

BALANCING DOMAIN DECOMPOSITION BY CONSTRAINTS AND PERTURBATION*

SANTIAGO BADIA[†] AND HIEU NGUYEN[‡]

Abstract. In this paper, we formulate and analyze a perturbed formulation of the balancing domain decomposition by constraints (BDDC) method. We prove that the perturbed BDDC has the same polylogarithmic bound for the condition number as the standard formulation. Two types of properly scaled zero-order perturbations are considered: one uses a mass matrix, and the other uses a Robin-type boundary condition, i.e, a mass matrix on the interface. With perturbation, the well-posedness of the local Neumann problems and the global coarse problem is automatically guaranteed, and coarse degrees of freedom can be defined only for convergence purposes but not well-posedness. This allows a much simpler implementation as no complicated corner selection algorithm is needed. Minimal coarse spaces using only face or edge constraints can also be considered. They are very useful in extreme scale calculations where the coarse problem is usually the bottleneck that can jeopardize scalability. The perturbation also adds extra robustness as the perturbed formulation works even when the constraints fail to eliminate a small number of subdomain rigid body modes from the standard BDDC space. This is extremely important when solving problems on unstructured meshes partitioned by automatic graph partitioners since arbitrary disconnected subdomains are possible. Numerical results are provided to support the theoretical findings.

Key words. BDDC, preconditioner, coarse space, parallel solver, scalability

AMS subject classifications. 65N55, 65N22, 65F08

DOI. 10.1137/15M1045648

1. Introduction. The development of highly scalable linear solvers for the solution of large scale linear systems arising from the finite element (FE) discretization of second-order elliptic problems on distributed-memory machines is of great importance in many applications. In this work, we consider nonoverlapping domain decomposition (DD) methods [49], which take advantage of the partition of the FE mesh into submeshes to define effective preconditioners that can exploit large levels of concurrency. In particular, we focus on a variant of the balancing DD by constraints (BDDC) method. The BDDC method was first introduced in 2003 by Dohrmann [18]. It can be regarded as an improved version of the balancing domain decomposition (BDD) method by Mandel [43]. It also has a very close connection with the dual primal finite element tearing and interconnecting (FETI-DP) method [25, 24]. In fact, the eigenvalues of the preconditioned systems in the two approaches are almost identical [44, 41, 15]. The BDDC method is particularly well suited for extreme scale simulations, since it allows for a very aggressive coarsening, the computations at different levels can be computed in parallel, the subdomain problems can be solved inexactly

*Received by the editors October 27, 2015; accepted for publication (in revised form) October 19, 2016; published electronically November 29, 2016.

<http://www.siam.org/journals/sinum/54-6/M104564.html>

Funding: The work of the authors was funded by the European Research Council under the FP7 Program Ideas through the Starting Grant 258443, COMFUS: Computational Methods for Fusion Technology, and the FP7 NUMEXAS project under grant agreement 611636. The work of the first author was supported by the Catalan Government through the ICREA Acadèmia Research Program.

[†]Universitat Politècnica de Catalunya, 08034 Barcelona, Spain, and CIMNE – Centre Internacional de Mètodes Numèrics en Enginyeria, Parc Mediterrani de la Tecnologia, UPC, 08860 Castelldefels, Spain (sbadia@cimne.upc.edu).

[‡]Institute of Research and Development, Duy Tan University, 3 Quang Trung, Danang, Vietnam, and CIMNE – Centre Internacional de Mètodes Numèrics en Enginyeria, Parc Mediterrani de la Tecnologia, UPC, 08860 Castelldefels, Spain (nguyentrunghieu14@dtu.edu.vn).

[19, 42] by, e.g., one AMG cycle, and it can straightforwardly be extended to multiple levels [50, 45]. All of these properties have been carefully exploited in the series of papers [6, 7, 8, 9], where an extremely scalable implementation of these algorithms has been proposed, leading to excellent weak scalability on nearly half a million cores in its multilevel version (see also [29, 30] for weak scalability at extreme scales of the FETI-DP method).

The BDDC method is based on a classification of nodes on the subdomain interface into corners, and members of edges and faces, based on geometrical arguments that can be used for any input mesh [49]. It is formulated on the partially subassembled FE space in which continuity among subdomains is only weakly enforced on some quantities, e.g., the value on subdomain corners and mean values on subdomain edges and subdomain faces. In two dimensions, enforcing continuity on corners only is enough to have a weakly scalable method; i.e., the number of iterations depends only on the local problem size but not on the number of processors being used or the global problem size. In three dimensions, continuity of mean values on edges and/or faces must also be enforced to have fast convergence [25, 38].

The role of the corner constraints in the BDDC method is not only to attain scalability. They must also ensure that the partially subassembled local problems and the global coarse problem are well-posed. For structured meshes and regular partitions, algorithms to find such constraints are simple, but it would become much more complicated when considering unstructured meshes of complex three-dimensional geometries and partitions resulting from parallel mesh partitioners, e.g., ParMETIS [27] or PT-Scotch [16]. In order to have well-posedness, different corner selection algorithms have been considered [18, 40, 37, 48]. The idea behind these algorithms is to introduce enough corner constraints such that the partially subassembled problems are provably well-posed even with corner constraints only.

The reason why edge and/or face constraints are not usually considered for the same purpose is twofold. First, it is harder to check how edge or face constraints fix rigid-body modes in elasticity problems. Second, if corner constraints are sufficient to have well-posedness, the partially subassembled systems, which are indefinite when edge and/or face constraints are used, can be reduced to positive definite systems by eliminating variables associated with corners [18]. Consequently, we can use more efficient and more robust sparse methods which only work for positive definite systems. An alternative allowing the use of positive definite sparse solvers and avoiding corner constraints is to consider a change of basis that explicitly enforces edge and face constraints [41, 37]. However, this approach is not straightforward to implement and notably changes the sparsity pattern of the partially subassembled matrices.

The corner selection algorithms also have their own drawbacks. Based on our experience, the implementation of this type of algorithm is an involved and time-consuming task which depends on the physical problem to be solved and also the type of FE formulation being used. For example, the implementation of the corner selection algorithm from [48] in the scientific software FEMPAR [4] has more than 1,500 lines of code. Furthermore, the situation becomes far more complicated when subdomains are disconnected, or connected only by corners or edges.¹ As far as we know, these situations have not been fully considered when formulating the corner selection algorithms [40, 18, 48], thus leading to the breakdown of the BDDC method. It is certainly possible to deal with disconnected subdomains, but it will increase the

¹As an example, one can think about elasticity in three dimensions with two subdomains sharing, e.g., one edge. The rigid body modes for the pair of subdomains are nontrivial.

complexity of the implementation of the BDDC method; for the solution of the same problem in the BNN preconditioner, see [5]. Unfortunately, ParMETIS [27] and PT-Scotch [16] parallel mesh partitioners cannot guarantee connected subdomains, and we are not aware of any parallel mesh partitioner that can do it. The situation is even more dramatic when using adaptive mesh refinement and space-filling curves for efficient load balancing, e.g., exploiting Morton order [1].

In the authors' opinion, finding corner constraints to guarantee the well-posedness is one of the main issues that reduces the simplicity of the BDDC method. In this work, we consider a different and simpler approach based on a perturbed formulation of the method. Instead of using corner constraints to guarantee well-posedness, we consider perturbed partially subassembled systems which are automatically well-posed, regardless of the choice of the constraints. Two types of perturbations are considered—a zero-order term, i.e., a mass matrix, and the use of Robin-type boundary conditions, i.e., a perturbation by interface mass matrix. Certainly, we still need to have a number of constraints to achieve scalability, but these can be simply chosen based on geometrical properties.

Perturbation has been used to make local problems well-posed in other nonoverlapping methods [17, 22]. However, in these previous works, the methods are either without a coarse space or with a geometrical coarse space. Therefore, using perturbation does not affect the coarse space and the convergence analysis of the methods. This does not hold for the BDDC methods where the coarse space depends on the formulation of the local problems. In fact, one can check that the scaling coefficients in [17, 22] are completely different from the ones proposed in this work which are required for scalability purposes.

We show that the perturbed BDDC method has the same polylogarithmic bound for the condition number and thus is scalable. In addition, when the chosen constraints are not enough and the BDDC space contains a small subspace of rigid body modes, we prove that most of the eigenvalues of the preconditioned system, except for a few large isolated ones, are bounded by the usual bounds of the standard BDDC method. In other words, the perturbed BDDC method is still scalable even though the number of iterations will be higher compared to the case when all rigid body modes are fixed.

Further, with the perturbed formulation, the well-posedness is guaranteed and sparse direct methods for positive definite systems can be used without the need of corner selection mechanisms, corner constraints enforcement, or a change of basis when there are only edge or face constraints.

Another very important benefit of using the perturbed formulation is the fact that we can reduce the size of the coarse problem, allowing even more aggressive coarsening strategies than the ones that can be used with the standard BDDC method. For three-dimensional problems, corners (including those introduced by the corner selection algorithms) do not significantly contribute to the convergence. With the perturbed BDDC method we can consider much smaller coarse problems only involving faces and/or edges. This point is particularly important when dealing with unstructured meshes and partitions generated by mesh partitioners due to the proliferation of corners. Since the coarse problem is the bottleneck that can destroy scalability, these strategies are better suited for large scale simulations.

The rest of the paper is organized as follows. In section 1, we introduce the model problems and the interface problem in nonoverlapping methods. In section 2, the formulation of the perturbed BDDC preconditioner is presented. Different types of perturbation are discussed in section 3. The proposed preconditioner is analyzed in section 4. In section 5, numerical experiments are provided to verify the theoretical

findings. We end by drawing some conclusions in section 6.

1.1. Model problems. Let Ω be a bounded domain with a polygonal boundary in \mathbb{R}^n , $n = 2, 3$. We consider two model problems.

Poisson's equation. For simplicity of exposition, we consider Poisson's equation with homogeneous Dirichlet conditions. Its weak formulation reads as follows: find $u(x) \in H_0^1(\Omega)$ such that

$$(1.1) \quad a(u, v) = \langle F, v \rangle \quad \forall v(x) \in H_0^1(\Omega),$$

where

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad \langle F, v \rangle = \int_{\Omega} f(x) v(x) \, dx.$$

The case with a general boundary condition only involves a modification of the right-hand side, and our theoretical results still apply. In fact, we consider nonhomogeneous boundary conditions in the numerical experiment section.

Linear elasticity. Let $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, and assume that $\partial\Omega_D$ is clamped with the homogeneous Dirichlet boundary condition, and $\partial\Omega_N$ is subjected to a surface force \mathbf{g} , i.e., a natural boundary condition. There is also a body force \mathbf{f} . Let $\mathbf{H}_0^1(\Omega, \partial\Omega_D)^n = \{\mathbf{v} \in H^1(\Omega)^n : \mathbf{v}|_{\partial\Omega_D} = 0\}$. We consider the problem of finding $\mathbf{u} \in \mathbf{H}_0^1(\Omega, \partial\Omega_D)^n$ such that

$$(1.2) \quad a(\mathbf{u}, \mathbf{v}) = \langle \mathbf{F}, \mathbf{v} \rangle \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega, \partial\Omega_D)^n,$$

where

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \int_{\Omega} 2\mu \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}) \, dx + \int_{\Omega} \lambda \operatorname{div} \mathbf{u} \operatorname{div} \mathbf{v} \, dx, \\ \boldsymbol{\epsilon}(\mathbf{u}) : \boldsymbol{\epsilon}(\mathbf{v}) &= \sum_{i,j=1}^3 \epsilon_{ij}(\mathbf{u}) \epsilon_{ij}(\mathbf{v}), \quad \epsilon_{ij}(\mathbf{u}) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \\ \langle \mathbf{F}, \mathbf{v} \rangle &= \int_{\Omega} \mathbf{f}^T \mathbf{v} \, dx + \int_{\partial\Omega_N} \mathbf{g}^T \mathbf{v} \, ds. \end{aligned}$$

The presentation in this work covers both the Poisson and linear elasticity problems. In general, proofs are given for the Poisson problem, and some comments are provided about how to extend it to linear elasticity when needed.

Let \mathcal{T}_h be a shape-regular triangulation of size h of Ω and $V_h \subset H_0^1(\Omega)$ be the corresponding \mathcal{P}_1 space, i.e., the space of continuous piecewise linear functions on \mathcal{T}_h that vanish on $\partial\Omega$. Then the discrete problem for (1.1) is to find $u_h \in V_h$ such that

$$(1.3) \quad a(u_h, v) = \langle F, v \rangle \quad \forall v \in V_h.$$

Let us also consider a partition of Ω into subdomains $\Omega = \cup_{j=1}^J \Omega_j$ with the inter-subdomain interface $\Gamma = \cup_{j=1}^J \partial\Omega_j \setminus \partial\Omega$. We denote by H the subdomain characteristic size. We further assume that the subdomain partition is obtained by aggregation of elements in \mathcal{T}_h .

We represent by $V_h(\Gamma)$ the *interface space* of piecewise discrete harmonic functions which are discrete harmonic on each subdomain. The *global Schur complement operator* $S_h : V_h(\Gamma) \rightarrow V_h(\Gamma)'$, where $V_h(\Gamma)'$ is the dual space of $V_h(\Gamma)$, can be defined by

$$(1.4) \quad \langle S_h v_1, v_2 \rangle = a(v_1, v_2) \quad \forall v_1, v_2 \in V_h(\Gamma).$$

Here $\langle \cdot, \cdot \rangle$ denotes the canonical duality pairing, i.e., $\langle \gamma, v \rangle = \gamma(v)$.

Using static condensation, namely the process of eliminating component functions in V_h that vanish on Γ to reduce (1.3) to a problem posed on the interface space $V_h(\Gamma)$ (see [14, section 7.5]), we have

$$(1.5) \quad S_h \bar{u}_h = f_h, \quad \bar{u}_h \in V_h(\Gamma), \quad f_h = F|_{V_h(\Gamma)}.$$

Our goal is to formulate an efficient preconditioner for S_h .

2. Preconditioner formulation.

2.1. Constraints. The constraints in BDDC usually rely on a topological classification of the nodes on the interface as subdomain corners, nodes of subdomain edges or subdomain faces. In this subsection, we will give concrete definitions of these geometrical objects.

We denote by Γ_h the discretization of Γ associated with \mathcal{T}_h . Let $\mathcal{N}(p)$ be the set of indices of subdomains that share node $p \in \Gamma_h$, i.e., $\mathcal{N}(p) = \{j : p \in \Gamma_h^j = \Gamma_h \cap \partial\Omega_j\}$. Then the nodes in Γ_h can be classified into objects, where each object \mathcal{G} is a maximal subset of nodes with the same set of indices, i.e., $\mathcal{N}(p) = \mathcal{N}(q)$ for any $p, q \in \mathcal{G}$. We denote $|\cdot|$ the cardinality of a set and $\mathcal{N}(\mathcal{G})$ the set of subdomain indices associated with \mathcal{G} , i.e., $\mathcal{N}(\mathcal{G}) := \mathcal{N}(p)$, $p \in \mathcal{G}$. Then, in three-dimensional space, \mathcal{G} is classified as a *face* if $|\mathcal{N}(\mathcal{G})| = 2$ and $|\mathcal{G}| > 1$, an *edge* if $|\mathcal{N}(\mathcal{G})| > 2$ and $|\mathcal{G}| > 1$, and a *corner* if $|\mathcal{G}| = 1$. In two-dimensional problems, \mathcal{G} is classified as an *edge* if $|\mathcal{N}(\mathcal{G})| = 2$ and $|\mathcal{G}| > 1$ and a *corner* if $|\mathcal{G}| = 1$. A similar classification was introduced in [37, 32, 48].

A constraint in BDDC can be the value at a subdomain corner or an average value on a subdomain edge or a subdomain face. As we will see in the following subsection, the constraints are no longer responsible for the well-posedness of the local Neumann problems and the global coarse problem in the proposed perturbed formulation. Therefore, they can be chosen based only on their potential to improve the convergence of the method. For simplicity, in this work we propose choosing \mathcal{C} , the set of constraints, by deciding the types of constraints we want to use. In other words, \mathcal{C} contains all the corner constraints, or all the edge constraints, or all the face constraints, or combinations of these options; we do not consider only corner constraints in three dimensions as this choice is known to be noncompetitive [24, 25].

2.2. Perturbed BDDC preconditioner. Considering $\tilde{a}(\cdot, \cdot)$ a perturbed bilinear form of $a(\cdot, \cdot)$, we assume that $\tilde{a}(\cdot, \cdot)$ satisfies the following assumptions.

Assumption 2.1. There exist two constants C_L and C_U which are independent of the size of the domain (D), the size of the subdomains (H), and the number of the subdomains (J) such that

$$(2.1) \quad C_L a(v, v) \leq \tilde{a}(v, v) \leq C_U a(v, v) \quad \forall v \in H_0^1(\Omega).$$

Assumption 2.2. The bilinear form $\tilde{a}(\cdot, \cdot)$ can be written as

$$\tilde{a}(u, v) = \sum_{j=1}^J \tilde{a}_j(u_j, v_j) \quad \forall u, v \in H_0^1(\Omega), \quad u_j = u|_{\Omega_j}, v_j = v|_{\Omega_j},$$

where each $\tilde{a}_j(\cdot, \cdot)$ is a positive definite bilinear form on $H_0^1(\Omega)|_{\Omega_j}$.

Assumption 2.3. The local perturbed bilinear form $\tilde{a}_j(\cdot, \cdot)$ introduced in Assumption 2.2 satisfies

$$(2.2) \quad C_\ell a_j(v_j, v_j) \leq \tilde{a}_j(v_j, v_j) \quad \forall v_j \in H_0^1(\Omega)|_{\Omega_j},$$

where $a_j(u, v) = \int_{\Omega_j} \nabla u \cdot \nabla v \, dx$ and $C_\ell > 0$ does not depend on the size of the domain (D) , the size of the subdomains (H) , or the number of the subdomains (J) .

In the next step, we will present the perturbed formulation for the BDDC method. We adopt notation similar to that in [14, section 7.8].

Let $\tilde{V}_h(\Gamma)$ be the space of piecewise discrete harmonic functions with respect to (w.r.t.) the bilinear form $\tilde{a}(\cdot, \cdot)$. We define the *perturbed global Schur complement operator* $\tilde{S}_h : \tilde{V}_h(\Gamma) \rightarrow \tilde{V}_h(\Gamma)'$ as follows:

$$(2.3) \quad \langle \tilde{S}_h v_1, v_2 \rangle = \tilde{a}(v_1, v_2) \quad \forall v_1, v_2 \in \tilde{V}_h(\Gamma).$$

Let $V_j := V_h|_{\Omega_j}$ and $V_h(\Omega_j)$ be the subspace of V_j whose members vanish on $\partial\Omega_j$. Also let $\tilde{\mathcal{H}}_j = \tilde{V}_h(\Gamma)|_{\Omega_j}$ be the space of local $\tilde{a}_j(\cdot, \cdot)$ -discrete harmonic functions. Then, it follows that $\tilde{\mathcal{H}}_j \subset V_j$ and $\tilde{\mathcal{H}}_j = \{v_j \in V_j : \tilde{a}_j(u_j, v_j) = 0 \, \forall u_j \in V_h(\Omega_j)\}$.

Now we can introduce the perturbed BDDC space associated with $\tilde{a}(\cdot, \cdot)$:

$$(2.4) \quad \tilde{\mathcal{H}}_c = \{\tilde{v} \in L^2(\Omega) : \tilde{v}_j = \tilde{v}|_{\Omega_j} \in \tilde{\mathcal{H}}_j, \tilde{v} \text{ has continuous values for constraints in } \mathcal{C}\}.$$

Clearly, $\tilde{V}_h(\Gamma) \subset \tilde{\mathcal{H}}_c$.

The symmetric positive definite operator $\tilde{S}_c : \tilde{\mathcal{H}}_c \rightarrow \tilde{\mathcal{H}}_c'$ is then defined by

$$(2.5) \quad \langle \tilde{S}_c u, v \rangle = \sum_{j=1}^J \tilde{a}_j(u_j, v_j) \quad \forall u, v \in \tilde{\mathcal{H}}_c, u_j = u|_{\Omega_j}, v_j = v|_{\Omega_j}.$$

We note that $\langle \tilde{S}_c u, v \rangle = \langle \tilde{S}_h u, v \rangle$ for $u, v \in \tilde{V}_h(\Gamma) \subset \tilde{\mathcal{H}}_c$.

Let $\tilde{E}_j : \tilde{\mathcal{H}}_j \rightarrow L^2(\Omega)$ be the trivial extension (by zero) operator. We can define $\tilde{\mathcal{H}}_j^\circ := \{v \in \tilde{\mathcal{H}}_j : \tilde{E}_j v \in \tilde{\mathcal{H}}_c\}$, the space of local $\tilde{a}_j(\cdot, \cdot)$ -discrete harmonic functions whose interface constraints in \mathcal{C} equal zero. Each $\tilde{\mathcal{H}}_j^\circ$ is then equipped with the symmetric positive definite operator $\tilde{S}_j : \tilde{\mathcal{H}}_j^\circ \rightarrow (\tilde{\mathcal{H}}_j^\circ)'$ defined by

$$(2.6) \quad \langle \tilde{S}_j u, v \rangle = \tilde{a}_j(u, v) \quad \forall u, v \in \tilde{\mathcal{H}}_j^\circ.$$

Now we define the coarse space of the perturbed BDDC preconditioner,

$$(2.7) \quad \tilde{\mathcal{H}}_0 = \{v \in \tilde{\mathcal{H}}_c : \tilde{a}_j(u_j, v_j) = 0 \, \forall u_j \in \tilde{\mathcal{H}}_j^\circ, v_j = v|_{\Omega_j}, 1 \leq j \leq J\},$$

and the symmetric positive definite operator $\tilde{S}_0 : \tilde{\mathcal{H}}_0 \rightarrow \tilde{\mathcal{H}}_0'$ as follows:

$$(2.8) \quad \langle \tilde{S}_0 u, v \rangle = \langle \tilde{S}_c u, v \rangle \quad \forall u, v \in \tilde{\mathcal{H}}_0.$$

Obviously, the space $\tilde{\mathcal{H}}_c$ can be written as

$$(2.9) \quad \tilde{\mathcal{H}}_c = \tilde{\mathcal{H}}_0 \oplus_{\tilde{S}_c} \tilde{\mathcal{H}}^\circ, \quad \text{where } \tilde{\mathcal{H}}^\circ|_{\Omega_j} = \tilde{\mathcal{H}}_j^\circ.$$

Remark 2.4. The member functions of $\tilde{\mathcal{H}}_0$ satisfy a constrained minimization property; namely, for any $v_0 \in \tilde{\mathcal{H}}_0$ we have

$$\tilde{a}_j(v_0, v_0) = \min_{v_j \in V_j, v_j = v_0 \text{ "at" } \mathcal{C}|_{\Gamma_h^j}} \tilde{a}_j(v_j, v_j).$$

Now in order to connect the spaces $\tilde{\mathcal{H}}_0$ and $\tilde{\mathcal{H}}_j^\circ$ to $\tilde{V}_h(\Gamma)$, we define the *averaging operator* $\tilde{P}_\Gamma : \tilde{\mathcal{H}}_c \rightarrow \tilde{V}_h(\Gamma)$ as follows:

$$(2.10) \quad (\tilde{P}_\Gamma v)(p) = \frac{1}{|\mathcal{N}(p)|} \sum_{j \in \mathcal{N}(p)} v_j(p) \quad \forall v \in \tilde{\mathcal{H}}_c, v_j = v|_{\Omega_j}, p \in \Gamma_h.$$

As we want to formulate a preconditioner for S_h , not \tilde{S}_h , we need a connection from $\tilde{V}_h(\Gamma)$ to $V_h(\Gamma)$. Let $Q_\Gamma : \tilde{V}_h(\Gamma) \rightarrow V_h(\Gamma)$ be the operator defined by

$$(2.11) \quad (Q_\Gamma \tilde{v})(p) = \tilde{v}(p) \quad \forall \tilde{v} \in \tilde{V}_h(\Gamma), p \in \Gamma_h.$$

The operator Q_Γ in (2.11) is well defined because functions in $V_h(\Gamma)$ and $\tilde{V}_h(\Gamma)$ are completely determined by their nodal values on the interface. Here, we emphasize that even though \tilde{v} and $Q_\Gamma \tilde{v}$ are identical on the interface Γ_h , they can be different. More specifically, \tilde{v} is the harmonic extension of $\tilde{v}|_{\Gamma_h}$ w.r.t. the bilinear form $\tilde{a}(\cdot, \cdot)$, while $Q_\Gamma \tilde{v}$ is the harmonic extension of $\tilde{v}|_{\Gamma_h}$ w.r.t. the bilinear form $a(\cdot, \cdot)$.

Finally, we are in a position to present our perturbed BDDC preconditioner $\tilde{B}_{\text{BDDC}} : V_h'(\Gamma) \rightarrow V_h(\Gamma)$ for S_h . It is defined as follows:

$$(2.12) \quad \begin{aligned} \tilde{B}_{\text{BDDC}} &= (Q_\Gamma \tilde{P}_\Gamma \tilde{E}_0) \tilde{S}_0^{-1} (Q_\Gamma \tilde{P}_\Gamma \tilde{E}_0)^t + \sum_{j=1}^N (Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j) \tilde{S}_j^{-1} (Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j)^t \\ &= \sum_{j=0}^J P_j \tilde{S}_j^{-1} (P_j)^t, \quad \text{where } P_j = Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j. \end{aligned}$$

Here, we use \tilde{E}_0 to denote the natural injection of $\tilde{\mathcal{H}}_0$ into $\tilde{\mathcal{H}}_c$.

Remark 2.5. We note that the perturbed BDDC preconditioner is formulated with perturbed operators in all of the local subdomain solves as well as the global coarse solve. For perturbed bilinear forms satisfying Assumptions 2.1 and 2.2, the invertibility of these operators is guaranteed regardless of the choice of the constraints.

Remark 2.6. We emphasize that \tilde{B}_{BDDC} is not the standard BDDC preconditioner formulated for the perturbed problem (the PDE associated with the perturbed bilinear form). It is a hybrid preconditioner for the original Schur operator S_h and not the Schur operator of the perturbed problem. Reducing the problem to the interface requires solving local Dirichlet problems in the original form, whereas forming the perturbed preconditioner requires solving local Neumann problems and a global coarse problem in the perturbed form.

3. Choices of perturbed bilinear form. In this section, we present two practical choices for the perturbed bilinear form $\tilde{a}(\cdot, \cdot)$ and its subdomain components.

3.1. Perturbation with full mass matrix. The bilinear forms $\tilde{a}(\cdot, \cdot)$ and $\tilde{a}_j(\cdot, \cdot)$ are chosen as follows.

$$(3.1) \quad \tilde{a}(u, v) = \sum_{j=1}^J \tilde{a}_j(u, v), \quad \tilde{a}_j(u, v) = a_j(u, v) + \frac{1}{D^2} \int_{\Omega_j} uv \, dx,$$

where D is the size (diameter) of Ω . We emphasize that the scaling factor used here, i.e., $1/D^2$, is different from the scaling factor $1/H^2$ used in [21].

LEMMA 3.1. *The bilinear forms $\tilde{a}(\cdot, \cdot)$ and $\tilde{a}_j(\cdot, \cdot)$ defined in (3.1) satisfy Assumption 2.1 with $C_L = 1$ and $C_U = 1 + C_\Omega$, Assumption 2.2, and Assumption 2.3 with $C_\ell = 1$. Here, the constant C_Ω depends only on the shape of Ω but not on its size.*

Proof. The first inequalities in Assumptions 2.1, 2.2, and 2.3 are straightforward. The second inequality in Assumption 2.1 follows from a Friedrichs inequality and a scaling argument. Here, C_Ω is the constant in a Friedrichs inequality, and it depends only on the shape of Ω but not on its size (see, e.g, [49, Corollary A.15]). In case Ω is convex, according to [12], after the scaling argument, $C_\Omega = 1/\pi$. \square

3.2. Robin perturbation. The bilinear form $\tilde{a}(\cdot, \cdot)$ is defined by

$$(3.2) \quad \tilde{a}(u, v) = \sum_{j=1}^J \tilde{a}_j(u, v), \quad \tilde{a}_j(u, v) = a_j(u, v) + \frac{H_j^{n-1}}{D^n} \int_{\Gamma_j} uv \, ds,$$

where D is the size of the domain Ω , H_j is the size of the subdomain Ω_j , and n is the dimension of the physical space Ω . We call this a Robin perturbation since the local Neumann problems, in this case, can be posed with Robin boundary condition $(H_j^{n-1}/D^n)u + \partial u/\partial n_j = 0$ on $\partial\Omega_j \setminus \partial\Omega$, where n_j is the outward normal vector of $\partial\Omega_j$.

LEMMA 3.2. *The bilinear forms $\tilde{a}(\cdot, \cdot)$ and $\tilde{a}_j(\cdot, \cdot)$ defined as in (3.2) satisfy Assumption 2.1 with $C_L = 1$ and $C_U = 1 + \hat{C}_\Omega$, Assumption 2.2, and Assumption 2.3 with $C_\ell = 1$. Here, the constant \hat{C}_Ω does not depend on the size of the domain or the size of the subdomains.*

Proof. We only need to focus on the second inequality in Assumption 2.1, as other inequalities are trivial.

For subdomain Ω_j , applying Lemma A.17 in [49] with $L = 1$, $f_1(v) = \|v\|_{L^2(\partial\Omega)}$, and using the trace theorem, for any $v \in H_0^1(\Omega)$, we find that

$$\|v\|_{L^2(\partial\Omega_j)}^2 \leq C_1 |v|_{H^{1/2}(\partial\Omega_j)}^2 + C_2 \|v\|_{L^2(\Omega_j)}^2 \leq C_1 |v|_{H^1(\Omega_j)}^2 + C_2 \|v\|_{L^2(\Omega_j)}^2.$$

By a scaling argument and the fact that the partition is quasi-uniform, there exist constants \hat{C}_1, \hat{C}_2 (which depend only on the shape of Ω_j) so that

$$\|v\|_{L^2(\partial\Omega_j)}^2 \leq \hat{C}_1 H |v|_{H^1(\Omega_j)}^2 + \hat{C}_2 H^{-1} \|v\|_{L^2(\Omega_j)}^2.$$

Summing these inequalities over all subdomains, we have

$$2\|v\|_{L^2(\Gamma)}^2 \leq \hat{C}_{1 \max} H |v|_{H^1(\Omega)}^2 + \hat{C}_{2 \max} H^{-1} \|v\|_{L^2(\Omega)}^2.$$

Since $v \in H_0^1(\Omega)$, we can bound the second term on the right using a Friedrichs inequality. With the appropriate scaling, we find

$$2\|v\|_{L^2(\Gamma)}^2 \leq \hat{C}_{1 \max} H |v|_{H^1(\Omega)}^2 + \hat{C}_{2 \max} D^2 H^{-1} |v|_{H^1(\Omega)}^2.$$

As $H < D$, it follows that $\|v\|_{L^2(\Gamma)}^2 \leq \hat{C}_\Omega \frac{D^n}{H^{n-1}} |v|_{H^1(\Omega)}^2$, where \hat{C}_Ω does not depend on D or H . \square

Remark 3.3. In our presentation, the diffusion coefficient, for simplicity of exposition, is assumed to be 1. However, the perturbed BDDC method can be extended to the case where the coefficient is constant in each subdomain and varies greatly from

one subdomain to another. For this case, the perturbed bilinear forms in (3.1) and (3.2) are replaced by

$$\tilde{a}(u, v) = \sum_{j=1}^J \tilde{a}_j(u, v), \quad \tilde{a}_j(u, v) = \alpha_j \int_{\Omega_j} \nabla u \cdot \nabla v \, dx + \frac{\alpha_j}{D^2} \int_{\Omega_j} uv \, dx,$$

and

$$\tilde{a}(u, v) = \sum_{j=1}^J \tilde{a}_j(u, v), \quad \tilde{a}_j(u, v) = \alpha_j \int_{\Omega_j} \nabla u \cdot \nabla v \, dx + \frac{\alpha_j H_j^{n-1}}{D^n} \int_{\Gamma_j} uv \, ds,$$

respectively. Here α_j denotes the (constant) diffusion coefficient associated with the subdomain Ω_j .

In the case where the coefficient jumps do not align with subdomain interface, the perturbed formulation can also be extended to *physics-based* BDDC [10], our recently developed BDDC method especially targeting heterogeneous problems.

Remark 3.4. For linear elasticity, similar perturbed bilinear forms can be used:

$$\begin{aligned} \tilde{a}(\mathbf{u}, \mathbf{v}) &= a(\mathbf{u}, \mathbf{v}) + \frac{2\mu}{D^2} \int_{\Omega} \mathbf{u}^T \mathbf{v} \, dx, \\ \tilde{a}(\mathbf{u}, \mathbf{v}) &= a(\mathbf{u}, \mathbf{v}) + \frac{2\mu H^{n-1}}{D^n} \int_{\Gamma} \mathbf{u}^T \mathbf{v} \, ds. \end{aligned}$$

Results similar to Lemmas 3.1 and 3.2 can also be obtained since the proof of these lemmas can be extended to vector-space cases with help from Korn's first inequality:

$$\|\mathbf{v}\|_{H^1(\Omega)} \leq C (\epsilon(\mathbf{v}), \epsilon(\mathbf{v}))_{L^2(\Omega)} \quad \left(\leq C \frac{1}{2\mu} a(\mathbf{v}, \mathbf{v}) \right).$$

Remark 3.5. If the aspect ratio of the physical domain is of order 1, the domain diameter D can be obtained from the mass matrix M by pre- and post-multiplying it with the vector of ones and raising the result to $1/n$, i.e., $D = (\mathbf{1}^t M \mathbf{1})^{\frac{1}{n}}$. No geometrical information is needed. A similar statement also holds for the diameters of subdomains obtained from a good mesh partitioner, e.g., METIS [26].

4. Convergence analysis. In this section, we analyze the convergence of the perturbed BDDC preconditioner in solving (1.5). First, we present a tool to connect functions in the standard and the perturbed BDDC spaces.

LEMMA 4.1 (energy minimizing properties). *If $v \in V_h(\Gamma)$ and $\tilde{v} \in \tilde{V}_h(\Gamma)$ satisfy $v = Q_{\Gamma} \tilde{v}$, then the following inequalities hold:*

$$(4.1) \quad a(v, v) \leq a(\tilde{v}, \tilde{v}),$$

$$(4.2) \quad \tilde{a}(\tilde{v}, \tilde{v}) \leq \tilde{a}(v, v).$$

Proof. Since $v = Q_{\Gamma} \tilde{v}$, v and \tilde{v} are identical on Γ . Consequently, $v_j = v|_{\Omega_j}$ and $\tilde{v}_j = \tilde{v}|_{\Omega_j}$ are identical on $\partial\Omega_j$. Using the energy minimizing property of harmonic functions, we have $a_j(v_j, v_j) \leq a_j(\tilde{v}_j, \tilde{v}_j)$ and $\tilde{a}_j(\tilde{v}_j, \tilde{v}_j) \leq \tilde{a}_j(v_j, v_j)$. As v and \tilde{v} are both continuous functions, we can add these inequalities for all $j = 1, \dots, J$ to obtain (4.1) and (4.2). \square

We continue with some notation. Let $\theta_{\mathcal{V}}$, $\theta_{\mathcal{E}}$, and $\theta_{\mathcal{F}}$ be the characteristic finite element functions of a subdomain corner \mathcal{V} , a subdomain edge \mathcal{E} , and a subdomain face \mathcal{F} , respectively. They are essentially the discrete harmonic functions that vanish at every nodal point of $\partial\Omega_j$ except at \mathcal{V} , those of \mathcal{E} , and those of \mathcal{F} where they equal 1.

We use the notation $f \lesssim g$ to denote that $f \leq Cg$, where C does not depend on the size of the domain (D), the size of the subdomains (H), the number of subdomains (J), or particularly the constants C_L , C_U , and C_ℓ . We want to explicitly derive the change of condition number coming from the perturbation.

The following lemma is quoted from [21, Lemma 4.5] (see also [38, Lemma 2] and [49, Lemma 4.24]).

LEMMA 4.2. *Assume Ω_j is the union of a finite number of shape-regular (triangular or tetrahedral) elements. Let \mathcal{F} be a subdomain face of Ω_j , and let I_h denote the interpolation operator onto the finite element space V_h . Then, for all $u \in V_j$*

$$(4.3) \quad \|I_h(\theta_{\mathcal{F}}u)\|_{H_{00}^{1/2}(\mathcal{F})}^2 \lesssim [1 + \ln(H_j/h_j)]^2 \left(|u|_{H^{1/2}(\mathcal{F})}^2 + \frac{1}{H_j} \|u\|_{L^2(\mathcal{F})}^2 \right),$$

where $H_{00}^{1/2}(\mathcal{F})$ -norm is defined as the $H^{1/2}(\partial\Omega_j)$ -norm of the function extended by zero onto $\partial\Omega_j \setminus \mathcal{F}$ (cf. [49, section A.2]).

We note that the assumption on the shape of Ω_j is a technical requirement only for analysis. Attempts to relax this assumption have been successful only in two dimensions (see [35]). However, BDDC and FETI-DP, whose analysis relies on Lemma 4.2, work well in practice for more general partitions, e.g., those generated by graph partitioners (see [24, 33, 34, 48, 7]).

We are now ready to state one of our main results.

THEOREM 4.3. *The smallest eigenvalue of the preconditioned system $\tilde{B}_{\text{BDDC}}S_h$ satisfies the following lower bound:*

$$\lambda_{\min}(\tilde{B}_{\text{BDDC}}S_h) \geq C_U^{-1} =: \lambda_m.$$

Proof. This proof is modeled after the proof of [14, Lemma 7.8.11]. The main difference is the use of Assumption 2.1 and the energy minimizing property (4.2) to connect the perturbed formulation with the original system.

Equation (2.11) implies that for any $v \in V_h(\Gamma)$, there exists a unique $\tilde{v} \in \tilde{V}_h(\Gamma)$ such that $v = Q_\Gamma \tilde{v}$. Due to (2.9) and the fact that $\tilde{V}_h(\Gamma) \subset \tilde{\mathcal{H}}_c$, we can write $\tilde{v} = \tilde{v}_0 + \tilde{v}^\circ$, where $\tilde{v}_0 \in \tilde{\mathcal{H}}_0$ and $\tilde{v}^\circ \in \tilde{\mathcal{H}}^\circ$. Then we have

$$(4.4) \quad \tilde{v} = \tilde{P}_\Gamma(\tilde{v}_0 + \tilde{v}^\circ) = \tilde{P}_\Gamma \left(\tilde{E}_0 \tilde{v}_0 + \sum_{j=1}^J \tilde{E}_j \tilde{v}_j \right) = (\tilde{P}_\Gamma \tilde{E}_0) \tilde{v}_0 + \sum_{j=1}^J (\tilde{P}_\Gamma \tilde{E}_j) \tilde{v}_j,$$

where $\tilde{v}_j = \tilde{v}^\circ|_{\Omega_j} \in \tilde{\mathcal{H}}_j^\circ$. Consequently,

$$(4.5) \quad v = Q_\Gamma \tilde{v} = Q_\Gamma \tilde{P}_\Gamma \tilde{E}_0 \tilde{v}_0 + \sum_{j=1}^J Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j \tilde{v}_j = \sum_{j=0}^J P_j \tilde{v}_j.$$

From (4.4), (2.6), (2.8), (2.5), and (2.9), we obtain

$$\begin{aligned}
 (4.6) \quad \sum_{j=0}^J \langle \tilde{S}_j \tilde{v}_j, \tilde{v}_j \rangle &= \langle \tilde{S}_0 \tilde{v}_0, \tilde{v}_0 \rangle + \sum_{j=1}^J \tilde{a}_j(\tilde{v}_j, \tilde{v}_j) = \langle \tilde{S}_c \tilde{v}_0, \tilde{v}_0 \rangle + \langle \tilde{S}_c \tilde{v}^\circ, \tilde{v}^\circ \rangle \\
 &= \langle \tilde{S}_c(\tilde{v}_0 + \tilde{v}^\circ), (\tilde{v}_0 + \tilde{v}^\circ) \rangle = \langle \tilde{S}_c \tilde{v}, \tilde{v} \rangle = \langle \tilde{S}_h \tilde{v}, \tilde{v} \rangle = \tilde{a}(\tilde{v}, \tilde{v}).
 \end{aligned}$$

Using the energy minimizing property (4.2) and Assumption 2.1, we have

$$(4.7) \quad \tilde{a}(\tilde{v}, \tilde{v}) \leq \tilde{a}(v, v) \leq C_U a(v, v) = C_U \langle S_h v, v \rangle.$$

From (4.5), (4.6), and (4.7), it follows that

$$C_U \langle S_h v, v \rangle \geq \tilde{a}(\tilde{v}, \tilde{v}) = \sum_{j=0}^J \langle \tilde{S}_j \tilde{v}_j, \tilde{v}_j \rangle \geq \min_{\substack{v = \sum_{j=0}^J P_j \tilde{v}_j \\ \tilde{v}_0 \in \tilde{\mathcal{H}}_0, \tilde{v}_j \in \tilde{H}_j^\circ, (1 \leq j \leq J)}} \sum_{j=0}^J \langle \tilde{S}_j \tilde{v}_j, \tilde{v}_j \rangle.$$

The estimate for $\lambda_{\min}(\tilde{B}_{\text{BDDC}} S_h)$ then follows directly from the abstract additive Schwarz framework [14, 7.1.22]. \square

Let \mathcal{H}_c and P_Γ be defined similarly to $\tilde{\mathcal{H}}_c$ and \tilde{P}_Γ , with $\tilde{a}(\cdot, \cdot)$ being replaced by $a(\cdot, \cdot)$ in all the steps leading to (2.4) and (2.10), respectively. We emphasize that the functions in \mathcal{H}_c satisfy the same constraints in \mathcal{C} as the functions in $\tilde{\mathcal{H}}_c$.

The following assumption introduces our simplified and combined definition of *acceptable edge path* [38] and *acceptable face path* [39] for the case of a constant diffusion coefficient. It is needed in the analysis of options in which not all geometrical objects are included in \mathcal{C} to define constraints. For the case where the coefficient varies from one subdomain to another, the original definitions of acceptable paths are required (cf. [38, 39, 49]).

Assumption 4.4. For two subdomains Ω_a, Ω_b that share an edge but not a face or a corner but not an edge, either this shared object, which we call \mathcal{G} , is included in the set of constraints \mathcal{C} , or there exists a sequence of subdomains $\{\Omega_a = \Omega_{j_1}, \Omega_{j_2}, \dots, \Omega_{j_k} = \Omega_b\}$ satisfying the following properties:

- (i) they all share the common object \mathcal{G} ;
- (ii) for $j = 1, \dots, k - 1$, subdomains Ω_{j_m} and $\Omega_{j_{m+1}}$ must share, apart from \mathcal{G} , an object in \mathcal{C} of type \mathfrak{T} ; here type \mathfrak{T} is either an edge or a face and is predefined.

LEMMA 4.5. *Let $w \in \mathcal{H}_c$ and $u = w - P_\Gamma w$. If Assumption 4.4 is true, then the following estimate holds:*

$$(4.8) \quad |u_j|_{H^1(\Omega_j)}^2 \lesssim [1 + \ln(H/h)]^2 \sum_{k \in \mathcal{N}_j} |w_k|_{H^1(\Omega_k)}^2,$$

where $u_j = u|_{\Omega_j}$, $w_k = w|_{\Omega_k}$, and $\mathcal{N}_j = \{k : \partial\Omega_k \cap \partial\Omega_j \neq \emptyset\}$. Consequently,

$$(4.9) \quad \sum_{j=1}^J a_j(u_j, u_j) \lesssim [1 + \ln(H/h)]^2 \sum_{j=1}^J a_j(w_j, w_j)$$

or, equivalently,

$$(4.10) \quad |P_\Delta w|_S \lesssim [1 + \ln(H/h)]^2 |w|_S,$$

where $P_\Delta := I - P_\Gamma$ and $|w|_S := \langle S_c w, w \rangle = \sum_{j=1}^J a_j(w_j, w_j)$.

The estimates (4.8), (4.9), and (4.10) are typical in the analysis of FETI-DP [24], a method closely related to BDDC. For their proofs, we refer the reader to [38, Lemma 10], [39, Lemma 7], and [49, Lemma 6.36]. Here, we only want to emphasize that with our choice of constraints \mathcal{C} , Assumption 4.4 holds.

Remark 4.6. Lemma 4.5 also holds for linear elasticity; however, the requirements of the constraints are different. In three dimensions, we need at least six constraints in order to control the rigid-body modes of a subdomain. For example, if we use only face constraints, then constraints on first-order moments can be added to fully control the rigid-body modes. For more information about nonoverlapping methods for elasticity, we refer the reader to [36] and [37].

THEOREM 4.7. *The largest eigenvalue of the preconditioned system $\tilde{B}_{\text{BDDC}}S_h$ satisfies the following upper bound:*

$$\lambda_{\max}(\tilde{B}_{\text{BDDC}}S_h) \lesssim \frac{C_U}{C_L \min\{C_\ell, C_L\}} [1 + \ln(H/h)]^2 =: \lambda_M.$$

Proof. For this proof, we follow the main steps of the proof of [14, Lemma 7.8.11] with significant changes to adapt to the perturbed formulation.

For any $v \in V_h(\Gamma)$, we consider an arbitrary decomposition

$$v = Q_\Gamma \tilde{P}_\Gamma \tilde{E}_0 \tilde{v}_0 + \sum_{j=1}^J Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j \tilde{v}_j = \sum_{j=0}^J P_j \tilde{v}_j,$$

where $\tilde{v}_0 \in \tilde{\mathcal{H}}_0$, $\tilde{v}_j \in \tilde{\mathcal{H}}_j^\circ$, $1 \leq j \leq J$. According to the abstract additive Schwarz framework [14, 7.1.21], Theorem 4.7 will follow if we can show that

$$(4.11) \quad \langle S_h v, v \rangle \lesssim \frac{C_U}{C_L \min\{C_\ell, C_L\}} [1 + \ln(H/h)]^2 \sum_{j=0}^N \langle \tilde{S}_j \tilde{v}_j, \tilde{v}_j \rangle.$$

First, we need to make a connection between the perturbed formulation and the standard one. Let \mathcal{H}_0 , \mathcal{H}_j° , $P_\Gamma : \mathcal{H}_c \rightarrow V_h(\Gamma)$ and $E_j : \mathcal{H}_j \rightarrow L^2(\Omega)$, where $1 \leq j \leq J$, be defined similarly to $\tilde{\mathcal{H}}_0$, $\tilde{\mathcal{H}}_j^\circ$, \tilde{P}_Γ , and \tilde{E}_j but w.r.t. $a(\cdot, \cdot)$ instead of $\tilde{a}(\cdot, \cdot)$ (they are essentially the spaces and operators defined in the standard formulation of BDDC [13, 14]). Since the functions in \mathcal{H}_0 , \mathcal{H}_j° , $\tilde{\mathcal{H}}_0$, and $\tilde{\mathcal{H}}_j^\circ$ are completely defined by their values on $\Gamma_j = \Gamma \cap \partial\Omega_j$, there uniquely exist $v_0 \in \mathcal{H}_0$ and $v_j \in \mathcal{H}_j^\circ$ such that $v_0|_{\Gamma_j} = \tilde{v}_0|_{\Gamma_j}$, $v_j|_{\Gamma_j} = \tilde{v}_j|_{\Gamma_j}$. Therefore, their averaged extensions are the same; i.e.,

$$(4.12) \quad P_\Gamma E_j v_j = Q_\Gamma \tilde{P}_\Gamma \tilde{E}_j \tilde{v}_j = P_j \tilde{v}_j, \quad j = 0, \dots, J.$$

Consequently, $v = P_\Gamma E_0 v_0 + \sum_{j=1}^J P_\Gamma E_j v_j$. Now let

$$(4.13) \quad \tilde{v} = \tilde{P}_\Gamma \tilde{E}_0 \tilde{v}_0 + \sum_{j=1}^J \tilde{P}_\Gamma \tilde{E}_j \tilde{v}_j =: \sum_{j=0}^J \tilde{P}_j \tilde{v}_j;$$

then $\tilde{v} \in \tilde{V}_h(\Gamma)$ and $v = Q_\Gamma \tilde{v}$. This implies that v and \tilde{v} are identical on the interface Γ . Using the energy-minimizing property (4.1) and Assumption 2.1, we have

$$(4.14) \quad a(v, v) \leq a(\tilde{v}, \tilde{v}) \leq C_L^{-1} \tilde{a}(\tilde{v}, \tilde{v}).$$

From (1.4), (2.3), and (4.14), it follows that

$$(4.15) \quad \langle S_h v, v \rangle = a(v, v) \leq C_L^{-1} \tilde{a}(\tilde{v}, \tilde{v}) = C_L^{-1} \langle \tilde{S}_h \tilde{v}, \tilde{v} \rangle.$$

On the other hand, from (4.13), we find that

$$(4.16) \quad \begin{aligned} \langle \tilde{S}_h \tilde{v}, \tilde{v} \rangle &= \left\langle \tilde{S}_h \sum_{j=0}^J \tilde{P}_j \tilde{v}_j, \sum_{j=0}^J \tilde{P}_j \tilde{v}_j \right\rangle \\ &\leq 2 \left(\left\langle \tilde{S}_h \tilde{P}_0 \tilde{v}_0, \tilde{P}_0 \tilde{v}_0 \right\rangle + \left\langle \tilde{S}_h \sum_{j=1}^J \tilde{P}_j \tilde{v}_j, \sum_{j=1}^J \tilde{P}_j \tilde{v}_j \right\rangle \right) \\ &\lesssim \sum_{j=0}^J \langle \tilde{S}_h \tilde{P}_j \tilde{v}_j, \tilde{P}_j \tilde{v}_j \rangle, \end{aligned}$$

where we have used the fact that the support of $\tilde{P}_j \tilde{v}_j$ lies only on a bounded number of subdomains. Since $P_j \tilde{v}_j = Q_\Gamma \tilde{P}_j \tilde{v}_j$, we have that $P_j \tilde{v}_j \in V_h(\Gamma)$ is identical with $\tilde{P}_j \tilde{v}_j \in \tilde{V}_h(\Gamma)$ on Γ . Using the minimizing property (4.2), it follows that

$$(4.17) \quad \begin{aligned} \langle \tilde{S}_h \tilde{P}_j \tilde{v}_j, \tilde{P}_j \tilde{v}_j \rangle &= \tilde{a}(\tilde{P}_j \tilde{v}_j, \tilde{P}_j \tilde{v}_j) \leq \tilde{a}(P_j \tilde{v}_j, P_j \tilde{v}_j) \\ &= \tilde{a}(P_\Gamma E_j v_j, P_\Gamma E_j v_j) \leq C_U a(P_\Gamma E_j v_j, P_\Gamma E_j v_j), \end{aligned}$$

where in the last line we have used (4.12) and Assumption 2.1.

In the next step, we will relate $a(P_\Gamma E_j v_j, P_\Gamma E_j v_j)$ to $a_j(v_j, v_j)$. Using the triangle inequality and Lemma 4.5 for $w = E_j v_j$ and $u = E_j v_j - P_\Gamma E_j v_j$, we obtain

$$(4.18) \quad \begin{aligned} a(P_\Gamma E_j v_j, P_\Gamma E_j v_j) &= \sum_{k=1}^J a_k(P_\Gamma E_j v_j, P_\Gamma E_j v_j) \\ &\leq \sum_{k=1}^J \left(a_k(E_j v_j, E_j v_j) + a_k(E_j v_j - P_\Gamma E_j v_j, E_j v_j - P_\Gamma E_j v_j) \right) \\ &\lesssim \sum_{k=1}^J a_k(E_j v_j, E_j v_j) + [1 + \ln(H/h)]^2 \sum_{k=1}^J a_k(E_j v_j, E_j v_j). \end{aligned}$$

For $j = 0$, we have

$$(4.19) \quad \begin{aligned} a(P_\Gamma E_0 v_0, P_\Gamma E_0 v_0) &\lesssim \sum_{k=1}^J a_k(v_0, v_0) + [1 + \ln(H/h)]^2 \sum_{k=1}^J a_k(v_0, v_0) \\ &\lesssim [1 + \ln(H/h)]^2 \sum_{k=1}^J a_k(v_0, v_0). \end{aligned}$$

For $j = 1, \dots, J$, (4.18) implies that

$$(4.20) \quad \begin{aligned} a(P_\Gamma E_j v_j, P_\Gamma E_j v_j) &\lesssim a_j(v_j, v_j) + [1 + \ln(H/h)]^2 a_j(v_j, v_j) \\ &\lesssim [1 + \ln(H/h)]^2 a_j(v_j, v_j). \end{aligned}$$

Using (4.19), (4.20), Assumption 2.1 (for $j = 0$), and Assumption 2.3 (for $j = 1, \dots, J$), we find that

$$\begin{aligned}
 (4.21) \quad a(P_\Gamma E_j v_j, P_\Gamma E_j v_j) &\lesssim [1 + \ln(H/h)]^2 a_j(\tilde{v}_j, \tilde{v}_j) \\
 &\leq \frac{1}{\min\{C_\ell, C_L\}} [1 + \ln(H/h)]^2 \tilde{a}_j(\tilde{v}_j, \tilde{v}_j) \\
 &= \frac{1}{\min\{C_\ell, C_L\}} [1 + \ln(H/h)]^2 \langle \tilde{S}_j \tilde{v}_j, \tilde{v}_j \rangle.
 \end{aligned}$$

Here we have a slight abuse of notation with $a_0(\cdot, \cdot) = \sum_{k=1}^J a_k(\cdot, \cdot)$ and $\tilde{a}_0(\cdot, \cdot) = \sum_{k=1}^J \tilde{a}_k(\cdot, \cdot)$.

Inequality (4.11) then follows from (4.15), (4.16), (4.17), and (4.21). □

Combining Theorems 4.3 and 4.7, we obtain the following main result.

THEOREM 4.8. *There exists a positive constant C , independent of h, H, J, C_U, C_L , and C_ℓ , such that*

$$\kappa(\tilde{B}_{\text{BDDC}} S_h) = \frac{\lambda_{\max}(\tilde{B}_{\text{BDDC}} S_h)}{\lambda_{\min}(\tilde{B}_{\text{BDDC}} S_h)} \leq C \frac{(C_U)^2}{C_L \min\{C_\ell, C_L\}} \left(1 + \ln \frac{H}{h}\right)^2.$$

We have claimed that the well-posedness of the perturbed BDDC does not depend on the choice of constraints or the type of partition utilized. However, as Theorems 4.7 and 4.8 rely on the validity of Lemma 4.5, the choice of constraints still plays an important role in achieving a good rate of convergence. The essential condition in the proof of Lemma 4.5 is that for any object \mathcal{G} , it is either in the set of constraints \mathcal{C} or there is an acceptable path connecting a list of subdomains associated with \mathcal{G} . In the next step, we will study what will happen if that condition does not hold.

LEMMA 4.9. *Assume that there are not enough constraints in \mathcal{C} for (4.8) to hold with subdomain index j , for some $1 \leq j \leq J$. Then, with the perturbed bilinear forms suggested in section 3, we still have the following, weaker, estimate:*

$$(4.22) \quad |u_j|_{H^1(\Omega_j)}^2 \lesssim [1 + \ln(H/h)]^2 (D/H)^m \sum_{k \in \mathcal{N}_j} \tilde{a}_k(w_k, w_k),$$

where $m = n$, the dimension of the physical space, if the Robin perturbation is used; otherwise, $m = 2$.

Proof. From the definition of P_Γ in (2.10), we can write that

$$(4.23) \quad u(x) = \sum_{k \in \mathcal{N}(x)} \frac{1}{|\mathcal{N}(x)|} (w_j(x) - w_k(x)) \quad \forall x \in \Gamma \cap \partial\Omega_j,$$

where $\mathcal{N}(x) = \{k : x \in \Gamma_k = \Gamma \cap \partial\Omega_k\}$.

We cut u_j using $\theta_\mathcal{V}$, $\theta_\mathcal{E}$, and $\theta_\mathcal{F}$ and write it as a sum of terms which are only nonzero on individual subdomain corners, edges, and faces, respectively, i.e.,

$$(4.24) \quad u_j = \sum_{\mathcal{V} \subset \partial\Omega_j} \theta_\mathcal{V} u_j + \sum_{\mathcal{E} \subset \partial\Omega_j} I^h(\theta_\mathcal{E} u_j) + \sum_{\mathcal{F} \subset \partial\Omega_j} I^h(\theta_\mathcal{F} u_j).$$

For simplicity, we only present the estimate for the face terms. Let \mathcal{F} be the face

shared between subdomains Ω_j and Ω_k . From (4.23) and Lemma 4.2, we find that

$$\begin{aligned} \|I^h(\theta_{\mathcal{F}}u_j)\|_{H_0^{1/2}(\mathcal{F})}^2 &= \left\| I^h\left(\theta_{\mathcal{F}}\frac{1}{2}(w_j - w_k)\right) \right\|_{H_0^{1/2}(\mathcal{F})}^2 \\ &\lesssim [1 + \ln(H_j/h_j)]^2 \left(|w_j|_{H^{1/2}(\mathcal{F})}^2 + \frac{1}{H_j}\|w_j\|_{L^2(\mathcal{F})}^2 + |w_k|_{H^{1/2}(\mathcal{F})}^2 + \frac{1}{H_k}\|w_k\|_{L^2(\mathcal{F})}^2 \right). \end{aligned}$$

Similar estimates for the edge and vertex terms can be obtained from the proof of [38, Lemma 10] or [49, Lemma 6.36] with some modification. Summing all of the estimates for individual objects, it follows that

$$(4.25) \quad |u_j|_{H^1(\Omega_j)}^2 \lesssim [1 + \ln(H/h)]^2 \sum_{k \in \mathcal{N}_j} \left(|w_k|_{H^{1/2}(\Gamma_k)}^2 + \frac{1}{H}\|w_k\|_{L^2(\Gamma_k)}^2 \right).$$

Since w_k is a harmonic function w.r.t. $a_k(\cdot, \cdot)$, we have $|w_k|_{H^{1/2}(\Gamma_k)} = a_k(w_k, w_k)$. In addition, when the perturbation with the full mass matrix is used, we note that $H\|w_k\|_{L^2(\Gamma_k)}^2 \leq \|w_k\|_{L^2(\Omega_k)}^2$. Consequently, (4.25) and (3.1) imply that

$$\begin{aligned} |u_j|_{H^1(\Omega_j)}^2 &\lesssim [1 + \ln(H/h)]^2 \sum_{k \in \mathcal{N}_j} \left(a_k(w_k, w_k) + \frac{1}{H^2}\|w_k\|_{L^2(\Omega_k)}^2 \right) \\ (4.26) \quad &\leq [1 + \ln(H/h)]^2 (D/H)^2 \sum_{k \in \mathcal{N}_j} \tilde{a}_k(w_k, w_k). \end{aligned}$$

In case Robin perturbation is used, from (4.25) and (3.2), we have

$$(4.27) \quad |u_j|_{H^1(\Omega_j)}^2 \leq [1 + \ln(H/h)]^2 (D/H)^n \sum_{k \in \mathcal{N}_j} \tilde{a}_k(w_k, w_k).$$

Combining (4.26) and (4.27), we obtain (4.22). \square

THEOREM 4.10. *If the perturbed bilinear forms introduced in section 3 are used, even in the extreme case where there is no constraint in \mathcal{C} , we still have*

$$(4.28) \quad \lambda_{\max}(\tilde{B}_{\text{BDDC}}S_h) \lesssim \frac{C_U}{C_L \min\{C_\ell, C_L\}} (D/H)^m [1 + \ln(H/h)]^2 =: \hat{\lambda}_M,$$

$$(4.29) \quad \kappa(\tilde{B}_{\text{BDDC}}S_h) \leq C (D/H)^m \left(1 + \ln \frac{H}{h} \right)^2,$$

where C does not depend on the mesh sizes, the size of the domain and subdomains, or the number of subdomains.

Proof. The proof is a combination of Theorem 4.3, which does not depend on the choice of \mathcal{C} , and a modified version of the proof of Theorem 4.7 where Lemma 4.9 is used instead of Lemma 4.5. \square

Remark 4.11. Theorem 4.10 gives a bound for the condition number of the preconditioned system in the case when no constraint is used. We can see that the bound increases as the number of subdomains increases. Even though the perturbed preconditioner without constraints is well-posed, it is not a scalable one, like any one-level DD method.

For the rest of this section, we will study a more practical case where the set of constraints \mathcal{C} fails to eliminate a small number of piecewise constant functions or rigid-body modes from the BDDC space. In other words, we consider $\ker S_c \neq \emptyset$ and $\dim(\ker S_c) \ll \dim(V_h(\Gamma))$, where $S_c : \mathcal{H}_c \rightarrow \mathcal{H}'_c$ is the counterpart of \tilde{S}_c in the standard formulation of the BDDC method, namely

$$(4.30) \quad \langle S_c u, v \rangle = \sum_{i=1}^J a_j(u_j, v_j) \quad \forall u, v \in \mathcal{H}_c, u_j = u|_{\Omega_j}, v_j = v|_{\Omega_j}.$$

THEOREM 4.12. *Assume that $\ker S_c \neq \emptyset$; then the spectrum of the preconditioned system, counting multiplicities, can be decomposed as*

$$(4.31) \quad \sigma(\tilde{B}_{\text{BDDC}} S_h) = \mathcal{S}_1 \cup \mathcal{S}_2,$$

where $|\mathcal{S}_1| \leq \dim(\ker(S_c))$, $\mathcal{S}_1 \subset [\lambda_m, \hat{\lambda}_M]$, and $\mathcal{S}_2 \subset [\lambda_m, \lambda_M]$. Here, the constants λ_m , λ_M , and $\hat{\lambda}_M$ are defined in Theorems 4.3, 4.7, and 4.10, respectively.

Proof. For the sake of brevity, we will use \tilde{B} and S to refer to \tilde{B}_{BDDC} and S_h , respectively. We define $V_0 = P_\Gamma \ker(S_c)$. Then $V_h(\Gamma)$ can be decomposed as

$$(4.32) \quad V_h(\Gamma) = V_0 \oplus_S V_0^\perp, \quad \text{i.e., } \langle S v_0, v_0^\perp \rangle = 0, \forall v_0 \in V_0, v_0^\perp \in V_0^\perp.$$

Define $((\cdot, \cdot)) := \langle \tilde{B}^{-1} \cdot, \cdot \rangle$, an inner product on $V_h(\Gamma)$. Clearly, $\tilde{B}S$ is SPD w.r.t. $((\cdot, \cdot))$. In addition, for any $v_0 \in V_0$ and $v_0^\perp \in V_0^\perp$,

$$(4.33) \quad ((\tilde{B}S v_0^\perp, v_0)) = ((\tilde{B}S v_0, v_0^\perp)) = \langle S v_0, v_0^\perp \rangle = 0.$$

Let $\mathcal{B}_1 = \{\phi_1, \dots, \phi_{n_1}\}$ and $\mathcal{B}_2 = \{\phi_{n_1+1}, \dots, \phi_{n_1+n_2}\}$ be some bases of V_0 and V_0^\perp ; then $\mathcal{B} = \{\phi_1, \dots, \phi_{n_1+n_2}\}$ is a basis of $V_h(\Gamma)$. Let \mathbf{A} be the matrix representation of $\tilde{B}S$ w.r.t. the basis \mathcal{B} . From (4.33), we have

$$(4.34) \quad ((BS\phi_j, \phi_i)) = ((BS\phi_i, \phi_j)) = 0 \quad \text{for } n_1 + 1 \leq j \leq n_1 + n_2, 1 \leq i \leq n_1.$$

As the j th column of \mathbf{A} is the coefficient in the representation of $\tilde{B}S\phi_j$ w.r.t. \mathcal{B} , (4.34) implies that \mathbf{A} is a block diagonal matrix and $\mathbf{A} = \text{diag}(\mathbf{A}^1, \mathbf{A}^2)$, where \mathbf{A}^1 and \mathbf{A}^2 are the matrix representations of $\tilde{B}S|_{V_0}$ and $\tilde{B}S|_{V_0^\perp}$ w.r.t. the bases \mathcal{B}_1 and \mathcal{B}_2 , respectively. This leads to $\sigma(\mathbf{A}) = \sigma(\mathbf{A}^1) \cup \sigma(\mathbf{A}^2)$. Consequently,

$$\sigma(\tilde{B}S) = \sigma(\tilde{B}S|_{V_0}) \cup \sigma(\tilde{B}S|_{V_0^\perp}).$$

Since $\dim(V_0) \leq \dim(\ker(S_c))$, it follows that

$$|\sigma(\tilde{B}S|_{V_0})| \leq \dim(\ker(S_c)).$$

As Theorems 4.3 and 4.10, where the lower bound and a weak upper bound of $\sigma(\tilde{B}S)$ are obtained, hold regardless of the choice of \mathcal{C} , the proof is finished if we can show that

$$(4.35) \quad \lambda_{\max}(\tilde{B}S|_{V_0^\perp}) \leq \lambda_M.$$

Since $\tilde{B}S$ is symmetric positive definite w.r.t. $((\cdot, \cdot))$, it follows that

$$(4.36) \quad \lambda_{\max}(\tilde{B}S|_{V_0^\perp}) = \max_{v \in V_0^\perp, v \neq 0} \frac{((\tilde{B}Sv, v))}{((v, v))} = \max_{v \in V_0^\perp, v \neq 0} \frac{\langle Sv, v \rangle}{\langle \tilde{B}^{-1}v, v \rangle}.$$

Since V_0^\perp does not contain the piecewise constant functions in the case of Poisson's equation or any of the subdomain rigid-body modes in the case of linear elasticity, the results in Lemma 4.5 and Theorem 4.7 hold, and we have the estimate (4.35) for the largest eigenvalue of $\tilde{B}S|_{V_0^\perp}$. \square

Remark 4.13. When the constraints in \mathcal{C} fail to eliminate a small number of constant functions or rigid-body modes, i.e., $\ker(S_c) \neq \emptyset$ and $\dim(\ker(S_c)) \ll \dim(V_h(\Gamma))$, Theorem 4.12 indicates that most of the eigenvalues of the preconditioned system can still be bounded by the usual bounds in the standard BDDC methods. Some of the remaining eigenvalues might be larger than the usual upper bound. However, they are isolated (the number of them is bounded from above by $\dim(\ker(S_c))$). As large isolated eigenvalues can only delay the convergence of the CG method by a few iterations (cf. [2]), the perturbed BDDC method is still scalable. Obviously, the number of iterations will be higher than in the case in which all the piecewise constant functions or all the rigid-body modes are eliminated from the BDDC space by additional constraints, since one considers a smaller coarse space.

5. Numerical experiments. In this section, we will study several numerical experiments to verify the robustness as well as the efficiency of the perturbed BDDC method.

The perturbed formulation is built on top of the BDDC preconditioner in the scientific computing software package FEMPAR (FE Multiphysics massively PARallel) [4], a parallel hybrid MPI+OpenMP, object-oriented framework for the massively parallel FE simulation of multiscale multiphysics problems governed by PDEs [6, 7]. All our experiments were run on MareNostrum III at BSC (Barcelona) and HLRN-III at RRZN (Hannover).

In our experiments, both the standard and the perturbed BDDC preconditioners with different options of constraints are used to solve the linear system associated with (1.5) by the CG method. The number of CG iterations and the time (in seconds) to reduce the residual by at least a factor of 1e-6 will be reported. Here, time includes both the set-up time and the iterative solver time.

In presenting the data, C, E, and F legends indicate that corner, edge, and face constraints are used, respectively. The suffix 0 expresses that the standard BDDC formulation (no perturbation) is used. The suffix CD emphasizes that the corner selection algorithm in [48] and the standard BDDC formulation are used. If the legend is without a suffix, then it represents a result with perturbed BDDC formulation and no corner selection algorithm is involved.

In addition, we present only results for Robin perturbation. Perturbation with a full mass matrix yields almost identical results (see [11]).

5.1. Poisson's equation in a cube. In this experiment, we compare the performance of the perturbed and standard BDDC methods for a Poisson equation discretized by structured meshes. We consider the problem described in (1.1), where Ω is the unit cube $[0, 1] \times [0, 1] \times [0, 1]$. The right-hand side function f is chosen so that $u = x + y + z$ is the analytical solution.

We use uniform structured hexahedral meshes which are partitioned into $k \times k \times k$, $k = 3, \dots, 11$, subdomains of exactly the same shape by regular partitions. In order to study the weak scalability of the methods, when k increases (the size of the subdomains, H , decreases), we use smaller mesh size, h , to keep H/h constant, i.e., a constant local problem size.

Figure 1 shows the weak scalability in CG iteration (the left column), and in time

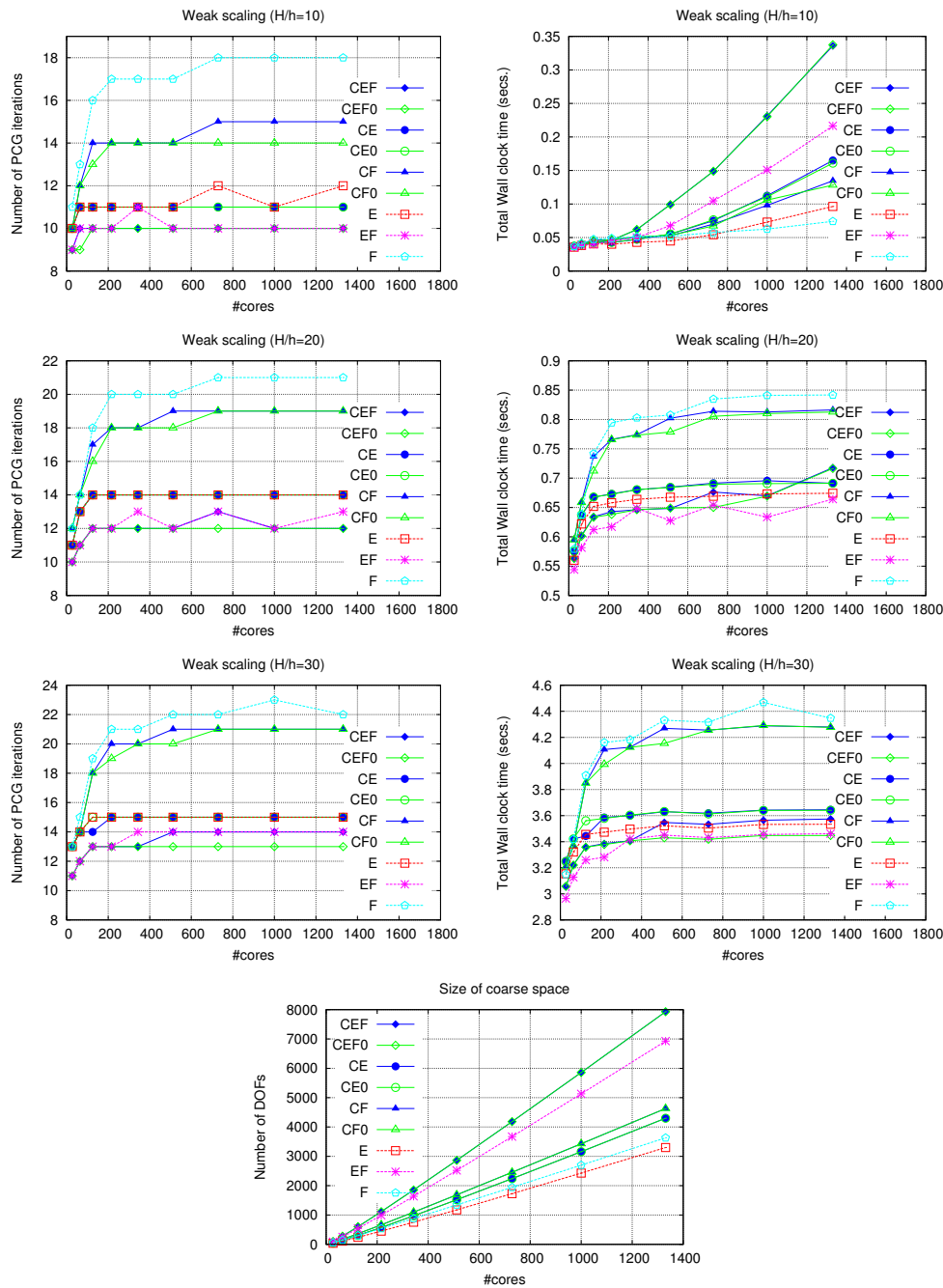


FIG. 1. Poisson's equation in a cube.

(the right column). Three different test cases, namely $H/h = 10, 20, 30$, are reported. The figure also presents (in the last row) the sizes of the coarse spaces, which depends only on the partition.

From Figure 1, we can conclude that the perturbed BDDC method, for all the

proposed choices of constraints, is weakly scalable; namely, the numbers of iterations are almost constant when the number of subdomains increases. The performance of the perturbed BDDC method in both iteration number and time is also very close to that of the standard BDDC method. In this experiment, compared to the standard method, the perturbed BDDC method only requires at most one extra iteration to reduce the residual to the same relative tolerance of $1e-6$.

On the different choices of constraints, we can reconfirm that corner constraints do not contribute much to the convergence of the methods as EF, E, F have the same number of iterations as CEF, CE, CF, and have almost the same number of iterations as CEF0, CE0, CF0, respectively. The choices without corner constraints (which need perturbation to be well-posed) always outperform the corresponding choices with corner constraints in terms of time whether perturbation is used or not. In terms of iteration number, the choices with larger coarse spaces (cf. the last rows of Figure 1) generally require fewer iterations, except among the ones with E and F. Obviously, the performance in time depends not only on the choices of constraints but also on the ratio H/h .

In order to understand the time performance, it is important to note that the BDDC method has fine tasks (solving local Neumann problems, computing the contribution to the coarse matrix, computing the fine corrections, etc.), which are performed in parallel on each processor, and a coarse task (computing the coarse correction), which is centralized in one processor. Roughly speaking, the fine tasks become more expensive (in both computation and communication) when the size of local problem (also the ratio H/h) increases, and the cost of coarse task rises when the size of the coarse space becomes larger. In the advanced implementation of BDDC methods in FEMPAR (cf. [7]), these two types of tasks run in parallel and are overlapped in run time. As long as the run time of the coarse task can be completely overlapped by the run time of the fine tasks, perfect weak scalability is attained.

For $H/h = 10$, the local problems are small and the coarse tasks are clearly dominant. The options with large coarse spaces require more time even though they use fewer iterations (the overlapping is not effective). In those situations, in order to make the overlapping approach effective, a multilevel BDDC method would be required [9]. For $H/h = 30$, the local problems are reasonably large, and the fine tasks dictate the time performance. The performance in terms of time resembles the performance in iteration number in this test case. The options with large coarse spaces arrive to the solution in shorter time. For $H/h = 20$, at first, the fine tasks are dominant. However, as the number of subdomains increases, the size of the coarse space increases. We can see that the coarse task becomes more and more significant and eventually dominant for the options with larger coarse spaces, namely CEF, CEF0, CE, and CE0. We will come back to this in the experiment for linear elasticity.

5.2. Poisson's equation in backward facing step domain. In this experiment, we compare the performance of the perturbed BDDC with standard BDDC for a Poisson equation discretized by unstructured meshes. We consider the problem described in (1.1), where Ω is the backward facing step domain depicted in Figure 2. Solutions for this type of problem are required in numerical simulations of (turbulent) incompressible flows using the pressure segregation techniques [23, 3].

For this problem, a homogeneous Dirichlet boundary condition is set on the whole boundary of Ω except for the leftmost face, where the unknown is constrained to be 1. We also set the force term $f = \cos(4\pi x) \cos(4\pi y) \cos(4\pi z)$.

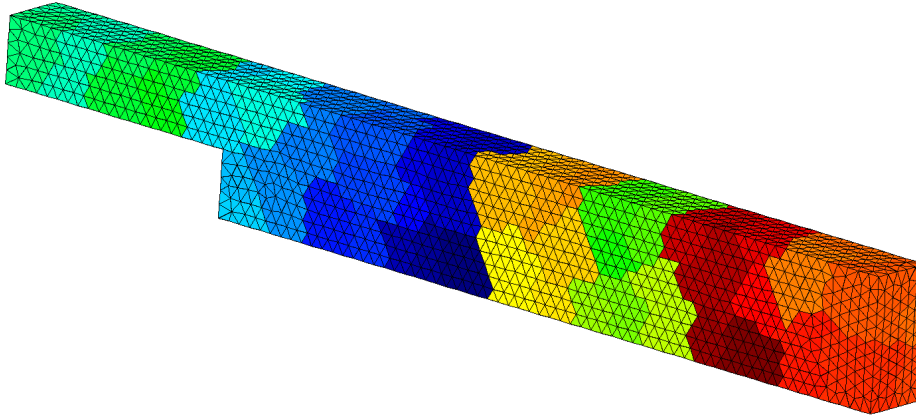


FIG. 2. Three-dimensional backward facing step domain with unstructured tetrahedral mesh and its partition into 64 subdomains using METIS.

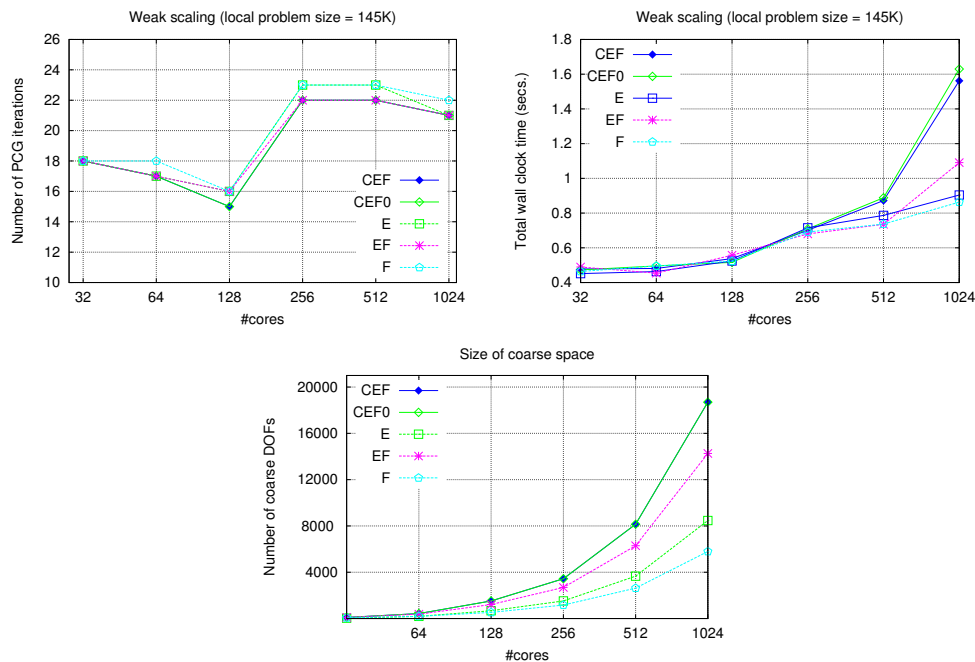


FIG. 3. Poisson's equation with backward facing step domain.

For this experiment, starting from four unstructured tetrahedral meshes with roughly 20K, 40K, 80K, and 160K elements, we use uniform refinements and METIS [26] to create a set of meshes with 32, 64, 128, 256, 512, and 1024 subdomains. The number of nodes in each subdomain is kept roughly around 145K.

In Figure 3, we plot the numbers of iterations, elapsed time, and number of coarse degrees of freedom (DOFs) over the number of subdomains (#cores). As results are close for different options of constraints, for better visibility, we show only the ones for CEF0, CEF, EF, E, and F.

From Figure 3, we can see that even though the iteration count for each option is bounded when the number of subdomains increases, all of them vary more than in the experiment in subsection 5.1, where structured meshes and regular partitions are used. However, as all of the partitions are created by METIS [26], we can control only the local problem size—not the shape and size of the subdomains or the ratio H/h . Taking this into account, we can safely say that each of the considered options of the constraints yields a scalable BDDC variant. In addition, the performance of perturbed BDDC (with CEF) is very close to that of the standard BDDC (with CEF0) in both iterations count and time. All options without corners (EF, E, and F), which are only available for the perturbed BDDC, work and have favorable performance in time.

5.3. Elasticity of a long beam. In this experiment, we compare the performance of the perturbed and standard BDDC methods for the linear elasticity problem in (1.2) with the Lamé parameters $\lambda = 0.1$ and $\mu = 1.0$ and the body force $\mathbf{f} = [0.0 \ -0.005 \ 0.0]^T$.

Let Ω be the long beam $[0, 2] \times [0, 0.5] \times [0, 0.5]$. We use the natural boundary condition on the whole boundary of Ω except on one side (the plane $x = 0$), where the homogeneous Dirichlet boundary condition is imposed. We also use uniform structured hexahedral meshes which are partitioned into $4k \times k \times k$, $k = 2, \dots, 11$, subdomains. Figure 4 shows the displacement of the beam (the solution).

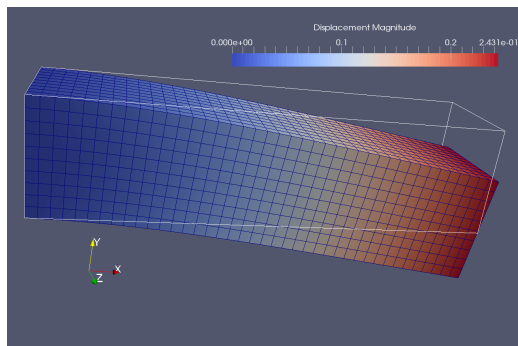


FIG. 4. *Linear elasticity—long beam: The displacement of the beam under the body force (right).*

For this test problem, simply choosing the constraints associated with geometrical objects as discussed in subsection 2.1 is not enough to make the standard BDDC method work. We need to use a corner selection algorithm or switch to the perturbed formulation. Among the possible options, we carry out the experiment for CEF, CEF_CD, EF, and E.

Figure 5 shows the weak scaling in terms of iteration number (left column) and time (right column), and the size of the coarse space (last row) for three different test cases with $H/h = 10, 20, 30$.

The result for this problem is a reconfirmation of what we have found earlier for the Poisson equation. From Figure 5, it is clear that all studied options of constraints succeed in making the BDDC methods scalable. The numbers of iterations are almost constant when the number of subdomains increases. The options with larger coarse spaces require fewer iterations to converge. However, the time performance depends on the size of local problems. For small local problems ($H/h = 10$), the coarse task is dominant, and options with smaller coarse spaces perform better. For large local problems ($H/h = 30$), the fine tasks dictate the result, and options with larger coarse

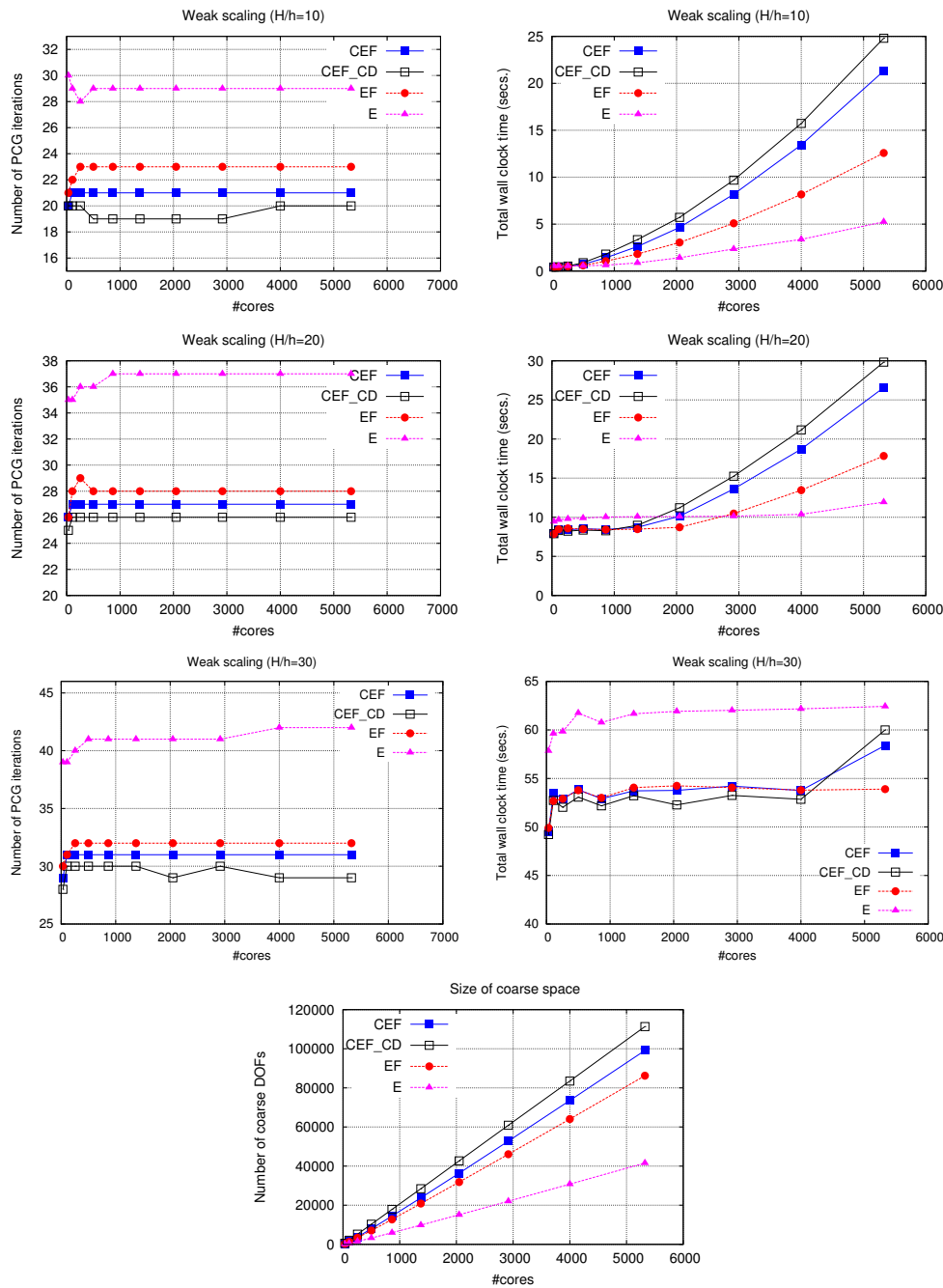


FIG. 5. Linear elasticity—long beam: Robin perturbation.

spaces converge to the solution in less time. For medium-size problems ($H/h = 20$), the fine tasks are dominant for a small number of subdomains but eventually are overtaken by the coarse task when the number of subdomains increases.

In conclusion, if time-to-solution is the ultimate goal, then it might not be optimal

to choose the constraints so that the number of iterations required is minimal. The choices of constraints should be based on the size of local problems and the number of subdomains used. Particularly, if one is interested in solving very large problems with many processors, then the options with minimal coarse spaces, such as E or F, are likely the winners. We emphasize that these two options are only available for the perturbed BDDC method.

5.4. Poisson's equation with heterogeneous coefficient. In this experiment, we verify our claim in Remark 3.3 that the perturbed BDDC can be extended to the case in which the diffusion coefficient is constant in each subdomain and varies greatly from one subdomain to another.

We consider the problem described in (1.1), where Ω is the unit square $[0, 1] \times [0, 1]$. The right-hand side function is $f = 1$. We consider structured triangular meshes and their regular partitions into $k \times k$, $k = 5, 10, 15$, subdomains. The ratio H/h is kept constant at 10. The diffusion coefficient is defined by

$$\log_{10}(\alpha_j) = \rho * \text{mod}(j, 5)/4,$$

where j is the subdomain index, numbering first in the x -direction and then in the y -direction. Figure 6 shows the coefficient and partition for the case $k = 10$ and $\rho = 6$. Basically, we have alternating channels with different values of the coefficient.

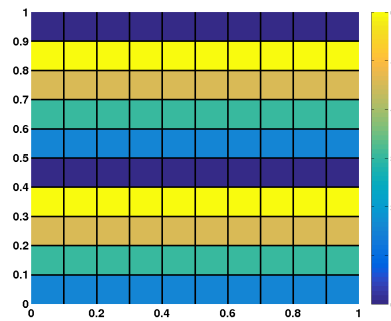


FIG. 6. Partition and coefficient distribution in the heterogeneous coefficient experiment.

In Table 1, iteration numbers for the options CE0, CE, and E are shown when ρ takes values in $\{2, 4, 6\}$. We can see that for all options, the number of iterations is almost constant when the number of subdomains or ρ increases. In addition, the performance of the perturbed BDDC method (CE) is almost the same as that of the standard BDDC method (CE0).

TABLE 1
Iteration numbers required in the Poisson equation with heterogeneous coefficient.

$k \rightarrow$	5			10			15		
	CE0	CE	E	CE0	CE	E	CE0	CE	E
2	11	11	14	11	12	15	11	12	16
4	11	12	15	12	12	16	12	12	16
6	12	12	16	12	12	17	12	12	17

5.5. Poisson's equation with insufficient constraints. In this experiment, we would like to verify Theorem 4.12 and Remark 4.13. We study the weak scalability of the perturbed BDDC method when there are not enough constraints to exclude all of the piecewise constant functions from the BDDC space. We repeat the experiment in subsection 5.1 but with some *floating subdomains*, where no constraint is imposed.

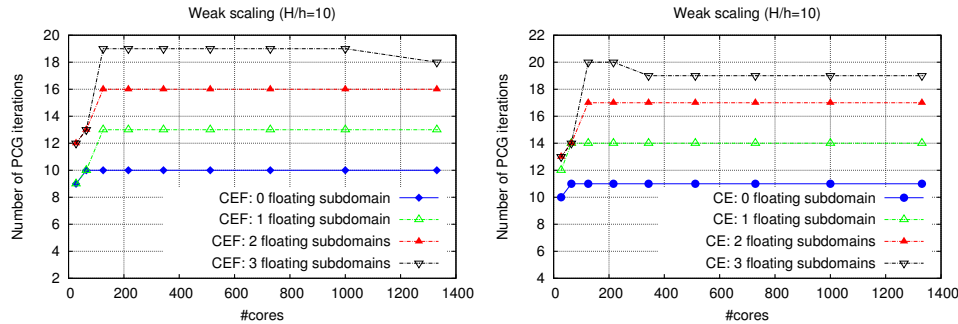


FIG. 7. *Poisson's equation: Some floating subdomains with no constraint.*

Figure 7 shows the iteration numbers required for the two options CEF and CE when different sets of subdomains fail to have any constraint. For comparison, original iteration numbers with all the constraints associated with these two options being enforced are also shown. We can see that even though the iteration numbers are higher than the original ones, they are almost constant when the number of subdomains increases. In other words, scalability is maintained even when we do not impose enough constraints to eliminate all of the piecewise constant functions from the BDDC space. This is predicted in Theorem 4.12 and Remark 4.13.

5.6. Elasticity of a long beam with disconnected subdomains. In this experiment, we study whether the perturbed BDDC method can maintain scalability when there are disconnected subdomains in the partition. We work with the same linear elasticity problem of the long beam described in the previous subsection. However, we consider a different partition of Ω into $8k \times 2k \times 2k$, $k = 1, \dots, 6$, subdomains where half of them consist of four disconnected parts. Figure 8 shows the partition when $k = 1$ on the left and a close-up of one disconnected subdomain on the right. For this experiment, the size of the local problem is the same as in the test case with $H/h = 20$ and regular partition. However, due to the special partition, the size of some of the subdomains are larger, and the ratio in this test case is $H/h = 30$.

For this experiment, even the corner detecting algorithm in [48] cannot make the standard BDDC method work. On the other hand, the perturbed BDDC method not only works but is also scalable (see Figure 9). This is even better than predicted in Theorem 4.12 as in this case the number of rigid-body modes which are not eliminated from the BDDC space is proportional to the number of subdomains. However, the numbers of iterations increase compared to the test case $H/h = 30$ with regular partition. When only few subdomains are disconnected, the increase in the number of iterations is less significant.

5.7. Elasticity of a cross link. In the last experiment, we work with a more realistic problem, where Ω is a mechanical object shown in Figure 10. With three unstructured tetrahedral meshes, using uniform refinements and METIS [26], we create 5 meshes with 32, 89, 256, 711, and 2048 subdomains. For each mesh, the number of

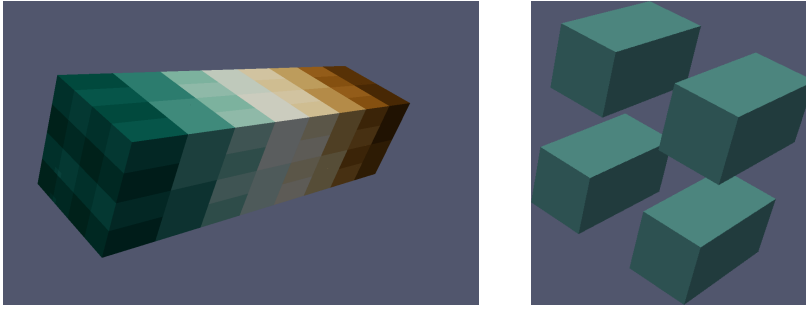


FIG. 8. *Linear elasticity—long beam: A partition with disconnected subdomains (left) and the close-up of one of its disconnected subdomains (right).*

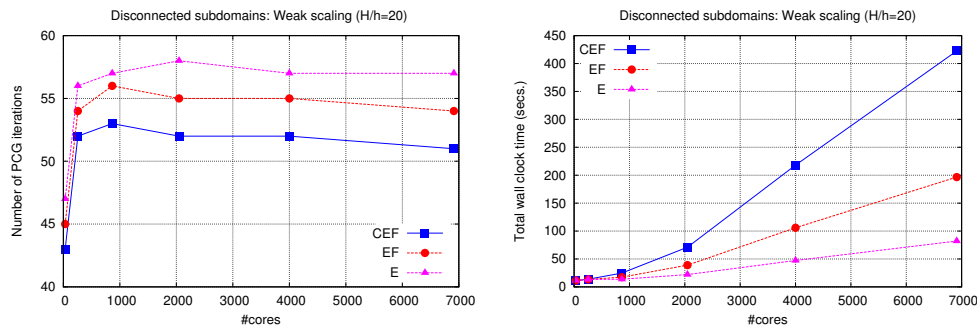


FIG. 9. *Linear elasticity—long beam: Disconnected subdomains.*

nodes in each subdomain is kept roughly the same at around 44K. The largest mesh has more than 25M elements and 4M nodes. Out of the 2048 subdomains of this mesh, one is disconnected.

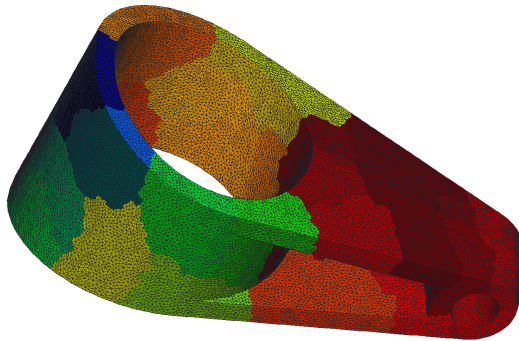


FIG. 10. *Linear elasticity—cross link: A mechanical object with a mesh and a partitioning.*

In Figure 11, we plot the results for the choices of constraints associated with CEF_CD, CEF, EF, E, and F. The standard BDDC with corner detection algorithm [48] fails to handle the largest mesh where there is a disconnected subdomain. All perturbed formulations (CEF, EF, E, and F) converge. Taking into account the fact that only the local problem size is fixed and not the ratio H/h , all options except F

have the numbers of iterations bounded independently of the number of subdomains. In other words, they are scalable. As expected, the method with only face constraints (F) is not weakly scalable. This is due to the fact that with face constraints only, not all rigid-body modes are fixed.

In terms of time, the options with smaller coarse spaces, namely EF, E, and F, have better performance. The advantage in time of these options becomes more and more significant when the number of subdomains increases. For the largest mesh, even though the option with only face constraints requires many more iterations, it is still the fastest option.

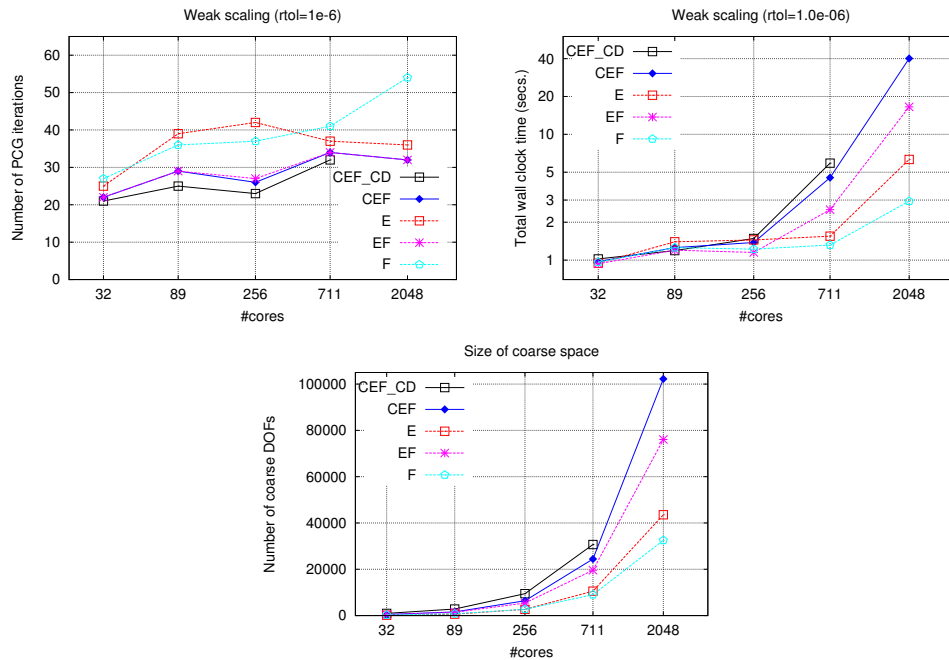


FIG. 11. *Linear elasticity—cross link: Weak scaling test.*

6. Concluding remarks. In this work, we have formulated and analyzed a perturbed BDDC method. The method is proved to be algorithmically weakly scalable. It is also verified by numerical experiments to be as scalable and efficient as the standard method for the same set of coarse constraints. Moreover, it does not rely on the constraints to guarantee the well-posedness and offers more freedom in choosing them. Notably, the implementation with just face and/or edge constraints does not require a change of basis for using sparse direct solvers for definite systems, and corner constraints and corner selection algorithms are not needed. These choices lead to small coarse spaces that are ideal for extreme scale computation. But the main advantage of the method is that the perturbed formulation brings extra robustness as it works even when the constraints fail to eliminate a small number of rigid-body modes from the BDDC space, or even when the partition has disconnected subdomains. This point is of great importance when dealing with unstructured meshes and automatic partitioners.

Even though our presentation is for problems with constant coefficients in each subdomain, our work can be extended to problems with truly heterogeneous coeffi-

cients. It is known that for these problems, more complicated strategies are needed to define the constraints; e.g., the perturbed formulation could also be combined with other approaches, such as deluxe scaling [20], adaptive coarse spaces [46, 31, 28, 47], and especially physics-based BDDC [10], to deal with heterogeneity of the coefficients.

The perturbation approach presented herein is a simple way to relax the roles of corner constraints. Based on our analysis, any selection of constraints known to result in a fast algorithm for the original problem would also result in a fast algorithm for the perturbed problem. Thus, strategies for choosing constraints such as those in [36, 37] could also be used in future work. On the other hand, the proof of Theorem 4.12 provides a way to rigorously analyze convergence properties of preconditioners with *local defects*. Even though we have developed this theory for the case where constraints do not fix all the rigid-body modes, it could also be applied to the analysis of fault-tolerant (resilient) preconditioners. This topic is critical for the forthcoming exascale supercomputers. These ideas will be further elaborated elsewhere.

Acknowledgments. We would like to thank the editor and the anonymous referees, whose suggestions helped improve the paper significantly. Finally, the authors thankfully acknowledge the computer resources, technical expertise, and assistance provided by the Red Española de Supercomputación and the North-German Supercomputing Alliance (HLRN).

REFERENCES

- [1] S. ALURU AND F. E. SEVILGEN, *Parallel domain decomposition and load balancing using space-filling curves*, in Proceedings of the 4th International Conference on High-Performance Computing, IEEE, Washington, DC, 1997, pp. 230–235.
- [2] O. AXELSSON AND G. LINDSKOG, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523, <https://doi.org/10.1007/BF01389448>.
- [3] S. BADIA AND R. CODINA, *Algebraic pressure segregation methods for the incompressible Navier-Stokes equations*, Arch. Comput. Methods Engrg., 15 (2008), pp. 343–369, <https://doi.org/10.1007/s11831-008-9020-3>.
- [4] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *FEMPAR: Finite Element Multiphysics Massively Parallel*, <https://web.cimne.upc.edu/groups/comfus/fempar.html>.
- [5] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *Enhanced balancing Neumann-Neumann preconditioning in computational fluid and solid mechanics*, Internat. J. Numer. Methods Engrg., 96 (2013), pp. 203–230, <https://doi.org/10.1002/nme.4541>.
- [6] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *Implementation and scalability analysis of balancing domain decomposition methods*, Arch. Comput. Methods Engrg., 20 (2013), pp. 239–262, <https://doi.org/10.1007/s11831-013-9086-4>.
- [7] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *A highly scalable parallel implementation of balancing domain decomposition by constraints*, SIAM J. Sci. Comput., 36 (2014), pp. C190–C218, <https://doi.org/10.1137/130931989>.
- [8] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *On the scalability of inexact balancing domain decomposition by constraints with overlapped coarse/fine corrections*, Parallel Comput., 50 (2015), pp. 1–24, <https://doi.org/10.1016/j.parco.2015.09.004>.
- [9] S. BADIA, A. F. MARTÍN, AND J. PRINCIPE, *Multilevel balancing domain decomposition at extreme scales*, SIAM J. Sci. Comput., 38 (2016), pp. C22–C52, <https://doi.org/10.1137/15M1013511>.
- [10] S. BADIA AND H. NGUYEN, *Physics-Based Balancing Domain Decomposition by Constraints for Heterogeneous Problems*, HAL:hal-01337968 [math.NA], 2016.
- [11] S. BADIA AND H. NGUYEN, *Relaxing the Roles of Corners in BDDC by Perturbed Formulation*, HAL:hal-01337852 [math.NA], 2016.
- [12] M. BEBENDORF, *A note on the Poincaré inequality for convex domains*, Z. Anal. Anwendungen, 22 (2003), pp. 751–756, <https://doi.org/10.4171/ZAA/1170>.

- [13] S. C. BRENNER, *A functional analytic framework for BDDC and FETI-DP*, in Domain Decomposition Methods in Science and Engineering XVII, U. Langer, M. Discacciati, D. E. Keyes, O. B. Widlund, and W. Zulehner, eds., Lect. Notes Comput. Sci. Eng. 60, Springer, Berlin, 2008, pp. 239–246, https://doi.org/10.1007/978-3-540-75199-1_27.
- [14] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008, <https://doi.org/10.1007/978-0-387-75934-0>.
- [15] S. C. BRENNER AND L.-Y. SUNG, *BDDC and FETI-DP without matrices or vectors*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1429–1435, <https://doi.org/10.1016/j.cma.2006.03.012>.
- [16] C. CHEVALIER AND F. PELLEGRINI, *PT-SCOTCH: A tool for efficient parallel graph ordering*, Parallel Comput., 34 (2008), pp. 318–331, <https://doi.org/10.1016/j.parco.2007.12.001>.
- [17] Y.-H. DE ROECK AND P. LE TALLEC, *Analysis and test of a local domain-decomposition preconditioner*, in Proceedings of the 4th International Symposium on Domain Decomposition Methods for Partial Differential Equations (Moscow, 1990), SIAM, Philadelphia, 1991, pp. 112–128.
- [18] C. R. DOHRMANN, *A preconditioner for substructuring based on constrained energy minimization*, SIAM J. Sci. Comput., 25 (2003), pp. 246–258, <http://doi.org/10.1137/S1064827502412887>.
- [19] C. R. DOHRMANN, *An approximate BDDC preconditioner*, Numer. Linear Algebra Appl., 14 (2007), pp. 149–168, <https://doi.org/10.1002/nla.514>.
- [20] C. R. DOHRMANN AND O. B. WIDLUND, *A BDDC algorithm with deluxe scaling for three-dimensional $H(\text{curl})$ problems*, Comm. Pure Appl. Math., 69 (2016), pp. 745–770, <https://doi.org/10.1002/cpa.21574>.
- [21] M. DRYJA, B. F. SMITH, AND O. B. WIDLUND, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal., 31 (1994), pp. 1662–1694, <https://doi.org/10.1137/0731086>.
- [22] M. DRYJA AND O. B. WIDLUND, *Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems*, Comm. Pure Appl. Math., 48 (1995), pp. 121–155, <https://doi.org/10.1002/cpa.3160480203>.
- [23] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, 2nd ed., Numer. Math. Sci. Comput., Oxford University Press, Oxford, UK, 2014, <https://doi.org/10.1093/acprof:oso/9780199678792.001.0001>.
- [24] C. FARHAT, M. LESOINNE, P. LE TALLEC, K. PIERSON, AND D. RIXEN, *FETI-DP: A dual-primal unified FETI method—part I: A faster alternative to the two-level FETI method*, Internat. J. Numer. Methods Engrg., 50 (2001), pp. 1523–1544, <https://doi.org/10.1002/nme.76>.
- [25] C. FARHAT, M. LESOINNE, AND K. PIERSON, *A scalable dual-primal domain decomposition method*, Numer. Linear Algebra Appl., 7 (2000), pp. 687–714, [https://doi.org/10.1002/1099-1506\(200010/12\)7:7/8<687::AID-NLA219>3.0.CO;2-S](https://doi.org/10.1002/1099-1506(200010/12)7:7/8<687::AID-NLA219>3.0.CO;2-S).
- [26] G. KARYPIS AND V. KUMAR, *A fast and high quality multilevel scheme for partitioning irregular graphs*, SIAM J. Sci. Comput., 20 (1998), pp. 359–392, <https://doi.org/10.1137/S1064827595287997>.
- [27] G. KARYPIS, K. SCHLOEGEL, AND V. KUMAR, *ParMETIS: Parallel Graph Partitioning and Sparse Matrix Ordering Library*, Tech. report, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, 1997.
- [28] H. H. KIM AND E. T. CHUNG, *A BDDC algorithm with enriched coarse spaces for two-dimensional elliptic problems with oscillatory and high contrast coefficients*, Multiscale Model. Simul., 13 (2015), pp. 571–593, <https://doi.org/10.1137/140970598>.
- [29] A. KLAWONN, M. LANSER, AND O. RHEINBACH, *Toward extremely scalable nonlinear domain decomposition methods for elliptic partial differential equations*, SIAM J. Sci. Comput., 37 (2015), pp. C667–C696, <https://doi.org/10.1137/140997907>.
- [30] A. KLAWONN, M. LANSER, AND O. RHEINBACH, *FE² TI: Computational scale bridging for dual-phase steels*, in Parallel Computing: On the Road to Exascale, G. R. Joubert, H. Leather, M. Parsons, F. Petersfe, and M. Sawyer, eds., Adv. Parallel Comput. 27, IOS Press, Amsterdam, 2016, pp. 797–806, <https://doi.org/10.3233/978-1-61499-621-7-797>.
- [31] A. KLAWONN, P. RADTKE, AND O. RHEINBACH, *FETI-DP methods with an adaptive coarse space*, SIAM J. Numer. Anal., 53 (2015), pp. 297–320, <https://doi.org/10.1137/130939675>.
- [32] A. KLAWONN AND O. RHEINBACH, *A parallel implementation of dual-primal FETI methods for three-dimensional linear elasticity using a transformation of basis*, SIAM J. Sci. Comput., 28 (2006), pp. 1886–1906, <https://doi.org/10.1137/050624364>.

- [33] A. KLAWONN AND O. RHEINBACH, *Inexact FETI-DP methods*, Internat. J. Numer. Methods Engrg., 69 (2007), pp. 284–307, <https://doi.org/10.1002/nme.1758>.
- [34] A. KLAWONN AND O. RHEINBACH, *Robust FETI-DP methods for heterogeneous three dimensional elasticity problems*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1400–1414, <https://doi.org/10.1016/j.cma.2006.03.023>.
- [35] A. KLAWONN, O. RHEINBACH, AND O. B. WIDLUND, *An analysis of a FETI-DP algorithm on irregular subdomains in the plane*, SIAM J. Numer. Anal., 46 (2008), pp. 2484–2504, <https://doi.org/10.1137/070688675>.
- [36] A. KLAWONN AND O. B. WIDLUND, *Selecting constraints in dual-primal FETI methods for elasticity in three dimensions*, in Domain Decomposition Methods in Science and Engineering, Lect. Notes Comput. Sci. Eng. 40, Springer, Berlin, 2005, pp. 67–81, https://doi.org/10.1007/3-540-26825-1_5.
- [37] A. KLAWONN AND O. B. WIDLUND, *Dual-primal FETI methods for linear elasticity*, Comm. Pure Appl. Math., 59 (2006), pp. 1523–1572, <https://doi.org/10.1002/cpa.20156>.
- [38] A. KLAWONN, O. B. WIDLUND, AND M. DRYJA, *Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients*, SIAM J. Numer. Anal., 40 (2002), pp. 159–179, <https://doi.org/10.1137/S0036142901388081>.
- [39] A. KLAWONN, O. B. WIDLUND, AND M. DRYJA, *Dual-primal FETI methods with face constraints*, in Recent Developments in Domain Decomposition Methods (Zürich, 2001), Lect. Notes Comput. Sci. Eng. 23, Springer, Berlin, 2002, pp. 27–40, https://doi.org/10.1007/978-3-642-56118-4_2.
- [40] M. LESOINNE, *A FETI-DP corner selection algorithm for three-dimensional problems*, in Domain Decomposition Methods in Science and Engineering XIV, I. Herrera, D. E. Keyes, O. B. Widlund, and R. Yates, eds., National Autonomous University of Mexico (UNAM), Mexico City, Mexico, 2003, pp. 217–224.
- [41] J. LI AND O. B. WIDLUND, *FETI-DP, BDDC, and block Cholesky methods*, Internat. J. Numer. Methods Engrg., 66 (2006), pp. 250–271, <https://doi.org/10.1002/nme.1553>.
- [42] J. LI AND O. B. WIDLUND, *On the use of inexact subdomain solvers for BDDC algorithms*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1415–1428, <https://doi.org/10.1016/j.cma.2006.03.011>.
- [43] J. MANDEL, *Balancing domain decomposition*, Comm. Numer. Methods Engrg., 9 (1993), pp. 233–241, <https://doi.org/10.1002/cnm.1640090307>.
- [44] J. MANDEL, C. R. DOHRMANN, AND R. TEZAUER, *An algebraic theory for primal and dual substructuring methods by constraints*, Appl. Numer. Math., 54 (2005), pp. 167–193, <https://doi.org/10.1016/j.apnum.2004.09.022>.
- [45] J. MANDEL, B. SOUSEDÍK, AND C. DOHRMANN, *Multispace and multilevel BDDC*, Computing, 83 (2008), pp. 55–85, <https://doi.org/10.1007/s00607-008-0014-7>.
- [46] J. MANDEL AND B. SOUSEDÍK, *Adaptive selection of face coarse degrees of freedom in the BDDC and the FETI-DP iterative substructuring methods*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1389–1399, <https://doi.org/10.1016/j.cma.2006.03.010>.
- [47] C. PECHSTEIN AND C. R. DOHRMANN, *A Unified Framework for Adaptive BDDC*, Tech. report RICAM-Report 2016-20, Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Linz, Austria, 2016.
- [48] J. ŠÍSTEK, M. ČERTÍKOVÁ, P. BURDA, AND J. NOVOTNÝ, *Face-based selection of corners in 3D substructuring*, Math. Comput. Simul., 82 (2012), pp. 1799–1811, <https://doi.org/10.1016/j.matcom.2011.06.007>.
- [49] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods—Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005.
- [50] X. TU, *Three-level BDDC in three dimensions*, SIAM J. Sci. Comput., 29 (2007), pp. 1759–1780, <https://doi.org/10.1137/050629902>.