Newton-Krylov-FETI-DP with Adaptive Coarse Spaces

Axel Klawonn[1], Martin Lanser[1], Balthasar Niehoff[1], Patrick Radtke[1], and Oliver Rheinbach[2]

1 Introduction

Newton-Krylov domain decomposition methods are approaches for solving nonlinear problems arising from the discretization of nonlinear partial differential equations. These methods are based on an iterative solution of linearized systems using a domain decomposition preconditioner. In this paper, we use FETI-DP as an iterative method and compute an adaptive coarse space, first introduced in [11], to improve the condition number and thus the convergence of the iterative method. A theory has been developed in [6] for this coarse space in two dimensions and later, in [4], for three dimensions. In this paper, several heuristic strategies are introduced to reduce the computational effort for nonlinear problems, where a sequence of related linear problems have to be solved. These approaches show the potential of reducing the number of eigenvalue problems necessary for the construction of adaptive coarse spaces. A different but related approach was presented in [2].

2 Newton-Krylov-FETI-DP

In order to solve a discrete nonlinear equation

$$\widehat{K}(\hat{u}) - \hat{f} = 0, \tag{1}$$

[1]Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany. e-mail:{axel.klawonn,martin.lanser,patrick.radtke}@uni-koeln.de .[2]Institut für Numerische Mathematik und Optimierung, Fakultät für Mathematik und Informatik, Technische Universität Bergakademie Freiberg, Akademiestr. 6, 09596 Freiberg, Germany. e-mail: oliver.rheinbach@math.tu-freiberg.de

associated with a computational domain Ω, we perform a Newton linearization of (1) and compute an update $\delta\hat{u}$ by solving the linearized system

$$D\widehat{K}(\hat{u})\,\delta\hat{u} = \widehat{K}(\hat{u}) - \hat{f}. \tag{2}$$

We always consider an iterative Krylov method such as CG to solve (2) using a domain decomposition preconditioner. In this paper, we always consider a FETI-DP (Finite Element Tearing and Interconnecting - Dual-Primal) preconditioner although a BDDC method could also be used. Therefore, we decompose Ω into nonoverlapping subdomains Ω_i, $i = 1, \dots, N$, and assume the subdomains to be unions of finite elements. We denote the finite element space associated with Ω by $\widehat{W}$ and the local finite element spaces associated with the subdomains by W_i, $i = 1, \dots, N$. Let us define local nonlinear problems in W_i, $i = 1, \dots, N$, by

$$K^{(i)}(u_i) = f_i. \tag{3}$$

These local problems arise from a finite element discretization on subdomains Ω_i, $i = 1, \dots, N$. The corresponding tangential matrices are defined as $DK^{(i)}(u_i)$. We introduce the block vectors

$$K(u) := \begin{pmatrix} K^{(1)}(u_1) \\ \vdots \\ K^{(N)}(u_N) \end{pmatrix},\ u := \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix},\ f := \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix}, \tag{4}$$

and the block tangential matrix

$$DK(u) = \begin{bmatrix} DK^{(1)}(u_1) & & \\ & \ddots & \\ & & DK^{(N)}(u_N) \end{bmatrix}. \tag{5}$$

In FETI-DP type methods, we divide all degrees of freedom into variables inside subdomains (I), dual interface variables (Δ), and primal variables (Π). Using the partial assembly operator R^T, well-known from the standard (linear) FETI-DP literature [1, 8, 10, 7], we can define the partially assembled operator $\widetilde{K}(\tilde{u}) := R^T K(R\tilde{u})$. Here, we perform a global assembly in all primal variables Π, but not in the remaining part of the interface. Equivalently, we partially assemble the right hand side $\tilde{f} := R^T f$ and the tangential matrix $D\widetilde{K}(\widetilde{u}) := R^T DK(\widetilde{u})R$. We define the space of partially assembled functions by $\widetilde{W} \subset W := W_1 \times \cdots \times W_N$. Introducing the standard FETI-DP jump operator B and Lagrange multipliers to enforce the constraint $B\,\widetilde{u} = 0$, the FETI-DP master system reads

$$\begin{pmatrix} D\widetilde{K}(\widetilde{u}) & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \delta\tilde{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} \widetilde{K}(\widetilde{u}) - \tilde{f} \\ 0 \end{pmatrix}. \tag{6}$$

ADAPTIVE-NEWTON-KRYLOV-FETI-DP

Init: $\widetilde{u}^{(0)} \in W$, continuous
for $k = 0, ..., convergence$

build: $\widetilde{K}(\widetilde{u}^{(k)})$ and $D\widetilde{K}(\widetilde{u}^{(k)})$
if cond_func(k, $r^{(0)}$,...,$r^{(k)}$, $its(0)$,..., $its(k-1)$**)**
compute adaptive coarse space using tangent $D\widetilde{K}(\widetilde{u}^{(k)})$
else
recycle adaptive coarse space from step $k-1$
end if
solve with preconditioned CG:
$M_{BP}^{-1}\, B\,(D\widetilde{K}(\widetilde{u}^{(k)}))^{-1}\, B^T\, \lambda = M_{BP}^{-1}\, B\,(D\widetilde{K}(\widetilde{u}^{(k)}))^{-1}\,(\widetilde{K}(\widetilde{u}^{(k)}) - \tilde{f})$
compute:
$\delta\tilde{u}^{(k)} = D\widetilde{K}(\widetilde{u}^{(k)})^{-1}\,(\widetilde{K}(\widetilde{u}^{(k)}) - \tilde{f} - B^T\lambda)$ // Compute $\delta\tilde{u}$ from λ.
compute: steplength $\alpha^{(k)}$
update: $\widetilde{u}^{(k+1)} := \widetilde{u}^{(k)} - \alpha^{(k)}\,\delta\tilde{u}^{(k)}$

end

Fig. 1 Algorithmic description of Adaptive-Newton-Krylov-FETI-DP.

At convergence, the solution $\delta\tilde{u}$ of (6) is continuous on the interface and thus can be assembled to the solution $\delta\hat{u}$ in (2). We finally obtain a solution of system (6) by eliminating all variables of $\delta\tilde{u}$ and using a preconditioned Krylov subspace method and solve

$$M_{BP}^{-1} F\,\lambda := M_{BP}^{-1}\, B\,(D\widetilde{K}(\widetilde{u}))^{-1}\, B^T\,\lambda = M_{BP}^{-1}\, B\,(D\widetilde{K}(\widetilde{u}))^{-1}\,(\widetilde{K}(\widetilde{u}) - \tilde{f}). \quad (7)$$

In this paper, we use the balancing preconditioner M_{BP}^{-1}, see, e.g., [9], for the Lagrange multipliers, implementing a second, adaptive coarse space computed from eigenvalue problems based on localized tangential matrices; see Section 3. The preconditioner M_{BP}^{-1} is defined by $M_{BP}^{-1} = (I-P)M^{-1}(I-P) + U(U^TFU)^{-1}U^T$, where $P = U(U^TFU)^{-1}U^TF$ is an F-orthogonal projection onto range$\,U$. The columns of U represent additional constraints of the form $U^TB\,\widetilde{u} = 0$. For more details on the balancing preconditioner applied to FETI-DP methods, we refer to [9]. We denote the resulting algorithm by Adaptive-Newton-Krylov-FETI-DP; see Fig. 1 for the algorithm.

3 Adaptive coarse space

In the following, we briefly describe an adaptive approach first introduced in [11]. For other uses of this coarse space and modifications, see, e.g., [12, 13, 6]. A theory is provided in [4, 6]. Due to space limitations, for further references on other adaptive coarse spaces, see, e.g., [15, 14, 3], and the references therein. Let the Schur complements S_l be obtained by eliminating

the interior degrees of freedom in $DK^{(l)}(u_l)$, $l = i, j$. We define $B_{D,ij}$ as the matrix with rows of $[B_D^{(i)} B_D^{(j)}]$ which correspond to Lagrange multipliers connecting degrees of freedom on $\partial\Omega_i \cap \partial\Omega_j$ and by B_{ij} the corresponding rows in $[B^{(i)} B^{(j)}]$. We then build a local operator $P_{D,ij} = B_{D,ij}^T B_{ij}$. Let $\widetilde{W}_{ij}$ be the subspace of functions in $W_i \times W_j$ which are continuous at those primal vertices that the two substructures Ω_i and Ω_j have in common. Let Π_{ij} be the l_2-orthogonal projection from $W_i \times W_j$ onto $\widetilde{W}_{ij}$. Let $\sigma > 0$ and $\overline{\Pi}_{ij}$ be the l_2-orthogonal projection that projects orthogonally the elements of $\ker(\Pi_{ij} S_{ij} \Pi_{ij} + \sigma(I - \Pi_{ij}))$ onto constants. In our computations we use $\sigma = \max(\mathrm{diag}\,(S_{ij}))$. To compute adaptive constraints, for each pair of substructures (Ω_i, Ω_j) having an edge in common, we solve the eigenvalue problem

$$\begin{aligned} &\overline{\Pi}_{ij} \Pi_{ij} P_{D_{ij}}^T S_{ij} P_{D_{ij}} \Pi_{ij} \overline{\Pi}_{ij} w_{ij,m} \\ =&\mu_{ij,m} (\overline{\Pi}_{ij} (\Pi_{ij} S_{ij} \Pi_{ij} + \sigma(I - \Pi_{ij})) \overline{\Pi}_{ij} + \sigma(I - \overline{\Pi}_{ij})) w_{ij,m}, \end{aligned} \tag{8}$$

for eigenpairs where $\mu_{ij,m} \geq \mathrm{TOL}$, $m = k, \ldots, n$. We implement the constraints $w_{ij,m}^T P_{D_{ij}}^T S_{ij} P_{D_{ij}} w_{ij} = 0$ for $w_{ij} \in W_i \times W_j$ and $m = k, \ldots, n$. The adaptive constraint vectors are then given by $u_{ij,m} = B_{D_{ij}} S_{ij} P_{D_{ij}} w_{ij,m}$. They are extended by zero on the remaining interface and aggregated in the matrix U.

In our Adaptive-Newton-Krylov-FETI-DP method, we also use heuristic strategies to decide if the adaptive coarse space can be recycled in a certain Newton step. Only if some condition $cond\_func(k, r^{(0)}, ..., r^{(k)}, its(0), ..., its(k-1))$ is fulfilled in the k-th Newton step, we do compute a new adaptive coarse space. Otherwise, we recycle the coarse space already used in the previous Newton step. We suppose, that conditions can be provided that depend on the nonlinear residuals $r^{(l)} := \widetilde{K}(u^{(l)}) - \tilde{f}$, $l = 0, \ldots, k$, the current iteration k, or the number of Krylov iterations $its(l)$ in the previous Newton steps $l = 0, \ldots k-1$. In the present paper, we propose three different strategies. **Strategy a)**: $cond\_func := true$, **Strategy b)**: $cond\_func := (k == 0)$, or **Strategy c)**: $cond\_func := ((its(k-1)/its(c) < 0.75) \vee (its(c)/its(k-1) < 0.75))$. For **Strategy a)**, we can prove a theoretical condition number bound for each linearization; see [6]. **Strategy b)** is based on the assumption that the optimal coarse space mainly depends on a coefficient function ρ. Therefore, the coarse space computed in the first Newton iteration can be recycled, since the coefficient function ρ does not change during the iteration. In **Strategy c)** we compute an adaptive coarse space in the first Newton step. In the following steps we consider the number of Krylov iterations in the previous Newton step ($its(k-1)$) and the last Newton step in which an adaptive coarse space has been computed ($its(c)$). We always compute a new coarse space if $its(k-1)$ and $its(c)$ differ strongly. This strategy is based on the assumption that the quality of the coarse space in the c-th Newton step is verified by theoretical results and thus we can recycle our current coarse space as long as we have similar iteration counts as in step c. Let us remark that Strategy b) will

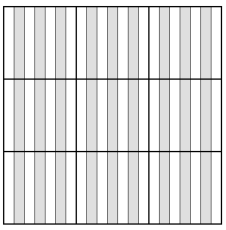

Fig. 2 Decomposition of $\Omega = [0,1] \times [0,1]$ into 3×3 subdomains. Each subdomain is intersected by 3 channels (gray color). All channels are unions of finite elements and the union of all channels is denoted by Ω_C.

not succeed for elastoplasticity problems, see [5], for which we suggest the use of **Strategy a)**. Alternatively, the knowledge of the plastic zones could be included into the heuristic function *cond_func*. This is ongoing research and will be published elsewhere.

4 Numerical Results

As a model problem, we consider the p-Laplace equation with $p = 4$

$$\begin{aligned} -\mathrm{div}(\rho\,|\nabla u|^2 \nabla u) &= 1 \qquad \text{in } \Omega \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{9}$$

where $\rho : \Omega \to \mathbb{R}$ is a coefficient function given by

$$\rho(x) = \begin{cases} 1e6 & \text{if } x \in \Omega_C \\ 1 & \text{elsewhere;} \end{cases} \tag{10}$$

see Fig. 2 for a definition of Ω_C. Let us remark that, given a finite element basis $\{\varphi_1, \ldots, \varphi_{N_i}\}$ on a subdomain Ω_i, we have

$$K^{(i)}(u_i) := \left(\int_{\Omega_i} \rho\,|\nabla u_i|^{p-2}\ \nabla u_i^T \nabla \varphi_1 dx\ ,\ \ldots\ ,\ \int_{\Omega_i} \rho\,|\nabla u_i|^{p-2} \nabla u_i^T \nabla \varphi_{N_i} dx \right)^T.$$

For the tangential matrices $DK^{(i)}(u_i)$, we obtain

$$\begin{aligned} (DK^{(i)}(u_i))_{j,k} := &\int_{\Omega_i} \rho\,|\nabla u_i|^{p-2} \nabla \varphi_j^T \nabla \varphi_k dx \\ &+ (p-2) \int_{\Omega_i} \rho\,|\nabla u_i|^{p-4}\ (\nabla u_i^T \nabla \varphi_j)\ (\nabla u_i^T \nabla \varphi_k) dx. \end{aligned}$$

This tangential matrix is symmetric positive definite for all nonconstant functions u. We present numerical results for model problem (9) in Table 1

TOL=1000										
N	**Strategy**	**Newton It.**	**Max. Krylov It.**	**Min. Krylov It.**	**Total Krylov It.**	**Max. cond.**	**Min. cond.**	**Interface d.o.f.**	**Avg. size U**	**EP Solves**
4	—	20	7	5	132	1.2	1.0	113	—	0
	a)	20	7	5	132	1.2	1.0	113	0	20
	b)	20	7	5	132	1.2	1.0	113	0	1
	c)	20	7	5	132	1.2	1.0	113	0	2
16	—	22	129	25	890	363 714.3	653.7	675	—	0
	a)	22	77	8	557	216.6	1.3	675	10.4	22
	b)	22	108	6	335	569.0	1.1	675	36.0	1
	c)	22	108	8	541	569.0	1.3	675	13.4	4
64	—	24	1 148	111	5 908	674 804.3	2 000.9	3 143	—	0
	a)	24	109	8	1 465	243.1	1.3	3 143	65.9	24
	b)	24	163	6	777	2 740.5	1.1	3 143	168.0	1
	c)	24	113	8	1 483	433.0	1.3	3 143	68.7	5
256	—	26	3 417	352	18 764	696 950.1	5 083.7	13 455	—	0
	a)	26	136	8	2 406	247.8	1.3	13 455	325.9	26
	b)	26	141	7	1 086	5 413.9	1.3	13 455	720.0	1
	c)	26	206	8	2 397	5 413.9	1.3	13 455	389.0	4

Table 1 Numerical results for model problem (9); each subdomain is a union of $2 \times 28 \times 28$ linear triangular finite elements; tolerance TOL= 1000 for adaptive coarse space; **N**: number of subdomains; **Strategy**: strategy chosen for cond_func,"—" denotes the case without an adaptive coarse space; **Max. / Min. Krylov It.**: maximal / minimal number of Krylov subspace iterations during the Newton iteration; **Total Krylov It.**: total number of Krylov subspace iterations during the Newton iteration; **Max. / Min. cond.**: maximal / minimal condition number during the Newton iteration; **Interface d.o.f.**: degrees of freedom on the interface; **Avg. size U**: average size of the adaptive coarse spaces during Newton iteration; **EP Solves**: Number of Newton steps in which a new adaptive coarse space is computed.

comparing Newton-Krylov-FETI-DP with Adaptive-Newton-Krylov-FETI-DP. We always make all subdomain vertex values primal initially. In all computations we use a moderate tolerance TOL= 1000 to keep our adaptive coarse spaces small. All three adaptive strategies reduce the number of CG iterations drastically in comparison to classical Newton-Krylov-FETI-DP. Using **Strategy a)** and computing a new coarse space in each Newton step, the condition number stays below the theoretical bound $C \cdot \mathrm{TOL}$. The coarse spaces generated are sufficiently small with a size of less than 5% of the size of the interface. Using **Strategy b)**, the number of CG iterations is even lower. This is caused by a comparably large coarse space computed in the first Newton step. In this approach, the adaptive coarse space has only to be computed once, which results in a large reduction of local computational work compared to **Strategy a)**. Unfortunately, the number of CG iterations in the different Newton steps and the average size of the coarse space strongly differs from the theoretically verified **Strategy a)** and thus a control using tolerance TOL is no longer possible. In contrast, **Strategy c)** can nearly reproduce the average size of the coarse space and the number of CG iterations of **Strategy a)**. Additionally, the number of adaptive coarse space computations and thus the number of local eigenvalue problems is reduced by a factor of 5.0 to 6.0. For a graphical comparison of all methods see also Fig. 3. Especially the similar behavior of **Strategies a)** and **c)** can be observed.

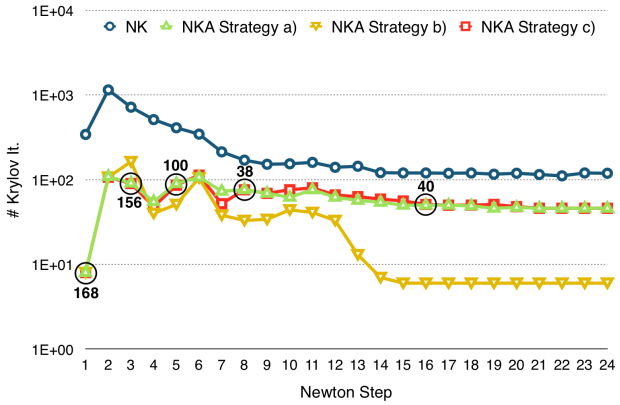

Fig. 3 Results for 64 subdomains from Table 1 showing the number of Krylov subspace iterations in each Newton step; **NK** (blue curve) denotes Newton-Krylov-FETI-DP without adaptive coarse spaces; **NKA Strategy a) / b) / c)** (green / yellow / red curve) denotes Strategy a) / b) / c); the five black circles mark the Newton steps in which **Strategy c)** decided to compute a new coarse space and the numbers give the sizes of the coarse spaces.

5 Conclusion

An adaptive Newton-Krylov-FETI-DP approach has been presented, where the condition numbers of all preconditioned tangential matrices are bounded by a constant. Additionally, heuristic strategies have been introduced saving local work by reducing the number of eigenvalue problems. Results for a p-Laplace model problem with highly heterogeneous coefficient have been presented, showing the ability of adaptive coarse spaces to save CG iterations.

Acknowledgement This work was supported in part by the German Research Foundation (DFG) through the Priority Programme 1648 "Software for Exascale Computing" (**SPPEXA**) under KL 2094/4-1, KL 2094/4-2, RH 122/2-1, RH 122/3-2 .

References

[1] Charbel Farhat, Michel Lesoinne, Patrick LeTallec, Kendall Pierson, and Daniel 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.*, 50:1523–1544, 2001.

[2] Pierre Gosselet, Christian Rey, and Julien Pebrel. Total and selective reuse of Krylov subspaces for the resolution of sequences of nonlinear structural problems. *Internat. J. Numer. Methods Engrg.*, 94(1):60–83, 2013.

[3] Hyea Hyun Kim and Eric T. Chung. A BDDC algorithm with enriched coarse spaces for two-dimensional elliptic problems with oscillatory and high contrast coefficients. *Multiscale Model. Simul.*, 13(2):571–593, 2015.

[4] Axel Klawonn, Martin Kühn, and Oliver Rheinbach. Adaptive Coarse Spaces for FETI-DP in Three Dimensions. *SIAM J. Sci. Comput.*, 38(5):A2880–A2911, 2016. Preprint 2015-11 available at `http://tu-freiberg.de/sites/default/files/media/fakultaet-fuer-mathematik-und-informatik-fakultaet-1-9277/prep/2015-11_fertig.pdf`.

[5] Axel Klawonn, Patrick Radtke, and Oliver Rheinbach. A Newton-Krylov-FETI-DP Method with an Adaptive Coarse Space Applied to Elastoplasticity. In *Domain Decomposition Methods in Science and Engineering XXII*, volume 104, pages 293–300. Springer LNCSE, 2016.

[6] Axel Klawonn, Patrick Radtke, and Oliver Rheinbach. A comparison of adaptive coarse spaces for iterative substructuring in two dimensions. *Electronic Transactions on Numerical Analysis (ETNA)*, 45:75–106, March 2016.

[7] Axel Klawonn and Oliver Rheinbach. Robust FETI-DP methods for heterogeneous three dimensional elasticity problems. *Comput. Methods Appl. Mech. Engrg.*, 196(8):1400–1414, 2007.

[8] Axel Klawonn and Oliver Rheinbach. Highly scalable parallel domain decomposition methods with an application to biomechanics. *ZAMM Z. Angew. Math. Mech.*, 90(1):5–32, 2010.

[9] Axel Klawonn and Oliver Rheinbach. Deflation, projector preconditioning, and balancing in iterative substructuring methods: connections and new results. *SIAM J. Sci. Comput.*, 34(1):A459–A484, 2012.

[10] Axel Klawonn and Olof B. Widlund. Dual-Primal FETI Methods for Linear Elasticity. *Comm. Pure Appl. Math.*, 59(11):1523–1572, 2006.

[11] Jan Mandel and Bedřich 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(8):1389–1399, 2007.

[12] Jan Mandel, Bedřich Sousedík, and Jakub Šístek. Adaptive BDDC in three dimensions. *Math. Comput. Simulation*, 82(10):1812–1831, 2012.

[13] Bedřich Sousedík, Jakub Šístek, and Jan Mandel. Adaptive-multilevel BDDC and its parallel implementation. *Computing*, 95:1087–1119, 2013.

[14] Nicole Spillane. Adaptive Multi Preconditioned Conjugate Gradient: Algorithm, Theory and an Application to Domain Decomposition. *`https://hal.archives-ouvertes.fr/hal-01170059`*, 2015.

[15] Nicole Spillane and Daniel J. Rixen. Automatic spectral coarse spaces for robust finite element tearing and interconnecting and balanced domain decomposition algorithms. *Internat. J. Numer. Methods Engrg.*, 95(11):953–990, 2013.