb{P}\mathbf{v}_{II}$ may be done element-wise and does not require assembling. More over, if $\mathbb{P}_\varepsilon$ is an appropriate prolongation matrix for the reference element, than it may be accepted for the restriction of $\mathbb{P}$ to any element τ_r.

We see, that the *DD* preconditioners under consideration require specification of only 3 matrices $\boldsymbol{\mathcal{K}}_{I,o}$, $\boldsymbol{\mathcal{S}}_o$ and $\mathbb{P}_\varepsilon$ for the reference element and provide a high level of parallelization, *i.e.,* element-wise and edge-wise.

Efficient *DD* preconditioners may be also associated with the similar to (2.1) factorization

$$\begin{gathered} \mathbf{K} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{K}^{(II,I)}\mathbf{K}_I^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{K}_I & \mathbf{0} \\ \mathbf{0} & \mathbb{S} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{K}_I^{-1}\mathbf{K}^{(I,II)} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \\ \mathbb{S} = \mathbf{K}^{(II)} - \mathbf{K}^{(II,I)}\mathbf{K}_I^{-1}\mathbf{K}^{(I,II)}, \end{gathered} \tag{2.2}$$

where now Schur complement $\mathbb{S}$ has the dimension $(\mathcal{N}_{II} + \mathcal{N}_{III}) \times (\mathcal{N}_{II} + \mathcal{N}_{III})$.

Again we may define the *DD* preconditioner by the similar to the used above formula $\boldsymbol{\mathcal{K}}^{-1} = \boldsymbol{\mathcal{K}}_I^{+} + \mathbf{P}\boldsymbol{\Psi}^{-1}\mathbf{P}^t$ with the same as above preconditioner $\boldsymbol{\mathcal{K}}_I$. The restriction of $\mathbf{P}$ to any element is again specified by an appropriate prolongation operator $\mathbf{P}_\varepsilon$ for the reference element, which, for the same reference element, is defined in quite similar to $\mathbb{P}_\varepsilon$ ways. Indeed, $\mathbb{P}_\varepsilon$ is the restriction of $\mathbf{P}_e$ to a more narrow space of definition. The use of factorization (2.2) is more suitable for Lagrange reference elements and assumes different approach for obtaining the preconditioner $\boldsymbol{\Psi}$ for $\mathbb{S}$, see [KJ2,KFOF].

3 Prolongations

Suppose that $\mathbb{K}_\wedge$ is an explicitly given preconditioner for $\mathbf{K}_\wedge$ with the blocks $\mathbb{K}_I$ and $\mathbb{K}_{II,I}$,– $\mathbb{K}_I$ not necessarily coinciding with $\boldsymbol{\mathcal{K}}_I$,– and that there is a fast exact solver for $\mathbb{K}_I$. Evidently, a good choice is to set $\mathbb{P}^t = (-\mathbb{K}_{II,I}\mathbb{K}_I^{-1}, \mathbf{I})$. When we do not have such $\mathbb{K}_\wedge$, we are forced to use other options.

1) Special prolongation operators $\mathbb{P}$ may be effectively implemented, at least in the case of complete reference elements. In *hp*-version, such algorithms were introduced in [IK],– for more details see [KFOF],– following the approach for *h*-version, presented, *e.g.,* in [N]. By the generalized angular quasyuniformity conditions, it is sufficient to define the prolongation operator $\mathcal{P}_\varepsilon : \mathcal{Q}_{p,x}(\partial\tau_0) \to \mathcal{Q}_{p,x}$ for the reference element, satisfying the inequalities

$$|v|_{1,\tau_0} \le c|\psi|_{1/2,\partial\tau_0}\,, \quad \|v\|_{1,\tau_0} \le c\|\psi\|_{1/2,\partial\tau_0}\,, \quad \forall\psi \in \mathcal{Q}_{p,x}(\partial\tau_0) := \mathcal{Q}_{p,x}|_{\partial\tau_0}\,. \quad (3.3)$$

Let $\mathbb{P}_\varepsilon$ be the restriction of the matrix representation of the operator $\mathcal{P}_\varepsilon$ to the edge nodes. It is clear, that the matrix $\mathbb{P}$ is uniquely defined by its restrictions $\mathbb{P}_r$ to the edge nodes of each finite element τ_r. As it was shown in the cited papers, for all of them we can accept $\mathbb{P}_r = \mathbb{P}_\varepsilon$ at appropriate local orderings of degrees of freedom. Suitable for the use in *DD* algorithms prolongation operators were suggested in [BDM1,BDM2]. Their use provides considerable reduction of operations in comparison with the condensation procedures, based on Gaussian elimination, see [KFOF]. It may be also expected, that the cost of their computer versions could be reduced, at a more elaborate arrangement of computations. It is important, that if $\boldsymbol{\mathcal{K}}_I \asymp \mathbf{K}_I$, and $\boldsymbol{\mathcal{S}} \asymp \mathbf{S}$ and the prolongation matrix $\mathbb{P}$ is defined in the described way, then $\boldsymbol{\mathcal{K}}_\wedge$ and $\mathbf{K}_\wedge$ are spectrally equivalent. In the same way we define the prolongation matrix $\mathbf{P}$ by setting $\mathbf{P}_r \equiv \mathbf{P}_\varepsilon$, where $\mathbf{P}_\varepsilon$is the matrix representation of $\mathcal{P}_\varepsilon$. This results in the spectral equivalence $\boldsymbol{\mathcal{K}} \asymp \mathbf{K}$ under conditions that $\boldsymbol{\mathcal{K}}_I \asymp \mathbf{K}_I$ and $\mathbb{S} \asymp \boldsymbol{\Psi}$.

2) Let $\mathbf{A} = \{\mathbf{A}_{i,j}\}_{i,j=1}^3$ be the reference element stiffness matrix and $\boldsymbol{\mathcal{A}}$ its preconditioner $\boldsymbol{\mathcal{A}} \asymp \mathbf{A}$. For the blocks on diagonal, we always use one index instead of two, *e.g.,* the notation $\boldsymbol{\mathcal{A}}_1$ instead of $\boldsymbol{\mathcal{A}}_{1,1}$. It may be accepted $\mathbb{P}_\varepsilon^t = (-\mathbf{A}_{II,I}\widetilde{\mathbf{A}}_I^{-1}, \mathbf{I})$, with

$$\widetilde{\mathbf{A}}_1^{-1} = [\mathbf{I} - \textstyle\prod_{k=1}^n (\mathbf{I} - \sigma_k \boldsymbol{\mathcal{A}}_1^{-1}\mathbf{A}_1)]\mathbf{A}_I^{-1} \ ,$$

some $n \geq 1$ and Chebyshev iteration parameters σ_k. Note that, if $\mathbf{u}^{(n)} := \widetilde{\mathbf{K}}_I^{-1}\mathbf{f}$, then $\mathbf{u}^{(n)}$ is produced by n iterations

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \sigma_k \boldsymbol{\mathcal{A}}_1^{-1}(\mathbf{A}_1\mathbf{u}^k - \mathbf{f}) \ , \qquad \mathbf{u}^0 = \mathbf{0} \ .$$

In the case of the reference element $\mathcal{E}_H$, the following assertion is true.

If $\boldsymbol{\mathcal{K}}_I \asymp \mathbf{K}_I$, $\boldsymbol{\mathcal{A}}_1 \asymp \mathbf{A}_1$, $\boldsymbol{\mathcal{S}} \asymp \mathbf{S}$ and $n = \mathcal{O}(\log p)$, then $\boldsymbol{\mathcal{K}}_\wedge \asymp \mathbf{K}_\wedge$. If $\mathbb{S} \asymp \boldsymbol{\Psi}$, then under same conditions on $\boldsymbol{\mathcal{K}}_I$, $\boldsymbol{\mathcal{A}}_1$ and n, we have $\boldsymbol{\mathcal{K}} \asymp \mathbf{K}$.

Although these facts are important, we omit the proof due to the restricted volume of the paper.

For the discretizations with the reference element $\mathcal{E}_L$, the prolongation matrices may be defined by the same iteration process, however, the necessary number of iterations require special analysis.

4 Schur complement preconditioning

The Schur complement preconditioning is the most thoroughly studied component of *DD* algorithms for *hp* discretizations of 2-d elliptic equations by complete reference elements. As for other components, it requires to develop a preconditioner only for the reference element. Well recognized facts serve the basement for the Schur complement preconditioning. If, *e.g.,* the matrix $\boldsymbol{\mathcal{S}}_A$ is spectrally equivalent to $\mathbf{S}_A := \mathbf{A}^{(2)} - \mathbf{A}^{(2,1)}\mathbf{A}_1^{-1}\mathbf{A}^{(1,2)}$, then, by the trace and prolongation theorems in complete polynomial spaces (see for these theorems [BDM1,BDM2]) it induces some discrete semi-norm, equivalent to $|\cdot|_{1/2,\partial\tau_o}$ on the space $\mathcal{Q}_{p,x}|_{\partial\tau_o}$. Since $|\cdot|_{1/2,\partial\tau_o}$ practically does not depend on the form of τ_o, suggestions made for the square in [IK,KJ1-KJ4] and for triangular reference elements in [KFOF] are similar. More distinctions are related to types of traces on $\partial\tau_o$ of coordinate functions, *e.g.,* hierarchical of different types, Lagrange etc.

In the case of the reference element $\mathcal{E}_H$, the matrix $\boldsymbol{\mathcal{S}}_o$ may be defined in the following way. Let $\boldsymbol{\mathcal{S}}_+ := \mathbf{W}^t\boldsymbol{\Lambda}\mathbf{W}$, where $\mathbf{W}$ is the explicitly given triangular matrix of the transformation from $\mathcal{M}_{1,p}$ to the trigonometric basis $\mathcal{M}_{1,p,tr} = (\mathrm{arc}\, k \cos s,\ k = 0, 1, .., p)$ and $\mathbf{\Lambda} = \mathrm{diag}[0, 1, .., p]$ is a diagonal matrix. If $\boldsymbol{\mathcal{S}}_o$ is obtained from $\boldsymbol{\mathcal{S}}_+$ by deleting the two first rows and columns corresponding to $\mathcal{L}_k,\ k = 0, 1$, then $\boldsymbol{\mathcal{S}} \prec \mathbf{S} \prec (1 + log\, p)^3 \boldsymbol{\mathcal{S}}$, see Section 3 in [KJ2]. Solving the system with the matrix $\boldsymbol{\mathcal{S}}_o$ requires $\mathcal{O}(p^2)$ arithmetic operations. Though not the cheapest, this is the simplest way of the Schur complement preconditioning, introduced in [IK], which allow to obtain an almost optimal *DD* algorithm.

Now, we turn to the reference element $\mathcal{E}_L$ and enumerate clockwise the edge nodes $\boldsymbol{\eta}^{(\boldsymbol{\alpha})} \in \partial\tau_{_o}$. Let $\mathbb{D}_\varepsilon$ denotes the cyclic $4p \times 4p$ matrix having 2 on diagonal and (-1) in the two adjacent diagonals and in the lower left and upper right corners. Let also $\boldsymbol{\mathcal{D}}_\varepsilon$ be such that $\boldsymbol{\mathcal{D}}_\varepsilon^2 = \mathbb{D}_\varepsilon$ and $\ker \boldsymbol{\mathcal{D}}_\varepsilon = \ker \mathbb{D}_\varepsilon$. The matrix $\boldsymbol{\mathcal{D}}$ is defined as the result of assembling $\mathcal{R}$ matrices $\boldsymbol{\mathcal{D}}_\varepsilon$, considered as stiffness matrices of elements τ_r, and deleting rows and columns for fixed nodes. Solving the system with the matrix $\boldsymbol{\mathcal{D}}$ by the preconditioned simple iteration process with a fixed number n of iterations defines another matrix, which is denoted by $\boldsymbol{\mathcal{S}}_\diamond^{-1}$. Formally, this iteration process is equivalent to setting $\boldsymbol{\mathcal{S}}_\diamond^{-1} = [\mathbf{I} - \prod_{k=1}^{n_s}(\mathbf{I} - \varkappa_k \mathbf{E}^{-1}\boldsymbol{\mathcal{D}})]\boldsymbol{\mathcal{D}}^{-1}$with Chebyshev iteration parameters $\varkappa_k$, the diagonal matrix $\mathbf{E}$, which diagonal coincides with $\boldsymbol{\mathcal{D}}$, and $n_s = \mathcal{O}(\sqrt{p})$. As it was established in [KJ1,KJ2], $\boldsymbol{\mathcal{S}}_\diamond \prec \boldsymbol{\mathcal{D}} \prec \mathbb{S} \prec \boldsymbol{\mathcal{D}} \prec \boldsymbol{\mathcal{S}}_\diamond$ and the cost of solving the system with the matrix $\boldsymbol{\mathcal{D}}$ is $\mathcal{O}(\mathcal{R}p\sqrt{p}\log p)$, see Section 6 in [KJ4]. The important fact for this estimate is that multiplications of vectors by $\boldsymbol{\mathcal{D}}$ may be completed in parallel element-wise by the 1-d FDFT's.

Since in the case under consideration we have established that Schur complement preconditioning is made by similar with h-version preconditioners, the further reduction of the computational cost of the interface problem may be also achieved in similar ways: *e.g.,* by applying the *DD* technique for solving the system with the matrix $\boldsymbol{\mathcal{D}}$, by implementation of BPX preconditioners etc.

5 Fast solvers for Dirichlet problems on subdomains of decomposition; hierarchical p-version

5.1. DD-like solver.

`5.1.1. Definition of the preconditioner.` Without lost of generality we assume $p = 2N + 1$. The subset $\mathcal{M}_p^0 := \big(L_{\boldsymbol{\alpha}} = \mathcal{L}_{\alpha_1}\mathcal{L}_{\alpha_2}\ ,:\ 2 \le \alpha_1, \alpha_2 \le p\big)$ of internal coordinate functions of the reference element may be rearranged in four groups, each having N^2 functions. The 1-st part contains $L_{\boldsymbol{\alpha}}$ with even α_1 and α_2, 2-nd – with even α_1 and odd α_2, 3-d – with odd and even, and 4-th – with odd and odd. This rearranging splits the internal stiffness matrix $\mathbf{A}_1$ of the reference element $\mathcal{E}_H$ in four independent blocks, so that $\mathbf{A}_1 = diag\,[\mathbf{A}_{e,e}, \mathbf{A}_{e,o}, \mathbf{A}_{o,e}, \mathbf{A}_{o,o}]$. Since all the matrices $\mathbf{A}_{a,b}$, $a, b = e, o$, are spectrally equivalent uniformly in p, it is sufficient to derive the fast solver for one of them, say, $\mathbf{A}_{e,e}$.

Let $\hbar = 1/(N+1)$, $\eta_i = i\hbar$, $\varphi_i = \eta_i^2$, $\mathbf{\Lambda}_{e,e}$ be the matrix of the system of equations[5]

$$-2\left(\varphi_i \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j-1}}{\hbar^2} + \varphi_j \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\hbar^2}\right) + (\frac{\varphi_i}{\varphi_j} + \frac{\varphi_j}{\varphi_i})u_{i,j} = f_{i,j}\,, \quad 1 \le i, j \le N\,,$$

with $u_{i,j} = 0$, if one of the indices is out of the pointed out range. Formally, this is a finite-difference approximation to the deteriorating elliptic equation in the unit square $\pi_1 = (\xi = (\xi_1, \xi_2) : 0 < \xi_1, \xi_2 < 1)$:

$$Łu \equiv -2\left(\xi_1^2 \partial^2 u/\partial\xi_2^2 + \xi_2^2 \partial^2 u/\partial \xi_1^2\right) + (\xi_1^2/\xi_2^2 + \xi_2^2/\xi_1^2)u = f(\boldsymbol{\xi})\,, \quad u|_{\partial\pi_1} = 0\,. \tag{5.4}$$

As it was established in [IK,KJ1], $\mathbf{A}_{e,e} \asymp \mathbf{\Lambda}_{e,e}$. By this reason, in order to get a fast *LS*, it is sufficient to obtain a fast solver for the matrix $\mathbf{\Lambda}_{e,e}$. The latter will be designed as *DD* solver for the discretization of (5.4) on the uniform grid. In order to distinguish this secondary decomposition and its main component from the global *DD*, we will use for them the abbreviations DD_l and LS_l. The first step will be approximation of (5.4) by the elliptic problem with the piece-wise constant coefficients.

The square mesh of the size $\hbar$ will be termed the *fine mesh*. We introduce the rectangular nonuniform mesh, termed the *coarse* or *decomposition mesh*, which decomposes π_1 in rectangular subdomains $\delta_{k,l}$, such that the coefficients of the equation (5.4) in either of them differ not too much from appropriate constants. The mesh lines of the coarse mesh are denoted $\overline{\xi}_{k,l} = \zeta_l$, where $k = 1, 2$, $\quad i = 0, 1, \ldots, N+1$, $\quad l = 0, 1, \ldots, l_0$, and numbers l_0 and ζ_l are specified below. Two parameters $q > 1$ and $n_0 \ge 1$ are used to generate the

[5] Do not confuse these η_i with the coordinates of the nodes of Gauss type quadratures, for which similar notations were introduced in Section 1.

coarse mesh as follows: $\zeta_0 = 0$, $\zeta_{l_0-1} = 1$,

$$\begin{aligned}
&\zeta_l = \eta_i \ , \qquad \text{for} \qquad i := \gamma(l) := \text{int}\ (q^l - 1)n_0 \ , \qquad l = 1, 2, \ldots, l_0 - 2 \ , \\
&\zeta_{l_0-1} = \eta_{\gamma(l_0-1)}, \quad \text{if} \quad \eta_{\gamma(l_0-1)} = 1, \\
&\zeta_{l_0-1} = \eta_i, \quad i = \text{int}\ \tfrac{1}{2}(\gamma(l_0-2) + N + 1), \quad \text{if} \quad \eta_{\gamma(l_0-1)} > 1 \, .
\end{aligned}$$

If $\text{int}\lfloor a \rfloor$ is the closest to a positive integer not less than a, then we conclude that $l_0 = \text{int}\ \lfloor (\log(\frac{N}{n_0} + 1))/\log q \rfloor$.

There exist the piece-wise constant functions $\psi(\zeta)$ and $b(\boldsymbol{\xi})$ of one and two variables, which are constants on intervals $\zeta \in (\zeta_{k-1}, \zeta_k)$, and subdomains $\boldsymbol{\xi} \in \delta_{k,l}$, respectively, and such that $c_1\psi(\zeta) \le \zeta^2 \le c_2\psi(\zeta)$ and $c_1 b(\boldsymbol{\xi}) \le (\xi_1/\xi_2)^2 + (\xi_2/\xi_1)^2 \le c_2 b(\boldsymbol{\xi})$ with positive constants independent of p. The functions $\psi(\zeta)$ and $b(\boldsymbol{\xi})$, are easily defined, *e.g.*, as the arithmetic means of the maximal and minimal values of ζ^2 and $(\xi_1/\xi_2)^2 + (\xi_2/\xi_1)^2$ on the corresponding sets, see [K1,K2].

It is convenient to define a new preconditioner as a finite element matrix. We subdivide each nest of the fine mesh by one of the diagonals in two triangles and introduce the space $\mathcal{H}_o(\pi_1)$ of piece-wise linear functions continuous on $\overline{\pi}_1$ subordinate to the boundary condition (5.4). We define $\mathbf{B} = \mathbf{B}_{e,e}$ as the finite element matrix, induced on this subspace by the bilinear form $\hbar^{-2} a_{\pi_1}(v, w)$,

$$a_{\pi_1}(v, w) = 2 \int_{\pi_1} \left(\psi(\xi_1) \frac{\partial v}{\partial \xi_2} \frac{\partial w}{\partial \xi_2} + \psi(\xi_2) \frac{\partial v}{\partial \xi_1} \frac{\partial w}{\partial \xi_1} \right) d\xi + \hbar^2 \sum_{i,j=1}^{N} \overline{b}_{i,j} v(\eta_i, \eta_j) w(\eta_i, \eta_j) \, ,$$

where $\overline{b}_{i,j} = \frac{1}{4}[b(\eta_i - 0, \eta_j - 0) + b(\eta_i + 0, \eta_j - 0) + b(\eta_i - 0, \eta_j + 0) + b(\eta_i + 0, \eta_j + 0)]$.

Lemma 2. *For $c_k > 0$ depending only on n_0 and q, we have $c_1\mathbf{B} \le \mathbf{\Lambda}_{e,e} \le c_2\mathbf{B}$.*

For the proof, see [K1]. The preconditioner $\mathbf{B}_{e,e}$ is optimal in condition and admits fast LS_l's on subdomains $\delta_{k,l}$, such as FDFT and multigrid methods.

Similarly with the primal decomposition, we represent the space $\mathcal{V} := \mathcal{R}^{N^2}$ of vectors $\mathbf{v} = (v_{i,j}\, ,\ 1 \le i, j \le N)$, by the direct sum of the subspaces $\mathcal{V}_m$, $m = 1, 2, 3$, of subvectors with the entries related to the internal for subdomains $\delta_{k,l}$, edge and vertex unknowns, respectively. For the restrictions of these spaces to $\overline{\delta}_{k,l}$, we use notations $\mathcal{U}$, $\mathcal{U}_m$. The above splitting of $\mathcal{V}$, assumes the block forms of $\mathbf{B}$, such as

$$\mathbf{B} = \{\mathbf{B}^{(i,j)}\}_{i,j=1}^2 = \{\mathbf{B}_{i,j}\}_{i,j}^3 \, , \qquad \mathbf{B}^{(1)} = \mathbf{B}_1 \, , \qquad \mathbf{B}_{3,1} = \mathbf{B}_{1,3} = \mathbf{0} \, .$$

Let $\mathbf{S}_{\mathrm{nod}} = \mathbf{B}_2 - \mathbf{B}_{2,1}\mathbf{B}_1^{-1}\mathbf{B}_{1,2}$, and $\mathbf{B}_\triangle = \mathbf{B}_3 - \mathbf{B}_{3,2}\mathbf{S}_{\mathrm{nod}}^{-1}\mathbf{B}_{2,3}$. Let also $\boldsymbol{\mathcal{S}}_{\mathrm{nod}}$ be such that $\mathrm{cond}[\boldsymbol{\mathcal{S}}_{\mathrm{nod}}^{-1}\mathbf{S}_{\mathrm{nod}}] \prec (1+\log N)^\nu$ and there exist an optimal solver for the systems with this matrix. For solving systems with matrices $\mathbf{K}_{I,r}$ by PCG (method of preconditioned conjugate gradients), we may use the *DD*$_l$-preconditioner $\boldsymbol{\mathcal{K}}_{I,o} := \mathrm{diag}\,[\boldsymbol{\mathcal{K}}_o, \boldsymbol{\mathcal{K}}_o, \boldsymbol{\mathcal{K}}_o, \boldsymbol{\mathcal{K}}_o]$ with the matrix $\boldsymbol{\mathcal{K}}_o$, which factorized form is $\boldsymbol{\mathcal{K}}_o = \mathbb{P}_o^T \boldsymbol{\mathcal{D}} \mathbb{P}_o$, where $\boldsymbol{\mathcal{D}} = \mathrm{diag}[\mathbf{B}_1, \mathbf{S}_{\mathrm{nod}}, \mathbf{B}_\triangle]$, and

$$\mathbb{P}_o = \begin{pmatrix} \mathbf{I}_1 & \mathbf{B}_1^{-1}\mathbf{B}_{1,2} & \\ & \mathbf{I}_2 & \mathbf{S}_{\mathrm{nod}}^{-1}\mathbf{B}_{2,3} \\ & & \mathbf{I}_3 \end{pmatrix} ,$$

assuming that systems with the matrix $\mathbf{S}_{\mathrm{nod}}$ are solved by $\mathcal{O}(\log p)$ the simple preconditioned iterations with the preconditioner $\boldsymbol{\mathcal{S}}_{\mathrm{nod}}$ and Chebyshev iteration parameters.

It is evident that $\mathbf{B}_1 = \mathrm{diag}\,[\mathbf{B}_1^{(k,l)}]_{k,l=1}^N$, where $\mathbf{B}_1^{(k,l)}$ is the block of the matrix $\mathbf{B}^{(k,l)}$ which is induced by the bilinear form $\hbar^{-2} a_{\delta_{k,l}}(\ ,\ )$,

$$a_{\delta_{k,l}}(v,w) = 2\int_{\delta_{k,l}} \left(\psi_k \frac{\partial v}{\partial \xi_2}\frac{\partial w}{\partial \xi_2} + \psi_l \frac{\partial v}{\partial \xi_1}\frac{\partial w}{\partial \xi_1}\right) d\xi + \hbar^2 \sum_{i,j=1}^N b_{k,l} v(\eta_i,\eta_j) w(\eta_i,\eta_j) ,$$

on the space $\mathcal{H}(\delta_{k,l})$. Here, ψ_k, $b_{k,l}$ are the values of functions ψ, b on the sets (ζ_{k-1}, ζ_k), $\delta_{k,l}$, respectively, and the bilinear form and the space are the restrictions of $a_{\pi_1}(,)$ and $\mathcal{H}_o(\pi_1)$ to a subdomain $\delta_{k,l}$. Each subsystem with the matrix $\mathbf{B}_1^{(k,l)}$ may be solved by FDFT, and the total cost for solving all subsystems is $\mathcal{O}(N^2 \log N)$.

Theorem 1. *Suppose that the cost of the multiplication of a vector by* $\mathbf{S}_{\mathrm{nod}}$ *is* $\mathcal{O}(N^2)$ *and the above assumptions hold. Then, except for multiplications by* $\mathbf{K}_{I,r}$*, the total cost of solving the system* $\mathbf{K}_{I,r}\mathbf{u}_{I,r} = \mathbf{f}_{I,r}$ *by the PCG with the preconditioner* $\boldsymbol{\mathcal{K}}_{I,o}$ *is* $\mathcal{O}(N^2(\log N)^{0.5\nu+3})$.

The additional implementation details and, in particular, the computation of $\boldsymbol{\mathcal{S}}_{\mathrm{nod}}$ are briefly discussed in Subsection 5.1.2.

The main term in the cost is due to the multiplications by $\mathbf{S}_{\mathrm{nod}}$ at calculation of $\mathbf{B}_\triangle$. By introducing an additional preconditioner for the interface component with cheaper multiplications, we may reduce the computational cost. Suppose, that there exists the matrix $\mathbf{C}^{(2)}$, which is spectrally equivalent to the matrix $\boldsymbol{\mathcal{B}}^{(2)}$ with the blocks $\boldsymbol{\mathcal{B}}_2 = \mathbf{S}_{\mathrm{nod}}$ and $\boldsymbol{\mathcal{B}}_{i,j} = \mathbf{B}_{i,j}$ for $(i,j) = (2,3), (3,2), (3,3)$, and such that the cost of the multiplication $\mathbf{C}^{(2)}\mathbf{v}^{(2)}$ is $\mathcal{O}(N(\log N)^2)$. Then we may set $\boldsymbol{\mathcal{K}}_o = \mathbb{P}_c^T \boldsymbol{\mathcal{D}}_c \mathbb{P}_c$, where $\boldsymbol{\mathcal{D}}_c = \text{diag}[\mathbf{B}_1, \mathbf{C}_2, \mathbf{C}_\triangle]$, and

$$\mathbb{P}_c = \begin{pmatrix} \mathbf{I}_1 & \mathbf{B}_1^{-1}\mathbf{B}_{1,2} & \\ & \mathbf{I}_2 & \mathbf{C}_2^{-1}\mathbf{C}_{2,3} \\ & & \mathbf{I}_3 \end{pmatrix}, \qquad \mathbf{C}_\triangle = \mathbf{C}_3 - \mathbf{C}_{3,2}\mathbf{C}_2^{-1}\mathbf{C}_{2,3}\,.$$

We assume, that the systems with $\mathbf{C}_2$ are solved by $\mathcal{O}(\log p)$ simple preconditioned iterations with the preconditioner $\boldsymbol{\mathcal{S}}_{\text{nod}}$ and variable Chebyshev iteration parameters. It was established in [K2], that the matrix $\mathbf{C}^{(2)}$, – given explicitly and satisfying the formulated conditions, – existed, the matrix $\boldsymbol{\mathcal{K}}_o$ was spectrally equivalent to $\mathbf{B}$ and the system with the matrix $\boldsymbol{\mathcal{K}}_o$ might be solved for $\mathcal{O}(N^2 \log N)$ arithmetic operations. We refer to this work also for definition of $\mathbf{C}^{(2)}$. In the proof, we essentially used the results of Maz'ya and Poborchi [MP] on the boundary norms for harmonic H^1-functions on the rectangular domains with the large aspect ratio.

`5.1.2. Additional implementation details.` The matrices $\mathbf{B}_\triangle$ and $\mathbf{C}_\triangle$ have the dimension $l_0^2 \times l_0^2$, and the exact elimination procedures are cheap. Therefore, in this subsection we touch only the solvers for the preconditioners $\mathbf{B}_1$ and $\boldsymbol{\mathcal{S}}_{\text{nod}}$.

As it was pointed out, the systems with the matrices $\mathbf{B}_1$ may be solved by 2-d FDFT's, applied to independent subsystems with matrices $\mathbf{B}_1^{(k,l)}$ and almost optimal in the operation count. It is not straight forward to apply iterative solvers to these systems. We see, that at one of the indices k, l, say k, approaching zero, the ratio $\psi(\xi_2)/\psi(\xi_1)$, $\xi \in \delta_{k,l}$, and the aspect ratio of the rectangle $\delta_{k,l}$ tend to infinity, making Dirichlet problems on subdomains $\delta_{k,l}$ highly orthotropic. However, some suggestions on anisotropic multigrid methods, made in [S,Pf,GO] and [AV], may be well adapted to LS_l-s. Thus, indeed, we may construct a variety of asymptotically optimal multilevel solvers for the matrices $\mathbf{B}_1^{(k,l)}$. One of them, formulated in [K2], is based on the spectrally equivalent to these matrices BPX-like preconditioners, which are denoted by $\mathbb{B}_1^{(k,l)}$. They result from the multiscale decomposition of the finite element spaces $\mathcal{H}(\delta_{k,l})$ on the special sequence of the imbedded meshes, see Section 7 in [K2].

Let us turn to preconditioning of the Schur complement. To simplify the presentation, here we use the notation $\boldsymbol{\Upsilon}$ for $\mathbf{S}_{\text{nod}}$. The Schur complement $\boldsymbol{\Upsilon} = \mathbf{B}_2 - \mathbf{B}_{2,1}\mathbf{B}_1^{-1}\mathbf{B}_{1,2}$ is the result of assembling of the local Schur complements

$\boldsymbol{\Upsilon}^{(k,l)} = \mathbf{B}_2^{(k,l)} - \mathbf{B}_{2,1}^{(k,l)}(\mathbf{B}_1^{(k,l)})^{-1}\mathbf{B}_{1,2}^{(k,l)}$. The latter has the block form $\boldsymbol{\Upsilon}^{(k,l)} = \{\boldsymbol{\Upsilon}_{i,j}^{(k,l)}\}_{i,j=1}^4$, with each block $\boldsymbol{\Upsilon}_{i,i}^{(k,l)}$ on diagonal corresponding to one of the edges $T_{s,t}$, which are ordered, *e.g.,* clockwise. The preconditioner $\boldsymbol{\mathcal{S}}_{\rm nod}$ is assembled of local matrices $\boldsymbol{\mathcal{S}}_{\rm nod}^{(k,l)}$, which may be defined in one of the two ways: in $\boldsymbol{\Upsilon}^{(k,l)}$, we place zero blocks instead of

1) the off-diagonal blocks $\boldsymbol{\Upsilon}_{i,j}^{(k,l)}$ with $i + j = \text{odd}$, or

2) all off-diagonal blocks.

By assembling so defined $\boldsymbol{\mathcal{S}}_{\rm nod}^{(k,l)}$, we come to $\boldsymbol{\mathcal{S}}_{\rm nod}$. It was established in [K2], that at the both choices we have

$$\operatorname{cond}[\boldsymbol{\mathcal{S}}_{\rm nod}^{-1}\boldsymbol{\Upsilon}], \operatorname{cond}[\boldsymbol{\mathcal{S}}_{\rm nod}^{-1}\mathbf{C}_2] \prec (1 + \log p)^2 ,$$

see (7.5) and Lemma 7.3 in [K2].

In our algorithm, we do not need to calculate $\boldsymbol{\mathcal{S}}_{\rm nod}$ and $\boldsymbol{\Upsilon}$. Instead we use the matrices $\boldsymbol{\mathcal{S}}_{\rm tr}^{(k,l)} = \boldsymbol{\mathcal{P}}^t\boldsymbol{\mathcal{S}}_{\rm nod}\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\Upsilon}_{\rm tr} = \boldsymbol{\mathcal{P}}^t\boldsymbol{\Upsilon}\boldsymbol{\mathcal{P}}$, which are the result of the transformation to the trigonometric basis in the space $\boldsymbol{\mathcal{V}}_2$. According to [K1],[K2], the matrices $\boldsymbol{\Upsilon}_{\rm tr}$ and $\boldsymbol{\mathcal{S}}_{\rm tr}$ may be directly calculated for $\mathcal{O}(N^2)$ operations, the system with the matrix $\boldsymbol{\mathcal{S}}_{\rm tr}$ is solved by the direct methods for $\mathcal{Q}_2$ operations, where $\mathcal{Q}_2 = \dim \mathcal{V}_2$ and, thus, is $\mathcal{O}(N \log N)$. By the operation $\boldsymbol{\Upsilon}^{-1}\mathbf{v}^{(2)}$ is implied the following sequence: the transformation of $\mathbf{v}^{(2)} \rightarrow \mathbf{v}_{\rm tr}^{(2)} = \boldsymbol{\mathcal{P}}^{-1}\mathbf{v}^{(2)}$ to the trigonometric basis by 1-d FDFT; solving the system $\boldsymbol{\Upsilon}_{\rm tr}\mathbf{u} = \mathbf{v}_{\rm tr}^{(2)}$ by the iterative procedure with the preconditioner $\boldsymbol{\mathcal{S}}_{\rm tr}$. In the case of the operation $\mathbf{C}_2^{-1}\mathbf{v}^{(2)}$ the matrix $\mathbf{C}_2$ is given, and the multiplication by it is cheap. Thus, we use the evident alternative procedure. The computational cost of 1-d FDFT needed for the transformation to the trigonometric basis and back is $\mathcal{O}(N(\log N)^2)$.

Now, we describe the arrangement of the trigonometric basis. Let us introduce the notations T_j , $j = 1, 2, .., l_0^2$, for different sides $T_{s,t} \in \pi_1$. The space $\mathcal{V}_2$ is the direct sum of the subspaces $\mathcal{V}_{2,j}$, each corresponding to the nodes of one edge T_i. The trigonometric basis in $\mathcal{V}_{2,j}$ is formed by the vectors $\boldsymbol{\mu}_{2,q,j} = (\sin(\pi q g/n_j))$, $q, g = 1, 2, .., n_j$, where $n_j = \dim[\mathcal{V}_{2,j}]$. Therefore, 1-$d$ FDFT is completed for each edge T_j independently in parallel. Let us also note, that at the choice 1), $\boldsymbol{\mathcal{S}}_{\rm tr}^{(k,l)}$ is block-diagonal with $2(N + 1 - l_0)$ independent blocks, every of which is a three-diagonal matrix $(l_0 - 1) \times (l_0 - 1)$. At the choice 2), $\boldsymbol{\mathcal{S}}_{\rm tr}$ is a diagonal matrix.

5.2. Multilevel solvers, complete and incomplete reference elements

5.2.1. `Algorithms.` The algorithm, which is presented here, was suggested for the reference element $\mathcal{E}_H$ in [Be1]. We will see that, with necessary changes, it is applicable to the case of the incomplete reference element $\mathcal{E}_{\Im,H}$. For the completeness we consider below the both cases in parallel.

Let us subdivide each square nest of the fine mesh in the two triangles by the diagonal passing through its left upper and the right lower vertices. Let also $\mathcal{H}^0(\pi_1)$ be the space of the piece-wise linear functions on π_1 vanishing on $\partial\pi_1$, $\mathbf{B}_{e,e}$ be the matrix of the bilinear form

$$a^{(1)}_{\pi_1}(u,v) := \int_{\Omega_1} (y^2 u_x v_x + x^2 u_y v_y) \quad \text{for all } \ u,v \in \mathcal{H}^{(0)}(\pi_1) \tag{5.5}$$

with the choice of the nodal basis in $\mathcal{H}^0(\pi)$, and $\mathbf{\Lambda}_{e,e}$ be now the matrix, which is obtained from $\mathbf{\Lambda}_{e,e}$, introduced in Section 5.1, by subtracting the diagonal matrix $\boldsymbol{\mathcal{D}}_\varphi = \mathrm{diag}[\varphi_i/\varphi_j + \varphi_j/\varphi_i]$.

The multigrid method, which we describe here, seem to be adjustable to incomplete, transition and orthotropic elements. However, in order to simplify the presentation, we consider the incomplete reference element $\mathcal{E}_{\Im,H}$ with $\omega_{\Im,o} = \widehat{\omega}_o$. In this case, as in the case of the complete reference element, the stiffness matrix may be represented in the form $\mathbf{A}_1 = \mathrm{d}iag\,[\mathbf{A}_{e,e}, \mathbf{A}_{e,o}, \mathbf{A}_{o,e}, \mathbf{A}_{o,o}]$, but now the blocks $\mathbf{A}_{a,b}$ have different dimensions. Nevertheless, they possess similar properties and the solver of the same type, as for the block $\mathbf{A}_{e,e}$, is applicable to all of them with the same asymptotically computational cost.

We retain the notations $\mathbf{\Lambda}_{e,e}$ and $\mathbf{B} = \mathbf{B}_{e,e}$ for the preconditioners of the matrix $\mathbf{A}_{e,e}$. Let π_2 be the half of π_1, which is below the diagonal of the square π_1, joining upper left and lower right vertices, $a^{(1)}_{\pi_2}(\ ,\ )$ be the restriction of the bilinear form $a^{(1)}_{\pi_1}(\ ,\ )$ to π_2 and $\mathcal{H}^0(\pi_2)$ be the space of functions from $\mathcal{H}^0(\pi_1)$, vanishing outside of π_2. The matrix $\mathbf{B}$ is defined as induced by the bilinear form $a^{(1)}_{\pi_2}(\ ,\ )$ on the space $\mathcal{H}^0(\pi_2)$ with the nodal basis. By $\mathbf{\Lambda}_{e,e}$ is understood the matrix, which is obtained from defined above $\mathbf{\Lambda}_{e,e}$ for the complete reference element in the following way: we cross out the rows and columns, corresponding to the coordinate functions $L_{\boldsymbol{\alpha}}$, $\alpha_1 + \alpha_2 > p$.

We will present the algorithm and the results on its properties for the both reference elements $\mathcal{E}_H$ and $\mathcal{E}_{\Im,H}$ simultaneously, pointing out the distinctions, when it is necessary.

Lemma 3. *For c_k independent of p, there are hold the inequalities*

$$c_1\mathbf{B} \le c_3\mathbf{\Lambda}_{e,e} \le \mathbf{A}_{e,e} \le c_2\log(1+p)\mathbf{\Lambda}_{e,e} \le c_4\log(1+p)\mathbf{B}\,. \tag{5.6}$$

It is sufficient to prove (5.6) for the element $\mathcal{E}_H$. In the case of this element, the proof, is the consequence of the results of [KJ1] and [Be].

Without loss of generality, we assume that $N+1 = 2^{-l_0}$ and introduce the sequence of l_0 imbedded orthogonal meshes of the sizes $\hbar_l = 2^{-l}$. The sets of the nodes $\mathbf{x} = \hbar_l(i,j)$ in π_k of these meshes are denoted by X_l. To this sequence correspond the sequences of the finite element spaces $\mathcal{H}_l^0(\pi_k)$ and finite element matrices $\mathbf{B}_l$, such that $\mathcal{H}_{l_0}^0(\pi_k) = \mathcal{H}^0(\pi_k)$ and $\mathbf{B}_{l_0} = \mathbf{B}$.

We also will need the notations:

$\mathcal{W}_l(\pi_k) := \mathcal{H}_l(\pi_k) \ominus \mathcal{H}_{l-1}(\pi_k)$, and $\mathbf{B}_{l,w}$ is the block on the diagonal of the matrix $\mathbf{B}_l$ corresponding to the subspace $\mathcal{W}_l(\pi_k)$,

U_l and W_l are the spaces of vectors with the entries which are the nodal values of functions from $\mathcal{H}_l^0(\pi_k)$ and from $\mathcal{W}_l(\pi_k)$ at the nodes of the sets X_l and $X_{l,w} := X_l \setminus X_{l-1}$, respectively,

$\mathbb{P}_{l-1} : U_{l-1} \to U_l$ is the interpolation matrix, such that, if $v \in \mathcal{H}_{l-1}^0(\pi_k)$ and $\mathbf{v}^{(l-1)} \in U_{l-1}$ is its vector representation, then $\mathbf{v}^{(l)} := \mathbb{P}_{l-1}\mathbf{v}^{(l-1)}$ is the vector representation of v in U_l,

$\mathbb{R}_l : U_l \to W_l$ is the matrix defining the restrictions $\mathbf{v}^{(l-1)} := \mathbb{R}_l\mathbf{v}^{(l)}$ of the vector $\mathbf{v}^{(l)}$ to the set of nodes $X_{l,w}$.

Suppose, we have good preconditioners $\boldsymbol{\mathcal{B}}_{l,w}$ for the matrices $\mathbf{B}_{l,w}$. Then, one multigrid itreration for the system

$$\mathbf{B}_l\mathbf{u} = \mathbf{f}\ , \tag{5.7}$$

producing $\mathbf{u}^{k+1} := \mathbf{Mgm}(l, \mathbf{u}^k, \mathbf{f})$ from a given approximation $\mathbf{u}^k$, is implemented as follows.

- **If** $l \ge 1$, then **do**
  1. Presmoothing in the space W_l:
     $\mathbf{v} := \mathbf{u}^k$;
     do ν **times** $\quad \mathbf{v} := \mathbf{v} - \tau^{-1}\mathbb{R}_l^T\boldsymbol{\mathcal{B}}_{l,w}^{-1}\mathbb{R}_l(\mathbf{B}_l\mathbf{v} - \mathbf{f})$;
  2. Coarse grid correction in the space V_{l-1}:
     $\mathbf{d}_{l-1} := \mathbb{P}_{l-1}^T(\mathbf{f} - \mathbf{B}_l\mathbf{v})$; $\mathbf{w}^0 = 0$;
     do μ_{l-1} **iterations** $\mathbf{w}^k = \mathbf{M}gm(l-1, \mathbf{w}^{k-1}, \mathbf{d}_{l-1})$;
     $\mathbf{v} := \mathbf{v} + \mathbb{P}_{l-1}\mathbf{w}^{\mu_{l-1}}$;

3. Postsmoothing in the space W_t:

do ν **times** $\quad\quad \mathbf{v} := \mathbf{v} - \tau^{-1}\mathbb{R}_l^T \boldsymbol{\mathcal{B}}_{l,w}^{-1}\mathbb{R}_l(\mathbf{B}_l\mathbf{v} - \mathbf{f})$;

$\mathbf{u}^{i+1} = \mathbf{v}$

- **else**

  solve (5.7) **by the exact method**

- **endif**

Here τ are the iteration parameters, which in the simplest case are taken constant and independent of l.

The choice of the preconditioner $\boldsymbol{\mathcal{B}}_{l,w}$ is very important. The optimal preconditioner, one of which is presented below, assumes the spectral equivalence with the matrix $\mathbf{B}_{l,w}$ and existence of the optimal solver for $\boldsymbol{\mathcal{B}}_{l,w}$. We denote by $\Im_i$ the line, which passes through the nodes $\max\,(\alpha_1, \alpha_2) = i$. In the case of the reference element $\mathcal{E}_H$, the line is composed of the two orthogonal segments of equal length: the segment $[0, i\hbar_l]$ on the line $x_2 = i\hbar_l$ and the segment $[0, i\hbar_l]$ on the line $x_1 = i\hbar_l$. In the case of the reference element $\mathcal{E}_{\Im,H}$, it also contains the two segments, which are the intersection of the line, defined above, with π_2. Evidently, for $i > 2^{l-1}$ the segments are disjoint.

The matrix $\mathbf{B}_{l,w}$ may be represented in the block form

$$\mathbf{B}_{l,w} = \mathrm{tridiag}\,[\mathbf{B}_{l,w}^{(i,i-1)},\ \mathbf{B}_{l,w}^{(i)},\ \mathbf{B}_{l,w}^{(i,i+1)}]_{i=1,2,..,N}\ , \tag{5.8}$$

with each block $\mathbf{B}_{l,w}^{(i)}$ corresponding to the unknowns at the nodes on the line $\Im_i$. It is easy to note that the blocks $\mathbf{B}_{l,w}^{(2k)}$ with the even numbers are diagonal and the blocks $\mathbf{B}_{l,w}^{(2k+1)}$ with the odd numbers are tridiagonal. Having set

$$\boldsymbol{\mathcal{B}}_{l,w} = \mathrm{diag}\,[\mathbf{B}_{l,w}^{(i)}]\,, \tag{5.9}$$

we are able to prove that

$$c_1\boldsymbol{\mathcal{B}}_{l,w} \le \mathbf{B}_{l,w} \le c_2\boldsymbol{\mathcal{B}}_{l,w} \tag{5.10}$$

with the constants independent of p and l.

The important factor for convergence of the described multigrid is the strengthened Cauchy inequality

$$a^{(1)}(u,w) \le c_0\sqrt{a^{(1)}(u,u)a^{(1)}(w,w)}\ , \quad \forall u \in \mathcal{H}_{l-1}(\pi_1) \quad \forall w \in \mathcal{W}_l(\pi_1)\ . \tag{5.11}$$

which is fulfilled with $c_0 \le 1/2$. For the case of the reference element $\mathcal{E}_H$, the inequalities (5.10) and (5.11) were proved by Beuchler [Be]. These inequalities retain for the reference element $\mathcal{E}_{\Im,H}$.

Theorem 2. *Let $\tau = 2/(c_1 + c_2)$, $\mu \ge 3$ and ν be greater than some $\nu_o(c_0, c_1, c_2)$. Then the convergence factor*

$$\rho_{l,\mathrm{mult}} := sup_{\;\mathbf{u}^k \in U_l} \; \|\mathbf{u}^{k+1} - \mathbf{u}\|_{\;\mathbf{B}_l} / \|\mathbf{u}^k - \mathbf{u}\|_{\;\mathbf{B}_l} \tag{5.12}$$

is estimated by the constant $\rho < 1$ which does not depend on p, l and $\mathbf{u}^k$.

With assumptions that (5.10) and (5.11) hold, Theorem 2 was proved independently by Schieweck [S] and Pflaum [Pf]. On the basis of easily obtained estimates for the constants c_k, $k = 0, 1, 2$, results of these works allow also to estimate ν. Under conditions of Theorem 2, it is sufficient to take $\nu \ge 3$.

Let $\mathbf{M}g_{\mu,\varkappa}$ be the linear operator defined by a fixed number $\varkappa$ of multigrid iterations, *e.g.,* for $\varkappa = 2$, $\mathbf{M}g_{\mu,\varkappa}\mathbf{v} = \mathbf{Mgm}(l_0, \mathbf{Mgm}(l_0, \mathbf{v}, \mathbf{f})\mathbf{f})$.

Corollary 1. *Let $\mu = 3$ and $\nu \ge 3$. Then*

$$c_1 \mathbf{M}g^{-1}_{\mu,\varkappa} \le \mathbf{A}_{e,e} \le c_2(1 + \log p)\mathbf{M}g^{-1}_{\mu,\varkappa} \,, \tag{5.13}$$

with the constants depending only on $\varkappa$.

Counting the number of operations for solving the systems with the matrix $\mathbf{A}_{e,e}$ by PCG with the preconditioner $\mathbf{M}^{-1}_{\mu,\varkappa}$, we conclude that we obtained an almost optimal solver for such systems, *i.e.,* it requires $\mathcal{O}(N^2(\log N)^{1/2})$ arithmetic operatons.

The above results rest on the fact that coupling between the unknowns inside of each line $\Im_i$ is stronger, than between unknowns of different lines. This basic fact allows to derive fast multilevel solvers in several other ways. In the above multigrid solver we applied smoothings only in the subspaces W_l with the smoother $\mathbf{T}_{l,w} = I - \tau \boldsymbol{\mathcal{B}}^{-1}_{l,w}\mathbf{B}_{l,w}$. A competing algorithm is obtained, if smoothing in the space W_l is replaced by the smoothing in the space U_l with the smoother $\mathbf{T}_l = I - \tau \boldsymbol{\mathcal{B}}_l^{-1}\mathbf{B}_l$, where the preconditioner is defined in a way close to the definition of $\boldsymbol{\mathcal{B}}_{l,w}$. Namely, we set

$$\boldsymbol{\mathcal{B}}_l = \mathrm{diag}\; [\mathbf{B}_l^{(i)}]_{i=1,2,..,N} \,, \tag{5.14}$$

with the blocks $\mathbf{B}_l^{(i)}$ appearing in the representation

$$\mathbf{B}_l = \mathrm{tridiag}\; [\mathbf{B}_l^{(i,i-1)},\; \mathbf{B}_l^{(i)}, \mathbf{B}_l^{(i,i+1)}]_{i=1,2,..,N} \,. \tag{5.15}$$

Let us note, that again each block $\mathbf{B}_l^{(i)}$ on diagonal corresponds to the unknowns at the nodes of one line $\Im_i$, but now all these blocks are tridiagonal. Our numerical experiments, as experiments of [Be1,Be2] show that this smoother is even more efficient.

Another alternative are the solvers for the systems (5.7), which are derived, if we use the pointed out above basic property in the general framework of AMLI (algebraic multilevel iterations) approach, see [AV]. We will briefly describe one of such solvers, emphasizing that it may be applied to incomplete and orthotropic elements. Let us introduce the matrix $\widehat{\boldsymbol{\mathcal{B}}}_{w,l} := c_2^{-1}\boldsymbol{\mathcal{B}}_{w,l}$ with the constant c_2 from (5.10) and the polynomial $P_2(t) = (1-2t/(1-\alpha))^2$ with α defined as the smallest positive solution of the polynomial equation

$$1-\gamma = (\frac{c_2}{c_1}-1)t + \frac{1}{4}\left(\frac{(1+\sqrt{t})^2+(1-\sqrt{t})^2}{\sum_{s=1}^{2}(1+\sqrt{t})^{2-s}(1-\sqrt{t})^{s-1}}\right)^2 .$$

Also, we represent the matrices $\mathbf{B}_l$ and $\mathbb{P}_{l-1}$ in the block form, corresponding to to the splitting $V_l = V_{l-1} \oplus W_l$:

$$\mathbf{B}_l = \begin{pmatrix} \mathbf{B}_{v,l} & \mathbf{B}_{vw,l} \\ \mathbf{B}_{wv,l} & \mathbf{B}_{w,l} \end{pmatrix}, \qquad \mathbb{P}_{l-1} = \begin{pmatrix} I \\ \mathbb{P}_{w,l-1} \end{pmatrix}.$$

The AMLI-type preconditioner for the matrix $\mathbf{B}_l$ is defined as

$$\boldsymbol{\mathcal{C}}_l = \begin{pmatrix} \mathbf{B}_l(I-P_2(\boldsymbol{\mathcal{C}}_{l-1}^{-1}\mathbf{B}_l))^{-1} & \mathbf{B}_{vw,l}+\mathbb{P}_{w,l}^T(\mathbf{B}_{w,l}-\widehat{\boldsymbol{\mathcal{B}}}_{w,l}) \\ \mathbf{0} & \widehat{\boldsymbol{\mathcal{B}}}_{w,l} \end{pmatrix} \times$$
$$\times \begin{pmatrix} I & \mathbf{0} \\ \widehat{\boldsymbol{\mathcal{B}}}_{w,l}^{-1}(\mathbf{B}_{wv,l}+(\mathbf{B}_{w,l}-\widehat{\boldsymbol{\mathcal{B}}}_{w,l})\mathbb{P}_{w,l}) & I \end{pmatrix}.$$

Theorem 3. *The matrices* $\mathbf{B}_l$ *and* $\boldsymbol{\mathcal{C}}_l$ *are spectrally equivalent,* i.e.,

$$c_1\boldsymbol{\mathcal{C}}_l \le \mathbf{B}_l \le \boldsymbol{\mathcal{C}}_l\ , \tag{5.16}$$

with the constant independent of p.

This theorem is a direct consequence of the results of [AL] on convergence of AMLI and the estimates for c_k, $k = 0, 1, 2$.

In order to solve the system with the matrix $\boldsymbol{\mathcal{C}}_l$, it is necessary to solve three times the systems with the matrix $\widehat{\boldsymbol{\mathcal{B}}}_{w,l}$ and twice the systems with the matrix $\boldsymbol{\mathcal{C}}^{(l)} := \mathbf{B}_l(I-P_2(\boldsymbol{\mathcal{C}}_{l-1}^{-1}\mathbf{B}_l))^{-1}$. The computational costs of solving the systems with the matrices $\widehat{\boldsymbol{\mathcal{B}}}_{w,l}$ and $\mathbf{B}_{w,l}$ are the same. Solving the system with the matrix $\boldsymbol{\mathcal{C}}^{(l)}$ requires, evidently, solving two systems with the matrix $\boldsymbol{\mathcal{C}}_{l-1}$. Estimating the number of arithmetic operations and taking into account Lemma 1, we come to the conclusion that the preconditioner $\boldsymbol{\mathcal{C}}_l$ provides an almost optimal solver for the system with the matrix $\mathbf{A}_1$.

`5.2.2. Numerical results.` Some of the algorithms were tested numerically and demonstrated very good agreement with *a priori* estimates. Indeed, in many cases, the numerical results demonstrate a better performance, than it was predicted theoretically. In this section, we describe the part of the results, obtained by means of multigrid solvers[6] and summarized in the table below.

The systems with the independent blocks $\mathbf{A}_{a,b},\ a,b = e,o$, of the stiffness matrix $\mathbf{A}_I$ of the reference element $\mathcal{E}_H$ were solved by the PCG with the multigrid preconditioner $\mathbf{M}_{\mu,\varkappa}$ with $\mu = 3,\ \nu = 3$. Instead of specifying $\varkappa$, the stopping criteria for multigrid iterations was reducing the preconditioned energy norm $(\mathbf{e}^T\mathbf{A}_{e,e}\mathbf{B}^{-1}\mathbf{A}_{e,e}\mathbf{e})^{1/2}$ of the error $\mathbf{e}$ in 10^9 times. The notations in the table are the following:

$p = 2N + 1$ - the highest degree of polynomials in the hierarchical basis,

$n = N^2$ - the number of the unknowns of the systems with the matrices $\mathbf{A}_{a,b},\ a,b = e,o$,

l - number of the fine grid, so that $N = 2^l$,

i_1,i_2 and i_3 - numbers of PCG-iterations for solving the systems with the matrices $\mathbf{A}_{e,e},\ \mathbf{A}_{e,o}$and $\mathbf{A}_{o,o}$, respectively,

M - the volume of the computer dynamic memory in bytes, used at solution of the system with one of the blocks $\mathbf{A}_{e,e}$,

$t_1,\ t_2$ and t_3 - times in seconds, spent for solving the systems $\mathbf{A}_{e,e},\ \mathbf{A}_{e,o}$and $\mathbf{A}_{o,o}$, respectively. The computer codes were run on PENTIUM-III, 750MHz with 256MB of operative memory.

[6] Computer codes and numerical experiments were completed by M. Nikitin

l	p	n	i_1	i_2	i_3	M	t_1	t_2	t_3
2	7	9	6	9	6	448	0	0	0
3	15	49	15	14	12	2,032	0	0	0
4	31	225	18	17	14	8,768	0.01	0.01	0.01
5	63	961	21	18	15	36,560	0.05	0.05	0.04
6	127	3,969	22	20	17	149,472	0.37	0.34	0.29
7	255	16,129	23	21	18	604,656	1.92	1.76	1.52
8	511	65,025	24	22	19	2,432,512	10.26	9.47	8.22
9	1,023	261,121	24	22	20	9,758,224	47.69	43.98	40.57
10	2,047	1,046,529	24	23	21	39,089,696	210.07	202.23	185.79
11	4,095	4,190,209	25	23	21	156,472,880	1,162.64	1,068.26	971.366

We see that the rate of the growth from $l = 2$ to $l = 3$, – i.e., for the rude fine grids evidently laying in the preasymptotic range, – of the number of PCG iterations is faster than $(\log N)^{1/2}$. Then, from $l = 4$ it is becoming slower, than $(\log N)^{1/2}$, predicted by the *a priori* estimates. In the asymptotic range, the computer time is growing slower, than $N^2(\log N)^{1/2}$. The results also clearly show, that all the blocks $\mathbf{A}_{e,e}$, $\mathbf{A}_{e,o}$and $\mathbf{A}_{o,o}$and the preconditioner $\mathbf{M}_{\mu,\varkappa}$ are indeed spectrally equivalent, although the chosen particular form $\mathbf{B}$ seem to be a little more adapted to $\mathbf{A}_{o,o}$, than to $\mathbf{A}_{e,e}$. The performance of the computer code seem to be improvable for the following reason. Practically, the multigrid solver was used as an exact one, since the stopping criteria required very high accuracy. Use of the appropriately chosen, *e.g.,* optimized, $\varkappa$ for exit from multigrid iterations may significantly reduce the computer time.

The results given above overlap with the results of numerical experiments, presented in [Be,Be1] with a very good correspondence between the both.

6 Fast solvers for Dirichlet problems on Lagrange p-version reference elements

6.1. DD-like solver.

The same ideas, as presented for the hierarchical *hp*-version, may be implemented for derivation of *LS*'s in the case of the Lagrange reference elements with GLC and GLL nodes, often used in the finite element and collocation methods. A nice property of such discretizations is that they may be efficiently preconditioned by the finite element matrices, generated by the assemblage of linear triangular elements with the same grids of vertices [P,KP]. Without loss of generality, we may assume $p = 2N$, $N = 2^{l_0-1}$. The step sizes $\hbar_k := \eta_k - \eta_{k-1}$ of the GLC grid $\eta_k = \cos((p-k)\pi/p)$, $k = 0, 1, .., p$, satisfy the relationship $\hbar_k \asymp k/p^2$, $k \le N$. Since the step sizes of the GLL grid has the same asymptotic, we may cover the both cases and many others, if we define the preconditioner on the grid

$$\eta_0 = -1\,, \quad \eta_i = \eta_{i-1} + \hbar_i\,, \quad \eta_N = 0\,, \quad \hbar_{i-1} \le \hbar_i \le i^\gamma/\aleph\,, \quad \gamma \ge 0\,, \quad \aleph = \sum_{i=1}^{N} i^\gamma\,, \tag{6.17}$$

symmetrically continued on $(0, 1)$.

Suppose, a grid $x_k = \eta_i$, $k = 1, 2$, is given. We subdivide each rectangular nest of this mesh in the two triangles by one of its diagonals and define the space $\mathcal{H}(\tau_o)$ of continuous piece-wise linear functions, vanishing on $\partial\tau_o$. The notations $\boldsymbol{\mathcal{A}}_G, \boldsymbol{\mathcal{A}}$ stand for the matrices, generated by the bilinear form

$$a_{\tau_o}(u, v) = \int_{\tau_o} \nabla u \cdot \nabla v \, d\mathbf{x}$$

on the spaces $\mathcal{H}(\tau_o)$, corresponding to the GLL or GLC grid and the grid (6.17) with $\gamma = 1$, respectively. Since, fast solvers for the matrices $\boldsymbol{\mathcal{A}}_G$ generated on GLL and GLC grids are quite similar, for definiteness we relate our considerations to GLL grids, more popular in the finite element method.

Lemma 4. *Let* $\mathbf{A}$ *be the stiffness matrix for the Lagrange reference element with the GLL grid of nodes. Then* $\mathbf{A} \asymp \boldsymbol{\mathcal{A}}_G \asymp \boldsymbol{\mathcal{A}}$.

The relation $\mathbf{A} \asymp \boldsymbol{\mathcal{A}}_G$ may be found in [Cas], for the proof see also [BM, Can]. The relation $\boldsymbol{\mathcal{A}}_G \asymp \boldsymbol{\mathcal{A}}$ is a direct consequence of the asymptotic equivalence of the grids.

Properties of the matrix $\boldsymbol{\mathcal{A}}_I$, which is the restriction of $\boldsymbol{\mathcal{A}}$ to the set of internal nodes, differ from properties of, *e.g.*, the preconditioner $\boldsymbol{\Lambda}_{e,e}$, used in the case of the hierarchical reference element. Nevertheless, fast solvers for the both may be derived in quite similar ways. Here, we briefly describe the domain decomposition type almost optimal solver.

The mesh, defined by the numbers η_i, is termed *fine* mesh. The numbers, denoted $\zeta_0 = -1$, ζ_k , $k = 1, 2, .., l_0$, $\zeta_{l_0} = 0$, will introduce an imbedded *coarse* mesh with the steps denoted $\varrho_k = \zeta_k - \zeta_{k-1}$. We use the notation $i(k)$ for the integers, which satisfy equalities $\eta_{i(k)} = \zeta_k$.

Proposition 1. *For any fixed* c_0 *and* n_1*, there exist* ζ_k*,* $k = 1, 2, 3, .., l_0$*, such that* $\zeta_1 = -1 + n_1\aleph^{-1}$*,* $l_0 \leq \gamma(\log c_0)\log N$ *and*

$$\hbar_i/\hbar_{i-1} \leq c_0\,, \quad \text{for all} \quad i_{k-1} < i \leq i_k\,, \quad k = 1, 2, .., l_0\,.$$

Such a grid is easily defined and symmetrically continued on the whole interval (-1,1). Nests of the corresponding tensor product grid are denoted as $\delta_{k,l} = (\mathbf{x} : \zeta_{k-1} < x_1 < \zeta_k\,, \zeta_{l-1} < x_2 < \zeta_l)$. Let $n_k = i(k) - i(k-1)$. We introduce the *piece-wise uniform fine grid* $x_1, x_2 = \vartheta_i$, such that $\vartheta_{i(k)} = \eta_{i(k)} = \zeta_k$, and $\vartheta_i = \vartheta_{i(k-1)} + (i - i(k-1))\sigma_i$, $\sigma_i = \varrho_k/n_k$ for $i(k-1) < i \leq i(k)$. The space of piece-wise linear continuous functions $\mathcal{H}(\tau_o)$, corresponding to this grid, induces the finite element matrix, which is denoted by $\mathbf{B}$. For the blocks of this and other matrices, similar to introduced earlier indices are used.

We define the DD_l preconditioner $\boldsymbol{\mathcal{K}}_{I,o}$ exactly by the same formula, which introduced $\boldsymbol{\mathcal{K}}_o$ in subsection 5.1.1, *i.e.,* we set $\boldsymbol{\mathcal{K}}_{I,o} = \mathbb{P}_o^T \boldsymbol{\mathcal{D}} \mathbb{P}_o$, but with newly defined $\mathbf{B}$. We also see that absolutely similar and with the same asymptotic cost procedures my be used for solving the systems with this preconditioner $\boldsymbol{\mathcal{K}}_{I,o}$. Indeed, the matrices $\mathbf{B}_I^{(k,l)}$ are again such that for solving systems with these matrices, we can use 2-d FDFT's for the total for all subdomains $\delta_{k,l}$ cost $\mathcal{O}(p^2 \log p)$ arithmetic operations. Optimal BPX-like preconditioners for $\mathbf{B}_I^{(k,l)}$, which were denoted $\mathbb{B}_I^{(k,l)}$, are also constructed in the same way, as in subsection 5.1.2. The Schur complement preconditioning and solving the interface problem implements exactly the same, as in subsection 5.1.2, approach. It requires 1-d FDFT of the vectors to the trigonometric basis and back and use of directly calculated matrices $\mathbf{S}_{\rm tr}$ and $\boldsymbol{\mathcal{S}}_{\rm tr}$. The costs of FDFT, calculation of these matrices and solving systems with the matrix $\boldsymbol{\mathcal{S}}_{\rm tr}$ are $\mathcal{O}(p(\log p)^2)$ and p^2 and $\mathcal{O}(p \log p)$, respectively, whereas $\mathrm{cond}\,[\boldsymbol{\mathcal{S}}_{\rm tr}^{-1}\mathbf{S}_{\rm tr}] \prec (1 + \log p)^3$. Thus, the estimate of Theorem 1 of the total computational work remains also valid.

Theorem 4. *Let* $\mathbf{K}_{I,r}$ *be the internal stiffness matrix of an element of hp-discretization with Lagrange reference element, which nodes be defined by the numbers (5.2). Then, the total cost of solving the system* $\mathbf{K}_{I,r}\mathbf{u}_{I,r} = \mathbf{f}_{I,r}$ *by the PCG with the preconditioner* $\boldsymbol{\mathcal{K}}_{I,o}$ *is* $\mathcal{O}(N^2(\log N)^{0.5\nu+3})$.

The extension of the more efficient solver, described in 5.1.2, requires adjusting of the preconditioner $\mathbf{C}$, which we do not discuss in this paper.

6.2. Multigrid solver.

Without loss of generality, we may assume $p = 2N$, $N = 2^{m-1}$, and introduce the sequence of m imbedded orthogonal meshes with the sets of nodes, denoted by $X_l = (\mathbf{x}_{l,i,j} = (x_{1,l,i}, x_{2,l,j})\ ,\ i,j = 0,1,..,2^l)$ and the set $X_m = (\mathbf{x} = (\eta_i, \eta_j))$ defined by the numbers (6.17). We assume that $x_{k,l,n} > x_{k,l,n-1}$ and that between each pair of mesh lines $x_k = x_{k,l,n}, x_{k,l,n-1}$ there is a mesh line $x_k = x_{k,l+1,2n-1}$. This sequence induces the sequences of the spaces $\mathcal{H}_l(\tau_o)$ and matrices $\mathbf{B}_l$, where $\mathcal{H}_l(\tau_o)$ is the space of continuous piece-wise bilinear functions on the mesh X_l, which vanish on $\partial\tau_o$, and $\mathbf{B}_l$ is the corresponding $(2^l-1)^2 \times (2^l-1)^2$ finite element matrix. The nodal basis function in $\mathcal{H}_l(\tau_o)$, which is equal unity at the node $\mathbf{x}_{l,i,j}$, is denoted as $\phi_{ij}^{(l)}$. Evidently, $\mathcal{H}_m(\tau_o) = \mathcal{H}(\tau_o)$ and $\mathbf{B}_m = \boldsymbol{\mathcal{A}}$.

We also introduce similar to the used in Section 5.2 notations:

V_l and W_l are the spaces of vectors with the entries, specified on the sets of nodes X_l and $X_{w,l} := X_l \setminus X_{l-1}$, so that $V_l = V_{l-1} \oplus W_l = W_l \oplus W_{l-1} \oplus ... \oplus W_2 \oplus V_1$;

$\mathbf{P}_{l-1} : V_{l-1} \to V_l$ is the interpolation matrix, such that, if $v \in \mathcal{H}_{l-1}(\tau_o)$ and $\mathbf{v}_{l-1} \in V_{l-1}$ is its vector representation, then $\mathbf{v}_l = \mathbf{P}_{l-1}\mathbf{v}_{l-1}$ is the vector representation of v in V_l;

$\mathbf{R}_l : V_l \to W_l$ is the restriction matrix such that $\mathbf{v}_{l-1} := \mathbf{R}_l\mathbf{v}_l$ contains the entries of $\mathbf{v}_l$ for the nodes of $X_{w,l}$;

$\mathbf{B}_{v,l}$, $\mathbf{B}_{w,l}$ and $\boldsymbol{\mathcal{B}}_{w,l}$ are the blocks on the diagonal of $\mathbf{B}_l$, corresponding to the subspaces V_l and W_l, and the preconditioner for $\mathbf{B}_{w,l}$, respectively.

One multigrid itreration for the system $\mathbf{B}_l\mathbf{u} = \mathbf{F}$ producing $\mathbf{u}^{k+1,l} := \mathbf{Mgm}(l, \mathbf{B}_l, \mathbf{F}, \mathbf{u}^{k,l})$ from a given approximation $\mathbf{u}^{k,l}$ is described as follows:

If $l \ge 1$, then **do**

1. Presmoothing in the space W_l:
   $\mathbf{v} := \mathbf{u}^{k,l}$;
   do ν **times** $\quad \mathbf{v} := \mathbf{v} - \tau^{-1}\mathbf{R}_l^t\boldsymbol{\mathcal{B}}_{w,l}^{-1}\mathbf{R}_l(\mathbf{B}_l\mathbf{v} - \mathbf{F})$;
2. Coarse grid correction in the space V_{l-1}:
   $\mathbf{d}_{l-1} := \mathbf{P}_{l-1}^t(\mathbf{F} - \mathbf{B}_l\mathbf{v})$; $\mathbf{w} = 0$;
   do μ_{l-1} **iterations** $\mathbf{w} = \mathbf{M}gm(l-1, \mathbf{B}_{l-1}, \mathbf{d}_{l-1}, \mathbf{w})$;
   $\mathbf{v} := \mathbf{v} + \mathbf{P}_{l-1}\mathbf{w}$;
3. Postsmoothing in the space W_l:
   do ν **times** $\quad \mathbf{v} := \mathbf{v} - \tau^{-1}\mathbf{R}_l^t\boldsymbol{\mathcal{B}}_{w,l}^{-1}\mathbf{R}_l(\mathbf{B}_l\mathbf{v} - \mathbf{F})$;
   $\mathbf{u}^{k+1,l} = \mathbf{v}$

else

solve $\mathbf{B}_l\mathbf{u} = \mathbf{F}$ **by the exact method**

endif

Here, τ is the iteration parameter, which in a simplest case may be taken constant. We may assume also that $\tau = 2/(\lambda_{min} + \lambda_{max})$, where λ_{min} and λ_{max} are minimal nonzero and maximal eigenvalues of the generalized eigenvalue problem $\mathbf{B}_{l,w}\mathbf{v} = \lambda \boldsymbol{\mathcal{B}}_{w,l}^{-1}\mathbf{v}$ or sufficiently good bounds for them from below and above, respectively. We see that in the accepted notations the algorithm is similar to the introduced in Section 5.2.1. However, essential difference of the discrete problems evinces in considerable differences in algorithms and in the results on convergence.

Again, the strengthened Cauchy inequality is in the base of the convergence analysis. We will show that there exists $\gamma \in [0, 1)$ such, that

$$\left(\mathbf{v}_{l-1}^t \mathbf{B}_l \mathbf{v}_{w,l}\right)^2 \le \gamma^2 \left(\mathbf{v}_{l-1}^t \mathbf{B}_{l-1} \mathbf{v}_{l-1}\right) \left(\mathbf{v}_{w,l}^t \mathbf{B}_{w,l} \mathbf{v}_{w,l}\right) \tag{6.18}$$

holds for all $\mathbf{v}_{l-1} \in V_{l-1}$ and $\mathbf{v}_{w,l} \in W_l$ with γ independent of p and l.

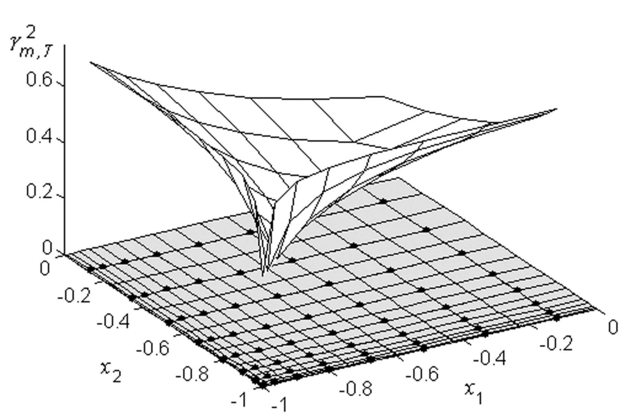
For bounding γ from above, we introduce the so called two-level hierarchical basis element stiffness matrices, which are denoted $\mathbf{C}_{l,T}$. Let T be a rectangular nest of the mesh X_{l-1} and $a_T(\,,\,)$ is the restriction of the bilinear form $a_{\tau_0}(\,,\,)$ to T. If to represent $\mathbf{C}_{l,T}$ in the block form

$$\mathbf{C}_{l,T} = \begin{pmatrix} \mathbf{C}_{l,T;1} & \mathbf{C}_{l,T;12} \\ \mathbf{C}_{l,T;21} & \mathbf{C}_{l,T;2} \end{pmatrix},$$

then

$$\begin{aligned}
\mathbf{C}_{l,T;1} &= \left\{ a_T(\phi_{ij}^{(l)}, \phi_{kn}^{(l)}) \right\}_{\mathbf{x}_{l,i,j},\ \mathbf{x}_{l,k,n} \in X_{w,l} \cap T}, \\
\mathbf{C}_{l,T;2} &= \left\{ a_T(\phi_{ij}^{(l-1)}, \phi_{kn}^{(l-1)}) \right\}_{\mathbf{x}_{l,i,j},\ \mathbf{x}_{l,k,n} \in X_{l-1} \cap T}, \\
\mathbf{C}_{l,T;12} = \mathbf{C}_{l,T;21}^t &= \left\{ a_T(\phi_{ij}^{(l)}, \phi_{kn}^{(l-1)}) \right\}_{\mathbf{x}_{l,i,j} \in X_{w,l} \cap T,\ \mathbf{x}_{l,k,n} \in X_{l-1} \cap T}.
\end{aligned}$$

Note, that $\mathbf{C}_{l,T;1}$ is a 5×5 matrix, $\mathbf{C}_{l,T;12}$ is a 5×4 and $\mathbf{C}_{l,T;2} = \mathbf{C}_{l-1,T}$ is a 4×4 matrices for all T, such that $T \cap \partial\tau_0 = \varnothing$. We introduce the notations γ_l and $\gamma_{l,T}$ for the constant γ in the strengthened Cauchy inequality for a fixed l and for fixed l and T, respectively. Let $\lambda_{l,T}$ be the minimal nonzero eigenvalue of the generalized problem $\mathbf{S}_{l,T}\mathbf{v}_{l,T} = \lambda \mathbf{C}_{l,T;2}\mathbf{v}_{l,T}$, where $\mathbf{S}_{l,T} = \mathbf{C}_{l,T;2} - \mathbf{C}_{l,T;21}\mathbf{C}_{l,T;1}^{-1}\mathbf{C}_{l,T;12}$, then $\gamma_{l,T}^2 = 1 - \lambda_{l,T}$ (see Theorem 4.3 in [EV]). Simple application of Cauchy unequality results in $\gamma_l^2 \le \max\ \gamma_{l;T}^2$, where maximum is taken over all elements T, that belong to the mesh X_{l-1}.



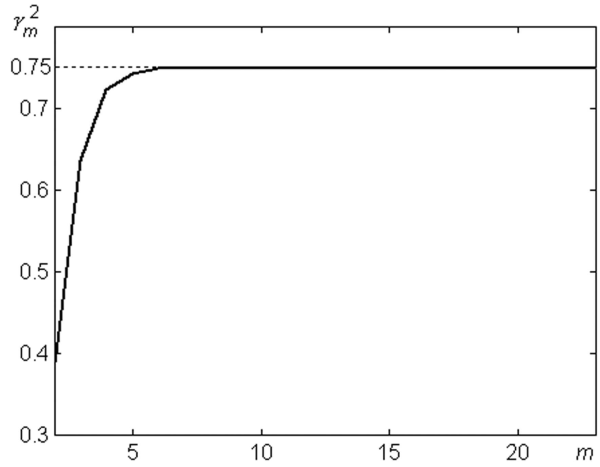


Fig. 1 Fig. 2

We calculated $\gamma^2_{l,T}$ for sufficiently large m and all m imbedded meshes numerically. Due to the symmetry of the meshes on τ_0, it is sufficient to calculate $\gamma^2_{m,T}$ only on the quarter of τ_0, say for $x_1 < 0$ and $x_2 < 0$. Numerical experiments indicate, that maximal value of $\gamma^2_{l,T}$ is obtained on the mesh X_{m-1}. Values of $\gamma^2_{m,T}$ on X_{m-1}, corresponding to $m = 6$, are presented on Fig. 1, and Fig. 2 contains the dependence of γ^2_m on m. The behavior of γ^2_m evidently shows, that γ^2_m is a monotonically growing function of m with $\lim \gamma^2_m = 0.75$. Therefore, $\gamma^2_m < 0.75$ for any fixed m.

Now we describe the preconditioner $\boldsymbol{\mathcal{B}}_{w,l}$ for pre- and postsmoothing iterations. Let us consider the quater $[-1, 0] \times [-1, 0]$ of the square τ_o. For each l-th mesh, we denote by $\Im_{l,\kappa}$ the mesh line, passing through the nodes with the indices satisfying $\max(i, j) = \kappa$, $\kappa = 1, 2.., 2^{l-1}$. For instance, for the m-th mesh the line $\Im_{l,\kappa}$ consists of the two parts: the segment $-1 \le x_1 \le \eta_\kappa$ of the line $x_2 = \eta_\kappa$ and the segment $-1 \le x_2 \le \eta_\kappa$ of the line $x_2 = \eta_\kappa$. On the rest three subsquares of τ_0 we arrange similar systems of lines by symmetry, so that indeed each line $\Im_{l,\kappa}$ for $\kappa = 1, 2, .., 2^{l-1}$ consists of four disjoint parts in four quarters of the coordinate system. For all l, the lines $\Im_{l,N_l}$ with $N_l = 2^{l-1}$ coincide with the intersection of the axes with τ_0.

We reorder the rows and columns, crossing the block $\mathbf{B}_{w,l}$, in such a way that we have sub-blocks $\mathbf{B}^{(\kappa)}_{w,l}$ on diagonal, each related to the nodes on the line $\Im_{l,\kappa}$. Having in mind the block structure, corresponding to the blocks $\mathbf{B}^{(\kappa)}_{w,l}$ on diagonal, the preconditioner $\boldsymbol{\mathcal{B}}_{w,l}$ is obtained from $\mathbf{B}_{w,l}$ by making zero all off-diagonal blocks. For the blocks on diagonal of $\boldsymbol{\mathcal{B}}_{w,l}$, we set $\boldsymbol{\mathcal{B}}^{(\kappa)}_{w,l} = \mathbf{B}^{(\kappa)}_{w,l}$. It is easy to see, that for even numbers κ, the blocks $\boldsymbol{\mathcal{B}}_{w,l}^{(\kappa)}$ are diagonal. In the blocks $\boldsymbol{\mathcal{B}}_{w,l}^{(\kappa)}$ with odd κ, only the neighbouring nodes on the line $\Im_{l,\kappa}$ are coupled, so that these blocks are tridiagonal. Indeed these blocks, which are defined for $\kappa < N_l$, contain four independent subblocks, each related to one connected part of the line $\Im_{l,\kappa}$.

It may be shown that $\boldsymbol{\mathcal{B}}_{w,l}$ and $\mathbf{B}_{w,l}$ are spectrally equivalent, *i.e.,* that $\varkappa_{m,l} := c_2/c_1$ is bounded independently of m and l, where c_1, c_2 appear in the inequalities (5.10) for the defined above $\boldsymbol{\mathcal{B}}_{w,l}$ and $\mathbf{B}_{w,l}$. Note that $\mathbf{B}_{w,l}$ is the result of assembling of the matrices $\mathbf{B}_{w,l,T} = \mathbf{C}_{l,T,2}$ over all nests T of the grid l. Let the respective preconditioning matrices, assembling of which renders $\boldsymbol{\mathcal{B}}_{w,l}$, be denoted $\boldsymbol{\mathcal{B}}_{w,l,T}$. We have $\varkappa_{m,l} \leq \sup_T \varkappa_{m,l,T} := \mathrm{cond}[\boldsymbol{\mathcal{B}}_{w,l,T}^{-1}\mathbf{B}_{w,l,T}]$. Then, we prove that $\sup_T \varkappa_{m,l,T}$ is always reached for one of square nests T. One diagonal of such a nest is always on the diagonal of τ_o. Another useful relationship is

$$\lim_{m\to\infty} \frac{x_{1,1,1} - x_{1,1,0}}{x_{1,1,2} - x_{1,1,1}} = \lim_{m\to\infty} \frac{x_{2,1,1} - x_{2,1,0}}{x_{2,1,2} - x_{2,1,1}} = \frac{1}{3}\,.$$

Let us consider the meshlines, deviding the square nest T of the mesh $l-1$ in four nests of the mesh l. From the above relationship, we conclude that in the coarsening procedure they may be shifted from the center of T not more, than 1/4-th of the size of T. From the pointed out facts, it follows that $\sup \varkappa_{m,l,T}$ is bounded. Now it is obvious, that the system of algebraic equations with the matrix $\mathbf{B}_{w,l}$ may be solved by PCG with the preconditioner $\boldsymbol{\mathcal{B}}_{w,l}$ for cN_l^2 arithmetic operations with c independent of m and l.

Also, we studied condition numbers $\varkappa_{m,l,T}$ numerically. For all initial finest grids X_m, the condition numbers $\varkappa_{m,l}$ increase with grid coarsening and reaches its peak value $\varkappa_{m,1}$ on X_1. Fig. 3 illustrates the dependence of $\varkappa_{m,1}$ on m for rather wide range of m.

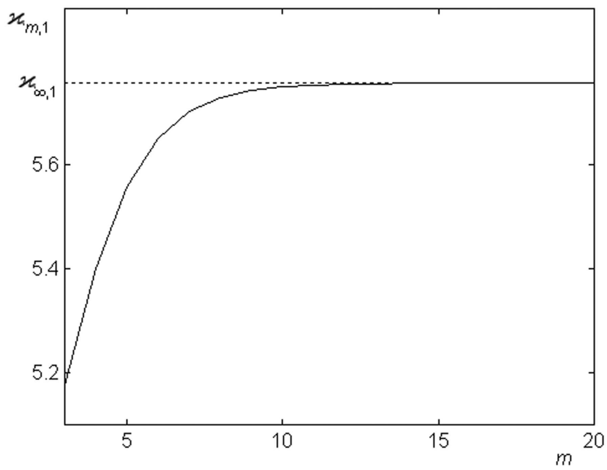


Fig.3

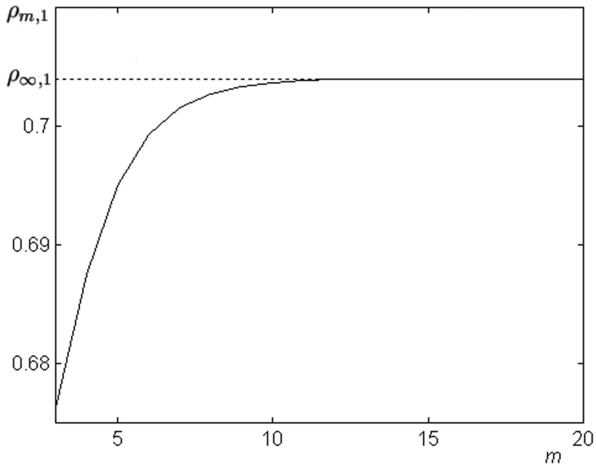


Fig.4

We can introduce the coarsest mesh X_1 , corresponding to the limit case $m = \infty$. In the square $[-1, 0] \times [-1, 0]$ this mesh consists of the lines $x_{1,1,0} = x_{2,1,0} = -1$, $x_{1,1,1} = x_{2,1,1} = -3/4$, $x_{1,1,2} = x_{2,1,2} = 0$. Denote by $\varkappa_{\infty,1}$ the generalized condition number of $\boldsymbol{\mathcal{B}}_{w,l}$ and $\mathbf{B}_{w,l}$ with respect to this mesh. We find, that $\varkappa_{\infty,1} \approx 5,7566$, see Fig. 3.

Convergence factors $\rho_{m,1}$ of the pre- or postsmoothing iterations and the upper bound for them $\rho_{\infty,1} \approx 0.7040$ were calculated in a similar way. Results are plotted on the Fig. 4.

With upper bounds for γ and $\rho_{m,1}$, we are able to estimate the convergence rate of the multigrid method and its major parameters. Let for simplicity, $\mu_l \equiv \mu$,

$$\parallel \mathbf{u}^{k+1,l} - \mathbf{u}^l \parallel \leq \delta_l \parallel \mathbf{u}^{k,l} - \mathbf{u}^l \parallel ,$$

$\mathbf{u}^{k+1,l}$ is obtained from the given approximation $\mathbf{u}^{k,l}$ by one multigrid iteration $\mathbf{u}^{k+1,l} := \mathbf{Mgm}(l, \mathbf{B}_l, \mathbf{F}, \mathbf{u}^{k,l})$ and $\mathbf{u}^l$ is the exact solution on the mesh X_l. Then (see Theorem 3.1 [S]), the numbers δ_l satisfy the recurrent sequence

$$\delta_1 = 0, \qquad \delta_l = \delta_{l-1}^{\mu} + (1 - \delta_{l-1}^{\mu})(\sigma + (1-\sigma)\gamma)^2, \quad l = 1, 2, \ldots ,$$

where σ is the convergence factor for ν pre- or postsmoothing iterations, *i.e.,* in our case $\sigma \leq \rho_{\infty,1}^{\nu}$. We can consider sufficiently large l irrespective of m, because the sequence is monotonically growing and the estimates, obtained for γ and $\rho_{m,1}$ are uniform in m. Therefore, having introduced the notation $\delta = \lim_{l\to\infty} \delta_l$, we need to find such μ and ν, that the equation

$$\delta = \delta^{\mu} + (1 - \delta^{\mu})(\rho_{\infty,1}^{\nu} + (1 - \rho_{\infty,1}^{\nu})\gamma)^2 \tag{6.20}$$

has a solution $\delta \in [0,1)$. Since $\inf_\nu \left\{(\rho_{\infty,1}^\nu + (1-\rho_{\infty,1}^\nu)\gamma)^2\right\}$ is obtained for $\nu = \infty$ and is equal to γ^2, by Lemma 3.2 of [S] we have $\delta \overline{\in} [0,1)$ for $\mu = 1,2,3$. Numerical study of (6.20) for $\mu = 4$ shows that there is no $\delta \in [0,1)$ even for sufficiently large ν as well. However, for $\mu \geq 5$ convergence of the multigrid iterations may be guaranteed. Dependence of δ on ν for several μ is plotted on Fig. 5. A more precise analysis of the convergence of the multigrid method is still possible, in particular, on the basis of the results of [Pf].

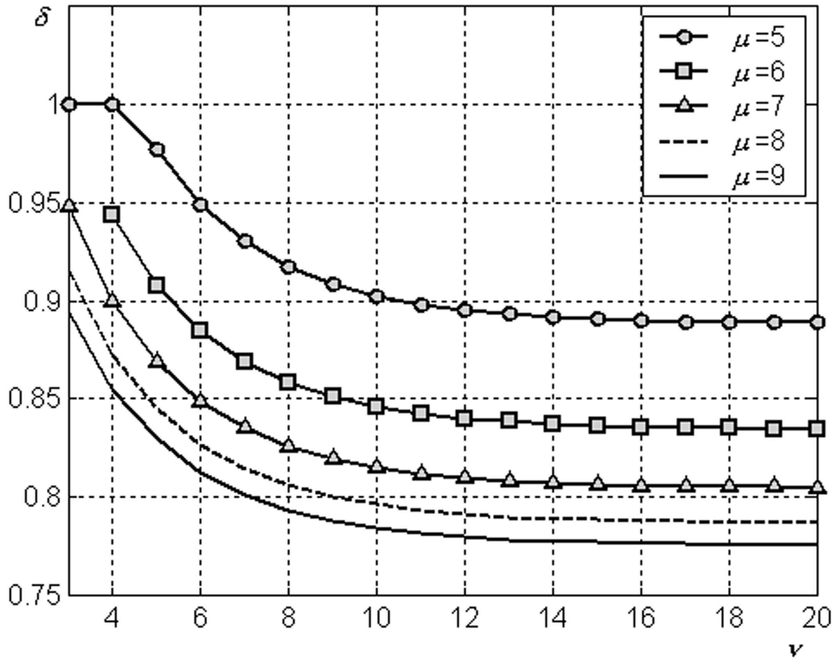


Fig.5

References:

[AV] O. Axelsson and P.S. Vassilevskii. Algebraic multilevel preconditioning methods ii. SIAM J. Numer. Anal., **27**(6) (1990), 1569–1590.

[B] I. Babuška, A. Craig, J. Mandel, and J. Pitkäranta. Efficient preconditioning for the p–version finite element method in two dimensions. SIAM J. Numer. Anal., **28** (1991), 624–661.

[BDM1] C. Bernardi, M. Dauge, Y. Maday. Relévements de traces presérvant les polinômes. C. R. Acad. Sci. Paris Sér. I Math., **315** (1992), 333–338 (in French).

[BDM2] C. Bernardi, M. Dauge, Y. Maday. Polynomials in weighted Sobolev spaces: Basics and trace liftings. Publications du Laboratoire d'Analyse Numerique, R92039. Universite Pierre et Marie Curie, Paris, 1993.

[BM] C. Bernardi, Y. Maday. Approximations spectrales de problemes aux limites elliptiques. Math. Appl., **10**, Springer-Verlag, France, Paris, 1992.

[Be] S. Beuchler. A preconditioner for solving the inner problem for the p-version of the FEM, Part II – algebraic multigrid proof. Technische Universität Chemnitz, Preprint SFB393/01 -07 (2001), 1–56.

[Be1] S. Beuchler. MTS-BPX-preconditioner for the p-Version of the fem. Technische Universität Chemnitz, Preprint SFB393/01-16 (2001), 1–7.

[Can] C. Canuto. Stabilization of spectral methods by finite element bubble functions. Comput. Methods Appl. Mech. Engrg., **116** (1994), 13–26.

[Cas] M. Casarin. Quasi-optimal Scharz methods for the conforming spectral element discretizations. SIAM J. Numer. Anal., **34**, N6 (1997), 2482–2502.

[GO] M. Griebel and P. Oswald. Tensor product type subspace splittings and multilevel iterative methods for anisothropic problems. Advances Comput. Math. **4** (1995), 171–201.

[IK] S.A. Ivanov and V.G. Korneev. Preconditioning in the domain decomposition methods for the p-version with the hierarchical bases. Matematicheskoe modelirovanie, **8**, 1996, 63–73.

[KP] S. Kim and S. Parter. Preconditioning Chebyshev spectral collocation by finite-difference operators. SIAM J. Numer. Anal., **34**, N3 (1997), 939–958.

[K1] V. Korneev. Almost optimal solving procedure for Dirichlet problems on subdomains of decomposition in hierarchical hp-version. Differentsialnye uravnenia, **37**, N7, 2001, 958-969.

[K2] V.G. Korneev. Local Dirichlet problems on subdomains of decomposition in hp discretizations, and optimal algorithms for their solution. Matematicheskoie modelirovanie, **14**, N5, 2202, 61–94.

[KJ1] V.G. Korneev and S. Jensen. Preconditioning of the p-version of the finite element method. Comput. Methods Appl. Mech. Engrg., **150** (1997), 215–238.

[KJ2] V.G. Korneev and S. Jensen. Domain decomposition preconditioning in the hierarchical p-version of the finite element method. Applied Numerical Mathematics, **29** (1999), 479–518.

[KJ3] V. Korneev and S. Jensen. Efficient preconditioning by domain decomposition method for hp-version with hierarchical basis, I. Izvestia VUZ'ov, Math., N 5 (444), 1999, 37-56. (In Russian)

[KJ4] V. Korneev and S. Jensen. Efficient preconditioning by domain decomposition method for hp-version with hierarchical basis, II. Izvestia VUZ'ov, Math., N 11 (450), 1999, 24-40.

[KFOF] V. Korneev, J.E. Flaherty, T. Oden and J. Fish. Hp-Version Additive Schwarz Algorithms on Triangular Meshes. Matematicheskoie modelirovanie, **14**, N2 (2202), 61–94.

[N] S.V. Nepomnyaschikh. Method of splitting into subspaces for solving elliptic boundary value problems in complex–form domains. Russian (cont. Sov.) J. Numer. Anal. & Math. Model. **6**:2 (1991) 151–168.

[MP] V. Maz'ya and S. Poborchi. Differentiable functions on bad domains. Singapore-New Jersy- London-Hong Kong: World Scientific, 1998, 504 p.

[P] S. Parter. Preconditioning Legendre spectral collocation methods for elliptic problems II: finite element operators. SIAM J. Numer. Anal., **39**, N1 (2001), 348–362.

[Pf] Ch. Pflaum. Robust confergence of multilevel algorithms for convection-difusion equations. SIAM J. Num. Anal., **37(2)** (2000), 443–469.

[S] N. Schieweck. A multigrid convergence proof by a strengthened Caushy inequality for symmetric elliptic boundary value problems. In: Second multigrid seminar, Garzau 1985 , G. Telschow, ed., 49–62.