Efficient Adaptive Elimination Strategies in Nonlinear FETI-DP Methods in Combination with Adaptive Spectral Coarse Spaces

Axel Klawonn and Martin Lanser

1 Introduction

Nonlinear domain decomposition methods (DDMs) are based on a decomposition of a discretized nonlinear partial differential equation instead of applying a linear DDM to the tangential systems in a Newton-type iteration. The advantages are a faster convergence and an improved ratio of local work to communication, at least for many problems. We focus on nonlinear FETI-DP (Finite Element Tearing and Interconnecting - Dual Primal) methods here, which build a class of nonlinear two-level approaches. These methods can have a (partially) nonlinear coarse level and the integrated nonlinear right-preconditioner is based on a partial elimination of arbitrary degrees of freedom collected in an index set E. In [2], it was shown that the combination of nonlinear FETI-DP with an adaptive coarse space [9], which was implemented with a transformation of basis approach, improves the convergence. Additionally, in [6] the concept of choosing an index set E adaptively based on the residual was investigated. Finally, in [3], both ideas are combined to a nonlinear FETI-DP algorithm iterating in the transformed space, which we abbreviate with NL-FETI-DP-XT here. Additionally, also in [3], an efficient implementation iterating in the original finite element space is suggested using local saddle point problems [8] instead of an explicit transformation of basis; this method is abbreviated by NL-FETI-DP-X here. For the latter approach, modifications have to be made to the elimination set E. We will compare different strategies to modify E and finally suggest and numerically test a completely new and more efficient and robust strategy for NL-FETI-DP-X based on an approximation of the transformed residual.

Axel Klawonn, Martin Lanser

Department of Mathematics and Computer Science, Division of Mathematics, University of Cologne, Weyertal 86-90, 50931 Cologne, Germany, e-mail: axel.klawonn@uni-koeln.de, martin.lanser@uni-koeln.de

2 Nonlinear FETI-DP

Let us briefly recall the unified framework of nonlinear FETI-DP methods. For
a detailed description, we refer to [5]. Throughout this article, we assume that we
have a computational domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, which is divided into N non-
overlapping subdomains Ω_i, i.e., $\Omega = \bigcup_{i=1}^{N} \Omega_i$. Each subdomain is the union of finite
elements and the associated finite element spaces are denoted by $W^{(i)}$. We denote
the product space of all finite element spaces as $W = W^{(1)} \times \cdots \times W^{(N)}$. In FETI-
DP methods, we partition all variables into interior (I), dual (Δ), and primal (Π)
variables, where only continuity in the primal variables is prescribed and continuity
in the dual variables is enforced by Lagrange multipliers λ iteratively. Therefore,
we further introduce a subspace $\widetilde{W} \subset W$ of all finite element functions from W
that are continuous in the primal variables. A simple choice of primal variables are
subdomain vertices. More advanced strategies are based on enforcing continuity in
certain weighted averages over the degrees of freedoms of an edge or face. The
weights can, for example, be computed adaptively by solving localized eigenvalue
problems related to edges. This approach results in provably robust linear FETI-DP
methods; see, e.g., [7, 9]. For nonlinear FETI-DP methods, the adaptive coarse space
can be computed using the tangential matrix linearized in the initial value; see [2].
We use this specific adaptive coarse space in all computations in this article.

For completeness, we also introduce the subspace $\widehat{W} \subset W$, which contains all
finite element functions that are continuous across the complete interface and it
holds $\widehat{W} \subset \widetilde{W} \subset W$. Let us introduce the primal assembly operator $\check{R}^T : W \to \widetilde{W}$
and the nonlinear function $K(u) : W \to W$ obtained by a finite element discretization
of a nonlinear partial differential equation. Let us note that $K(u)$ is not necessarily
continuous on the interface.

As it was shown in [4], finding the solution of the fully assembled finite element
problem is equivalent to solving the nonlinear FETI-DP saddle point system

$$A(\tilde{u}, \lambda) = \begin{bmatrix} \widetilde{K}(\tilde{u}) + \check{R}^T B^T \lambda - \check{R}^T f \\ B\check{R}\tilde{u} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tilde{u} \in \widetilde{W}, \ \widetilde{K}(\tilde{u}) := \check{R}^T K(\check{R}\tilde{u}) \in \widetilde{W}. \quad (1)$$

This system is the basis for all nonlinear FETI-DP methods. Here, the linear con-
straints $B\check{R}\tilde{u} = 0$ together with Lagrange multipliers $\lambda \in V := \text{range}(B)$ enforce
continuity in all dual variables.

To implement arbitrary coarse constraints, as, e.g., adaptive constraints, one can
use a transformation of basis approach. The underlying idea is to transform the
complete system into a space W_T, where all primal constraints are again point-
wise constraints and can be enforced by a simple assembly operator as before.
A transformation matrix $T : W_T \to W$ with orthonormal rows, that is, with $T^T T = I$,
can be computed for all coarse spaces based on edge and face averages; see [3] for
details. Then, the transformed nonlinear FETI-DP saddle point system writes

$$A_T(\tilde{u}_T, \lambda) = \begin{bmatrix} \widetilde{K}_T(\tilde{u}_T) + \check{R}^T T^T B^T \lambda - \check{R}^T T^T f \\ BT\check{R}\tilde{u}_T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \widetilde{K}_T(\tilde{u}_T) := \check{R}^T T^T K(T\check{R}\tilde{u}).$$

(2)

As introduced in [5], we use a nonlinear right-preconditioner $M_T(\tilde{u}_T, \lambda)$ that is nonlinear in $\tilde{u}_T$ and linear in λ; see [5] for some desirable properties of M_T. Instead of $A_T(\tilde{u}_T, \lambda) = 0$, we now solve $A_T(M_T(\tilde{u}_T, \lambda)) = 0$ with a Newton-Krylov method. Following [5], the application of a nonlinear right-preconditioner can be interpreted as a (partial) nonlinear elimination process, where different choices of M_T lead to different elimination sets. With this interpretation, it is obvious to divide the overall set of variables into two different subsets E and L, where E contains all variables that should be nonlinearly eliminated by the preconditioner M_T, and L contains the remaining variables in which will be linearized.

After an appropriate rearrangement, we can split Eq. (2) according to the subsets E and L. We can write the nonlinear saddle point system (Eq. (2)) as

$$A_T(\tilde{u}_{T,E}, \tilde{u}_{T,L}, \lambda) = \begin{bmatrix} A_{T,E}(\tilde{u}_{T,E}, \tilde{u}_{T,L}, \lambda) \\ A_{T,L}(\tilde{u}_{T,E}, \tilde{u}_{T,L}, \lambda) \\ BT\check{R}\tilde{u}_T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

With the application of the nonlinear right-preconditioner, we now aim to eliminate all variables $\tilde{u}_E$, which correspond to the subset E. Thus, our preconditioner is implicitly defined by solving the nonlinear equation

$$A_{T,E}(M_{T,\tilde{u}_{T,E}}(\tilde{u}_{T,L}, \lambda), \tilde{u}_{T,L}, \lambda) = 0,$$

(3)

where we have $M_T(\tilde{u}_{T,E}, \tilde{u}_{T,L}, \lambda) := (M_{T,\tilde{u}_{T,E}}(\tilde{u}_L, \lambda), \tilde{u}_L, \lambda)$, since, by construction, M_T is linear in $\tilde{u}_{T,L}$ and λ. After we have computed the nonlinear preconditioner M_T by solving Eq. (3) with Newton's method, we obtain the nonlinear Schur complement system

$$\begin{bmatrix} A_{T,L}(M_{T,\tilde{u}_{T,E}}(\tilde{u}_{T,L}, \lambda), \tilde{u}_{T,L}, \lambda) \\ BT\check{R}\tilde{u}_T \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This can be solved with the traditional Newton-Krylov-FETI-DP approach; see [5]. Putting it all together, in each of these (outer) Newton iterations, M_T has to be recomputed, which is typically done by an inner Newton iteration.

Both Newton loops iterate in the transformed space, that is, all outer Newton updates $\delta\tilde{u}_T$ and inner Newton updates $\delta\tilde{u}_{T,E}$ have to be projected back to the original space after convergence. As in linear FETI-DP methods, the explicit usage of T leads to denser linear systems and thus, especially in three dimensions using rich coarse spaces, to a higher memory demand and a slower time to solution. As in the linear case, it is possible to reformulate nonlinear FETI-DP in the original nodal space using local saddle point systems and some further tricks; see [3] for details. It is possible to get rid of the matrix T in all computations of nonlinear FETI-DP without changing the nonlinear and linear convergence. Unfortunately, the additional

assumption has to be made that T has the block structure

$$T = \begin{bmatrix} T_E & 0 \\ 0 & T_L \end{bmatrix};$$

see [3] for details. To enforce this, all primal edges or faces, i.e., edges or faces with at least one primal constraint, have to be either included in E or L completely. In contrast, iterating in the transformed space, E can be chosen arbitrarily. In this article, we discuss different strategies of how to choose an appropriate E adaptively and compare nonlinear FETI-DP in the iterating in the nodal space (NL-FETI-DP-X) with nonlinear FETI-DP iterating in the transformed space (NL-FETI-DP-XT).

3 Adaptive selection of E

NL-FETI-DP-XT allows for completely arbitrary elimination sets E. For the more efficient NL-FETI-DP-X we will fulfill the necessary assumption on T by either choosing an edge to be part of E or, respectively, L completely, i.e., we will not split any edge. More precisely, in theory, it is sufficient to not split faces or edges which carry primal constraints. The adaptive selection of E used in this article is a modification of the procedure suggested in [6]. This heuristic strategy is based on the assembled nonlinear residual and is inspired by [1].

We first define the nonlinear residual in the k-th outer Newton iteration

$$r^{(k)} := K(u^{(k)}) + B^T \lambda^{(k)} - f = K(T \check{R} \tilde{u}_T^{(k)}) + B^T \lambda^{(k)} - f$$

and the assembled and transformed residual by

$$r_T^{(k)} := R^T T^T r^{(k)},$$

where $R^T : W \to \widehat{W}$ assembles all degrees of freedom on the interface. We now eliminate all variables, where the residual is comparably high, that is, if for the i-th component $r_{T,i}^{(k)}$ of $r_T^{(k)}$ the inequality

$$r_{T,i}^{(k)} \geq \rho_E ||r_T^{(k)}||_\infty$$

holds, the i-th degree of freedom is eliminated. That means, the index i is added to the elimination set E. Let us remark that we only describe the scalar case here. A procedure for systems of equations with more degrees of freedom in each physical node, e.g., elasticity problems, is discussed in [6] and out of the scope of this article. Here, $\rho_E < 1$ is a user defined parameter and smaller values immediately result in larger elimination sets E. To avoid single and isolated physical points in the elimination set, δ_E layers of finite element nodes surrounding E are added to the preselected set E. This is comparable to the procedure of selecting an overlap of a nonoverlapping subdomain. The resulting E can immediately be used within NL-FETI-DP-XT

and we denote this procedure to find E by **basic strategy**. Nonetheless, the resulting E can not be used in NL-FETI-DP-X, where we are not allowed to split edges. We suggest two different strategies to overcome this issue.

Strategy 1: After choosing E with the basic algorithm, all edges which have a nonempty intersection with E are added to E completely. Around these edges, δ_E layers of finite element nodes belonging to the interior of the adjacent subdomains are also added to E. We refer to Fig. 1 for a visualization of this strategy.

A disadvantage of strategy 1 and the basic approach is the need for computing T which is used to compute $r_T^{(k)}$. We will introduce a third strategy avoiding T, since, in the efficient implementation NL-FETI-DP-X, T is not necessary at all. Only the rows of T^T belonging to primal variables are usually known. Sorting the variables properly, we have

$$T^T = \begin{bmatrix} T_{\Pi\Pi}^T & T_{\Pi B}^T \\ T_{B\Pi}^T & T_{BB}^T \end{bmatrix}$$

and only the block $C := \begin{bmatrix} T_{\Pi\Pi}^T & T_{\Pi B}^T \end{bmatrix}$ is available. Let us note that $T_{\widehat{B},:}^T = [I\ 0]$, where $\widehat{B} := [I\ \widehat{\Delta}]$ and $\widehat{\Delta}$ is the index set of all dual variables belonging to edges which do not carry primal constraints. Therefore, we have

$$r_{T,\widehat{B}}^{(k)} = \left. \left(R^T r^{(k)} \right) \right|_{\widehat{B}}$$

which can be computed without knowing T. For the computation of the primal part $r_{T,\Pi}^{(k)}$ solely C is necessary. Only the part related to the dual part of the primal edges $r_{T,\Delta\setminus\widehat{\Delta}}^{(k)}$ cannot be computed without using T^T. In NL-FETI-DP-X, all edges carrying primal constraints have to be either completely part of E or, respectively, L. Assuming to have an appropriate coarse space and that all important information about the primal edges are transformed to the coarse space and thus to the vector $r_{T,\Pi}^{(k)}$, we can rely on $r_{T,\Pi}^{(k)}$ for the decision if an edge, which is part of $\Delta \setminus \widehat{\Delta}$, is chosen to be part of E or L. We therefore suggest the following modification.

Strategy 2: Choose the initial set E applying the basic algorithm to $\bar{r}_T^{(k)} := [r_{T,\widehat{B}}^{(k)T}, r_{T,\Pi}^{(k)T}, 0_{\Delta\setminus\widehat{\Delta}}]$ instead of $r_T^{(k)}$. Then proceed as in Strategy 1.

4 Problem and numerical results

We consider the nonlinear problem

$$\begin{aligned} -\alpha\Delta_p u &= 1 && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{4}$$

with the scaled p-Laplace operator $\alpha\Delta_p u := \mathrm{div}(\alpha|\nabla u|^{p-2}\nabla u)$. Within this article, we use $p = 4$ and a coefficient function $\alpha \colon \Omega \to \mathbb{R}$ with jumps. Moreover, we always
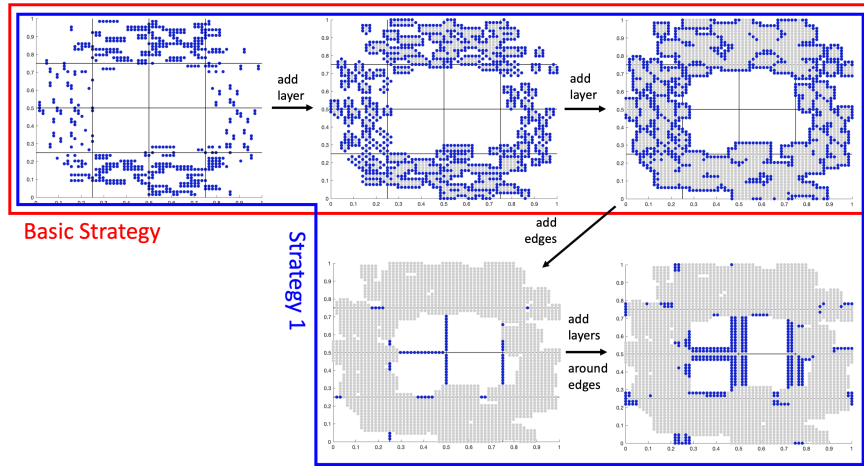
Fig. 1 Illustration of Strategy 1 to compute the elimination set E with $\delta_E = 2$. The starting point or initial set E is obtained based on the residual. In Strategy 1, first two layers are added, then all necessary edges are included in E, and finally an overlap of δ_E layers is added in the interior of the subdomains adjacent to those edges. The result of the basic algorithm can solely be used in NL-FETI-DP-XT.
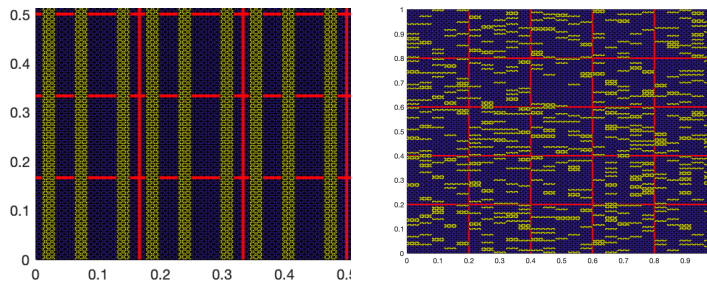

Fig. 2 Coefficient distributions and domain decompositions used in the numerical computations. **Left:** Channels with $\alpha = 1e3$ crossing material with $\alpha = 1$, zoomed in to a quarter of the unit square. **Right:** Randomly generated distribution with $\alpha = 1e6$ in the small yellow stripes.

use the unit square $\Omega = [0, 1] \times [0, 1]$ as the computational domain, a discretization with piecewise linear finite elements, and a structured domain decomposition into square subdomains. We consider two different coefficient distributions, which can be found in Fig. 2. We always choose $u^{(0)}(x, y) = x(1 - x)y(1 - y)$ as initial value in all computations.

For all edges we compute the adaptive coarse constraints introduced in [9] using the first linearized system and a tolerance of $tol = 10$ for the localized eigenvalue problems. For linear problems, there is a provable condition number bound of $N_{\mathcal{E}}^2 \cdot tol$ for FETI-DP using this coarse space, where $N_{\mathcal{E}}$ is the maximum number of edges

Table 1 Results for model problems with randomly generated coefficients and channels; always using vertices plus adaptive edge constraints; **outer it.** gives the total number of global Newton iterations and in brackets the number of Newton-Krylov-FETI-DP steps used for stability is shown; **inner it.** gives the number of inner Newton iterations summed up over the outer Newton iterations; **PCG it.** gives the number of PCG iterations summed up over the outer Newton iterations; $|E_{avg}|$ gives the average size of the elimination set in percentage of the number of degrees of freedom; **NL-FETI-DP-X** and **NL-FETI-DP-XT** stand for the adaptive selection of the elimination set; **NL-FETI-DP-2** stands for eliminating all variables, i.e., $E = [B \ \Pi]$; **NK-FETI-DP** stands for Newton-Krylov-FETI-DP. The best results are marked in bold.

_p_-**Laplace random; see Fig. 2 (right)**							
$p = 4$; $H/h = 25$; 25 subdomains; $tol = 10$							
	Stra-				**outer**	**inner**	**PCG**
method	**tegy**	δ_E	ρ_E	$|E|_{avg}$	**it.**	**it.**	**it. (sum)**
NL-FETI-DP-2	-	-	-	37.5%	8(5)	57	142
NK-FETI-DP	-	-	-	-	15(15)	-	284
NL-FETI-DP-XT	basic	2	0.01	6.9%	13(10)	36	254
NL-FETI-DP-XT	basic	5	0.01	33.3%	**4(0)**	**53**	81
NL-FETI-DP-X	1	2	0.01	11.1%	11(2)	74	200
NL-FETI-DP-X	1	5	0.01	41.1%	**4(0)**	55	78
NL-FETI-DP-X	2	2	0.01	8.3%	13(6)	68	237
NL-FETI-DP-X	2	5	0.01	38.7%	**4(0)**	**53**	**77**
_p_-**Laplace channels; see Fig. 2 (left)**							
$p = 4$; $H/h = 32$; 36 subdomains; $tol = 10$							
	Stra-				**outer**	**inner**	**PCG**
method	**tegy**	δ_E	ρ_E	$|E|_{avg}$	**it.**	**it.**	**it. (sum)**
NL-FETI-DP-2	-	-	-	71.4%	7(2)	43	80
NK-FETI-DP	-	-	-	-	19(19)	-	237
NL-FETI-DP-XT	basic	2	0.01	5.2%	16(12)	61	197
NL-FETI-DP-XT	basic	5	0.01	7.5%	14(9)	101	180
NL-FETI-DP-X	1	2	0.01	20.2%	6(0)	53	72
NL-FETI-DP-X	1	5	0.01	43.8%	**4(0)**	**45**	**54**
NL-FETI-DP-X	2	2	0.01	19.9%	6(0)	53	73
NL-FETI-DP-X	2	5	0.01	43.8%	**4(0)**	**45**	**54**

a subdomain can have; see [7] for the proof. In our computations, the outer Newton iteration is stopped if a relative reduction of 10^{-5} of the globally assembled residual is reached. The inner iteration is stopped, if the inner Newton update is smaller than 10^{-5} in the l_2-norm. Let us finally note that, for stability reasons, we will always switch to a Newton-Krylov-FETI-DP approach, if no further reduction of the residual is reached in the outer loop. We never switch back to nonlinear FETI-DP.

We always compute the average size of the elimination set E to give a rough estimate on the computational cost of the elimination process in the inner loop. We therefore compute the value

$$|E|_{avg} := \frac{1}{N_o} \sum_{k=1}^{N_o} \frac{|E^{(k)}|}{N_{dof}} \cdot 100\%,$$

where $|E^{(k)}|$ is the number of degrees of freedom in the elimination set of the k-th outer iteration, N_{dof} is the number of total degrees of freedom, and N_o the number of outer iterations. Let us remark that $|E^{(k)}| = 0$ for each Newton-Krylov iteration and thus $|E|_{avg}$ can be small if many Newton-Krylov steps have to be made.

The results for both model problems can be found in Table 1. It can be observed that NL-FETI-DP-X can compete with NL-FETI-DP-2 in terms of nonlinear and linear convergence; at least if appropriate elimination sets are chosen. Let us remark that NL-FETI-DP-X has in all setups less than 44% of the local computational cost. Additionally, both approaches outperform classical NK-FETI-DP. Strategies 1 and 2 have only been introduced in order to implement the theoretical need for edges not being split up in the efficient implementation of NL-FETI-DP-X. Nonetheless, splitting edges, which happens often in the basic strategy used in NL-FETI-DP-XT, actually deteriorates the performance, which was not expected. The most efficient Strategy 2, which does not need T explicitly, is competitive to Strategy 1 and therefore it is our suggestion to use this approach in NL-FETI-DP-X. Of course more tests and also three dimensional problems have to be investigated in the future.

References

1. Gong, S. and Cai, X.-C. A nonlinear elimination preconditioned inexact Newton method for heterogeneous hyperelasticity. *SIAM Journal on Scientific Computing* **41**(5), S390–S408 (2019).
2. Heinlein, A., Klawonn, A., and Lanser, M. Adaptive nonlinear domain decomposition methods with an application to the p-Laplacian. *SIAM Journal on Scientific Computing* **0**(0), S152–S172 (0).
3. Klawonn, A. and Lanser, M. Nonlinear FETI-DP domain decomposition methods: On the efficient implementation of arbitrary coarse spaces and nonlinear elimination sets (2022). In preparation.
4. Klawonn, A., Lanser, M., and Rheinbach, O. Nonlinear FETI-DP and BDDC methods. *SIAM J. Sci. Comput.* **36**(2), A737–A765 (2014).
5. Klawonn, A., Lanser, M., Rheinbach, O., and Uran, M. Nonlinear FETI-DP and BDDC methods: a unified framework and parallel results. *SIAM J. Sci. Comput.* **39**(6), C417–C451 (2017).
6. Klawonn, A., Lanser, M., and Uran, M. Adaptive nonlinear elimination in nonlinear FETI-DP methods. Preprint, Universität zu Köln (2021).
7. Klawonn, A., Radtke, P., and Rheinbach, O. A comparison of adaptive coarse spaces for iterative substructuring in two dimensions. *Electron. Trans. Numer. Anal.* **45**, 75–106 (2016).
8. Klawonn, A. and Widlund, O. B. Dual-Primal FETI methods for linear elasticity. *Communications on Pure and Applied Mathematics* **59**(11), 1523–1572 (2006).
9. Mandel, J. and Sousedík, B. 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).