Technical Report Series
Center for Data and Simulation Science

Alexander Heinlein, Martin Lanser

Additive and Hybrid Nonlinear Two-Level Schwarz Methods and Energy Minimizing Coarse Spaces for Unstructured Grids

Technical Report ID: CDS-2019-17
Available at https://kups.ub.uni-koeln.de/id/eprint/9845
Submitted on July 23, 2019

ADDITIVE AND HYBRID NONLINEAR TWO-LEVEL SCHWARZ METHODS AND ENERGY MINIMIZING COARSE SPACES FOR UNSTRUCTURED GRIDS*

ALEXANDER HEINLEIN[†‡] AND MARTIN LANSER[†‡]

Abstract. Nonlinear domain decomposition (DD) methods, such as, e.g., ASPIN (Additive Schwarz Preconditioned Inexact Newton), RASPEN (Restricted Additive Schwarz Preconditioned Inexact Newton), Nonlinear-FETI-DP, or Nonlinear-BDDC methods, can be reasonable alternatives to classical Newton-Krylov-DD methods for the solution of sparse nonlinear systems of equations, e.g., arising from a discretization of a nonlinear partial differential equation. These nonlinear DD approaches are often able to effectively tackle unevenly distributed nonlinearities and outperform Newton's method with respect to convergence speed as well as global convergence behavior. Furthermore, they often improve parallel scalability due to a superior ratio of local to global work.

Nonetheless, as for linear DD methods, it is often necessary to incorporate an appropriate coarse space in a second level to obtain numerical scalability for increasing numbers of subdomains. In addition to that, an appropriate coarse space can also improve the nonlinear convergence of nonlinear DD methods.

In this paper, four variants how to integrate coarse spaces in nonlinear Schwarz methods in an additive or multiplicative way using Galerkin projections are introduced. These new variants can be interpreted as natural nonlinear equivalents to well-known linear additive and hybrid two-level Schwarz preconditioners. Furthermore, they facilitate the use of various coarse spaces, e.g., coarse spaces based on energy-minimizing extensions, which can easily be used for irregular domain decompositions, as, e.g., obtained by graph partitioners. In particular, Multiscale Finite Element Method (MsFEM) type coarse spaces are considered, and it is shown that they outperform classical approaches for certain heterogeneous nonlinear problems.

The new approaches are then compared with classical Newton-Krylov-DD and nonlinear one-level Schwarz approaches for different homogeneous and heterogeneous model problems based on the p-Laplace operator.

Key words. Nonlinear preconditioning, inexact Newton methods, nonlinear Schwarz methods, nonlinear domain decomposition, multiscale coarse spaces, ASPIN, RASPEN

AMS subject classifications. 65F08, 65F10, 65H10, 65H20, 65N12, 65N22, 65N30, 65N55

1. Introduction.

We are concerned with the effective solution of nonlinear systems of equations using nonlinear domain decomposition (DD) methods. These nonlinear systems arise, e.g., by finite element discretization of the variational formulation of a nonlinear partial differential equation. Let therefore V be a finite element space discretizing a computational domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, and

$$F(u) = 0, \tag{1}$$

a certain nonlinear system of equations given by the nonlinear function $F : V \to V$, with the solution $u \in V$. In this paper, for the sake of clarity, we will restrict ourselves to the two-dimensional case, however our approaches can be easily extended to three dimensions as well.

Besides classical Newton-Krylov-DD methods, where the nonlinear system (1) is linearized with Newton's method and the tangential system is solved with a conjugate

*Preliminary ideas on one of four new nonlinear two-level Schwarz methods together with first Matlab experiments have been already presented in a proceedings paper; see [30]. The present paper significantly extends those preliminary ideas and also contains other new algorithms and numerical results.

[†]Department of Mathematics and Computer Science, University of Cologne, Weyertal 86-90, 50931 Köln, Germany, {alexander.heinlein, martin.lanser}@uni-koeln.de, url: http://www.numerik.uni-koeln.de

[‡]Center for Data and Simulation Science, University of Cologne, http://www.cds.uni-koeln.de

gradient (CG) or generalized minimal residual method (GMRES) approach preconditioned by some DD preconditioner, in recent years, nonlinear domain decomposition methods became popular. These approaches often yield faster convergence, especially if the initial value is outside the area of quadratic convergence of Newton's method. In nonlinear DD methods, the computational domain Ω is decomposed into nonoverlapping or overlapping subdomains. A corresponding decomposition of the nonlinear problem is then used to construct nonlinear left- or right-preconditioners. In contrast, to linear DD preconditioners or methods, which improve only the convergence of the linear solvers, nonlinear DD preconditioners can also positively affect the nonlinear convergence.

Nonlinear right-preconditioners are often associated with a nonlinear elimination procedure, as, e.g., described in [9]. Many different variants have been developed in the last two decades leading to different nonlinear DD methods, such as nonlinear FETI (Finite Element Tearing and Interconnecting) [46, 47, 31], Nonlinear-FETI-DP (Finite Element Tearing and Interconnecting - Dual Primal) [41, 39, 38, 37], Nonlinear BDDC (Balancing Domain Decomposition by Constraints) [37, 40], or nonlinear elimination preconditioned inexact Newton [33, 35]. In these approaches, the sets of variables which are eliminated in a nonlinear fashion are either chosen based on a nonoverlapping domain decomposition [46, 47, 31, 41, 39, 38, 37, 40] or problem dependent using a heuristic approach, e.g., based on the nonlinear residual [33, 35].

In this paper, we consider nonlinear left-preconditioners based on a nonoverlapping domain decomposition, i.e., nonlinear Schwarz preconditioners. While nonlinear Schwarz methods as iterative approaches have already been developed and analyzed in [4, 17, 44, 50], a nonlinear Schwarz preconditioner was first suggested in [5] and the resulting method was called ASPIN (Additive Schwarz Preconditioned Inexact Newton). In [5], also the corresponding exact approach was derived, which can be denoted by ASPEN (Additive Schwarz Preconditioned Exact Newton). Both approaches often yield superior nonlinear convergence compared to classical Newton-Krylov-DD approaches. This is also investigated numerically in [2], where ASPIN is compared with various combinations of nonlinear solvers designed for a fast nonlinear convergence. Both, ASPIN and ASPEN, were introduced as one-level methods and the numerical scalability of the preconditioned linear systems for an increasing number of subdomains is therefore generally not ensured. Thus, several approaches to implement a second level have been proposed: an additive nonlinear coarse problem based on a coarser mesh [6, 45], an additive linear coarse problem [34], and a multiplicative variant using an FAS (full approximation scheme) update [15]. In the latter publication, also a restricted variant of ASPEN, called RASPEN (Restricted Additive Schwarz Preconditioned Exact Newton), is introduced; restricted Schwarz methods typically improve the linear convergence compared to standard overlapping Schwarz methods.

In this paper, we introduce four different additively and multiplicatively coupled two-level (R)ASPEN and (R)ASPIN methods based on Galerkin projections instead of an FAS update. The multiplicative coupling between coarse space and local corrections is comparable to the MSPIN approach [43]. One of our four approaches was already discussed in [30]. The inexact tangential systems of our methods are naturally equivalent to linear systems preconditioned with the well-known additive or hybrid Schwarz preconditioners described in, e.g., [51]; therefore, the methods can be interpreted as the natural nonlinear equivalents to classical linear two-level Schwarz methods. We combine these methods with coarse spaces designed for irregular domain decompositions as provided by graph partitioners, as, e.g., METIS [36], and numerically prove a superior nonlinear as well as linear convergence behavior compared to

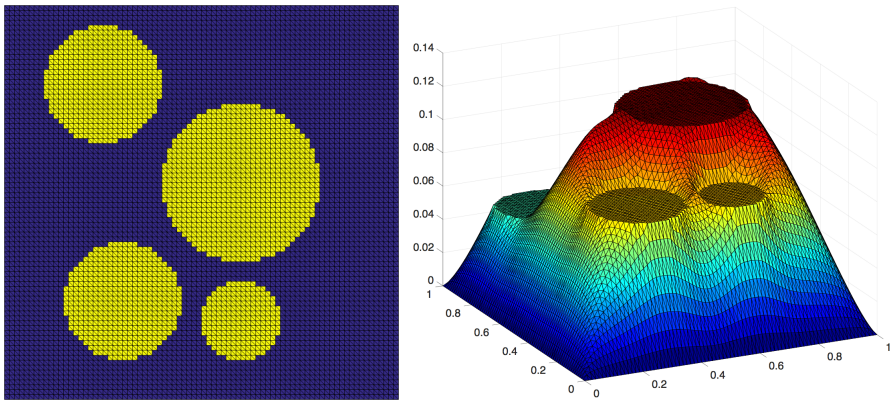

FIG. 1. **Left:** *Circular inclusions; coefficient distribution with* $\alpha = 10^3$ *or* $\alpha = 10^6$ *in the yellow part and* $\alpha = 1$ *in the remaining blue part.* **Right:** *Solution of the corresponding heterogeneous model problem with a high coefficient of* $\alpha = 10^6$.

one-level (R)ASPIN or (R)ASPEN methods and the corresponding Newton-Krylov-Schwarz approaches. Let us remark that these coarse spaces can theoretically also be used with FAS-RASPEN, but in this case, the coarse basis functions have to be provided in the dual space; see subsection 5.2 for more details. This is not necessary in our approaches. Despite of this, our nonlinear two-level algorithms and FAS-RASPEN share similar building blocks, which have similar computational costs.

2. Model problems. Our model problems are based on the scaled p-Laplace operator, which is defined by

$$\alpha\Delta_p u := \operatorname{div}(\alpha|\nabla u|^{p-2}\nabla u)$$

for $p \geq 2$. We then consider the model problem: find $u \in H_0^1(\Omega)$, such that

$$\begin{aligned} -\alpha\Delta_p u &= 1 && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{2}$$

where $\alpha : \Omega \to \mathbb{R}$ is a coefficient function.

Here, we always consider the case $p = 4$ and two different coefficient distributions: first, a **homogeneous model problem**, i.e., $\alpha(x) = 1$ for all $x \in \Omega$, and second, a **heterogeneous model problem** with a high coefficient in four circles and $\alpha(x) = 1$ in the remaining area; see Figure 1 (left) for the coefficient distribution and Figure 1 (right) for the corresponding solution. We present numerical results in section 8.

3. Nonlinear one-level Schwarz methods. Let us first describe the one-level nonlinear Schwarz method introduced in [5], which is also the basis for our new methods. The inexact variant is well-known under the name ASPIN and a restricted reformulation named RASPEN was first introduced in [15]. In both approaches, the nonlinear problem (1) is reformulated into an equivalent nonlinear problem $\mathcal{F}(u) = 0$ before Newton's method is applied. The reformulation is obtained from the solution of many local nonlinear problems on parts of Ω and can be interpreted as a nonlinearly left-preconditioned function $\mathcal{F}(u) = G(F(u))$. Here, G is only known implicitly.

We define a decomposition of Ω into nonoverlapping subdomains Ω_i, $i = 1, ..., N$, where each subdomain is the union of finite elements. By adding k layers of finite

elements to the subdomains, overlapping subdomains Ω_i', $i = 1, ..., N$ with overlap $\delta = kh$ are constructed, where h is the diameter of a finite element. We denote by V_i the local finite element spaces associated to the overlapping subdomains Ω_i'. With restrictions $R_i : V \to V_i$ and prolongations $P_i : V_i \to V$, we can define local nonlinear corrections $T_i(u)$ as the solutions of the local problems

$$R_i F(u - P_i T_i(u)) = 0, \ i = 1, ..., N. \tag{3}$$

In [5], it is shown that the nonlinear equation

$$\mathcal{F}_A(u) := \mathcal{F}_{\text{ASPEN}}(u) := \sum_{i=1}^{N} P_i T_i(u) = 0 \tag{4}$$

has the same solution as (1). Solving (4) using Newton's method, i.e., with the iteration

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_A(u^{(k)}) \right)^{-1} \mathcal{F}_A \left(u^{(k)} \right),$$

yields the ASPEN approach. The derivative of $\mathcal{F}_A(u)$ writes

$$D\mathcal{F}_A(u) = \sum_{i=1}^{N} P_i DT_i(u) = \sum_{i=1}^{N} P_i \left(R_i DF(u_i) P_i \right)^{-1} R_i DF(u_i), \tag{5}$$

with $u_i = u - P_i T_i(u)$ and $DT_i(u)$ obtained by deriving (3). Replacing u_i by u in (5) yields the inexact tangent of the ASPIN approach. Consequently, the tangent of the ASPIN approach is equal to $M_{\text{OS}}^{-1} DF(u)$, where M_{OS}^{-1} is the linear one-level additive Schwarz preconditioner applied for the tangent $DF(u)$.

Let us note that for both evaluations, $D\mathcal{F}_A(u^{(k)})$ and $\mathcal{F}_A(u^{(k)})$, the local nonlinear problems defined in (3) have to be solved in each Newton iteration. This is usually done by an inner Newton iteration carried out independently on each overlapping subdomain. In the remainder of this paper, we call Newton iterations on the subdomains *inner iterations* and denote global Newton iterations by the *outer iteration*. For an algorithmic description of ASPIN or ASPEN, see, e.g., [5, 15] or section 7.

Alternatively, one can also use the restricted Schwarz approach described in [7] to construct a nonlinear preconditioner; cf. [15]. We therefore define restricted prolongation operators $\widetilde{P}_i : V_i \to V$, $i = 1, ..., N$, such that

$$\sum_{i=1}^{N} \widetilde{P}_i R_i = I$$

is fulfilled; this means that the prolongation operators form a partition of unity on Ω. With the equation

$$\mathcal{F}_{RA}(u) := \mathcal{F}_{\text{RASPEN}}(u) := \sum_{i=1}^{N} \widetilde{P}_i T_i(u) = 0, \tag{6}$$

linearization with Newton's method leads to the RASPEN method

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{RA}(u^{(k)}) \right)^{-1} \mathcal{F}_{RA} \left(u^{(k)} \right),$$

with the derivative

$$D\mathcal{F}_{RA}(u) = \sum_{i=1}^{N} \widetilde{P}_i DT_i(u) = \sum_{i=1}^{N} \widetilde{P}_i \left(R_i DF(u_i) P_i\right)^{-1} R_i DF(u_i). \tag{7}$$

Therefore, ASPEN and RASPEN only differ by the prolongation from the subdomains to the complete domain and thus by the combination of the nonlinear correction terms $T_i(u)$; the local correction terms themselves will be identical if they are computed for the same function u. Moreover, the inexact tangent of the nonlinear RASPIN method is equivalent to the restricted two-level Schwarz preconditioned linear system.

4. Nonlinear two-level Schwarz methods. In this section, we describe our approach to add a nonlinear coarse level to ASPEN or RASPEN. Our approach differs from the two-level variants described in [15, 34, 6, 45] and we will discuss the differences later, in section 5. In this section, we assume that a restriction $R_0 : V \to V_0$ to a given coarse space V_0 and a corresponding prolongation $P_0 : V_0 \to V$ are given. Throughout this article, we always use $P_0 := R_0^T$. Coarse spaces for linear Schwarz methods are typically given by a finite element discretization on an additional coarse triangulation or by constructing coarse basis functions exploiting the domain decomposition. In section 6, we will discuss the construction of coarse spaces for nonlinear Schwarz methods and unstructured domain decompositions without the need for an additional coarse triangulation.

Based on R_0 and P_0, let us define the nonlinear coarse correction $T_0(u)$ as the solution of the nonlinear equation

$$R_0 F(u - P_0 T_0(u)) = 0 \tag{8}$$

for a given $u \in V$. Therefore, we exclusively consider nonlinear coarse functions $F_0 : V_0 \to V_0$ which can be defined by a Galerkin approach as $F_0(u_0) := R_0 F(P_0 u_0)$ for any $u_0 \in V_0$. In this paper, we will consider four different approaches to add a coarse level to the ASPEN or RASPEN method:

- in an additive fashion,
- in a multiplicative fashion **before** the nonlinear subdomain corrections are applied,
- in a multiplicative fashion **after** the nonlinear subdomain corrections are applied,
- or in a multiplicative fashion **before and after** the nonlinear subdomain corrections are applied.

Let us remark that we already discussed the third variant in [30] and presented some preliminary numerical results. All four variants lead to different nonlinear problems and therefore also different linearized systems. Consequently, both, the nonlinear and the linear convergence behavior, may differ; cf. section 8. We will now give a detailed description of the four approaches based on the ASPEN framework.

4.1. Additive coupling. We first define the additive nonlinear two-level operator

$$\mathcal{F}_{\text{add}}(u) := \mathcal{F}_{\text{additive}}(u) := \sum_{i=1}^{N} P_i T_i(u) + P_0 T_0(u). \tag{9}$$

A linearization of $\mathcal{F}_{\text{add}}(u) = 0$ using Newton's method leads to an additive two-level ASPEN method

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\text{add}}(u^{(k)})\right)^{-1} \mathcal{F}_{\text{add}}\left(u^{(k)}\right),$$

with the derivative

$$
\begin{aligned}
D\mathcal{F}_{\text{add}}(u) = &\sum_{i=1}^{N} P_i DT_i(u) + P_0 DT_0(u) \\
= &\sum_{i=1}^{N} P_i \left(R_i DF(u_i) P_i\right)^{-1} R_i DF(u_i) \\
&+ P_0 \left(R_0 DF(u_0) P_0\right)^{-1} R_0 DF(u_0),
\end{aligned} \tag{10}
$$

where $u_i = u - P_i T_i(u),\ i = 0, ..., N$, as before. The derivative $DT_0(u)$ is obtained by deriving (8).

In accordance to linear Schwarz operators, as defined, e.g., in [51], we introduce nonlinear Schwarz operators

$$
Q_i(u) := P_i \left(R_i\, DF(u)\, P_i\right)^{-1} R_i\, DF(u), \tag{11}
$$

such that (10) can be rewritten as

$$
D\mathcal{F}_{\text{add}}(u) = \sum_{i=0}^{N} Q_i(u_i). \tag{12}
$$

We refer to this method as A-ASPEN, where "A" indicates the additive coupling of the levels.

4.2. Multiplicative coupling – coarse problem first. The second nonlinear two-level operator is defined by

$$
\mathcal{F}_{\text{h},1}(u) := \mathcal{F}_{\text{hybrid},1}(u) := \sum_{i=1}^{N} P_i T_i(u - P_0 T_0(u)) + P_0 T_0(u). \tag{13}
$$

Applying Newton's method to the corresponding nonlinear equation $\mathcal{F}_{\text{h},1}(u) = 0$ yields a hybrid two-level ASPEN method

$$
u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\text{h},1}(u^{(k)})\right)^{-1} \mathcal{F}_{\text{h},1}\left(u^{(k)}\right),
$$

with the derivative

$$
\begin{aligned}
D\mathcal{F}_{\text{h},1}(u) = &\sum_{i=1}^{N} P_i DT_i\left(u - P_0 T_0(u)\right)\left(I - P_0 DT_0(u)\right) + P_0 DT_0(u) \\
= &\ I - \left(I - \sum_{i=1}^{N} Q_i(v_i)\right)\left(I - Q_0(u_0)\right).
\end{aligned} \tag{14}
$$

Here, we have $u_0 = u - P_0 T_0(u)$, $v_i = u_0 - P_i T_i(u_0)$, and Q_i as defined in (11).

We refer to this method as H1-ASPEN, where "H1" indicates a multiplicative coupling where the coarse operator is applied before the local corrections.

4.3. Multiplicative coupling – coarse problem second. We define another nonlinear two-level operator with multiplicative coupling, where the coarse problem is solved after the local corrections have been applied. It is defined by

$$
\mathcal{F}_{\text{h},2}(u) := \mathcal{F}_{\text{hybrid},2}(u) := \sum_{i=1}^{N} P_i T_i(u) + P_0 T_0(u - \sum_{i=1}^{N} P_i T_i(u)). \tag{15}
$$

The corresponding Newton iteration for solving $\mathcal{F}_{\mathrm{h},2}(u) = 0$ is given by

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\mathrm{h},2}(u^{(k)})\right)^{-1} \mathcal{F}_{h,2}\left(u^{(k)}\right),$$

with the derivative

$$\begin{aligned} D\mathcal{F}_{\mathrm{h},2}(u) = & \sum_{i=1}^{N} P_i DT_i\,(u) \\ & + P_0 DT_0\left(u - \textstyle\sum_{i=1}^{N} P_i T_i(u)\right)\left(I - \textstyle\sum_{i=1}^{N} P_i DT_i(u)\right) \\ = & \; I - (I - Q_0(v_0))\left(I - \sum_{i=1}^{N} Q_i(u_i)\right), \end{aligned} \tag{16}$$

where $u_i = u - P_i T_i(u)$ and $v_0 = u - \sum_{i=1}^{N} P_i T_i(u) - P_0 T_0(u - \sum_{i=1}^{N} P_i T_i(u))$, and Q_i as defined in (11).

We refer to this method as H2-ASPEN, where "H2" indicates a multiplicative coupling where the coarse operator is applied after the local corrections.

4.4. Multiplicative coupling – symmetric variant. In order to simplify the notation, we first define $v := u - P_0 T_0(u)$ and $w := v - \sum_{i=1}^{N} P_i T_i(v)$. Then, the symmetric nonlinear two-level operator is then defined by

$$\mathcal{F}_{\mathrm{h}}(u) = \mathcal{F}_{\mathrm{hybrid}}(u) := P_0 T_0(w) + \sum_{i=1}^{N} P_i T_i(v) + P_0 T_0(u). \tag{17}$$

We obtain the Newton iteration of the symmetric hybrid two-level ASPEN method

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\mathrm{h}}(u^{(k)})\right)^{-1} \mathcal{F}_{\mathrm{h}}\left(u^{(k)}\right)$$

for the solution of $\mathcal{F}_{\mathrm{h}}(u) = 0$. The derivatives of v and w with respect to u are given by

$$Dv = I - P_0 DT_0(u) = I - Q_0(u_0)$$

and

$$\begin{aligned} Dw = & \; I - P_0 DT_0(u) - \sum_{i=1}^{N} P_i DT_i(v)\,(I - P_0 DT_0(u)) \\ = & \left(I - \sum_{i=1}^{N} Q_i(v_i)\right)(I - Q_0(u_0)), \end{aligned}$$

with $u_0 = u - P_0 T_0(u)$, and $v_i = v - P_i T_i(v), i = 1, ..., N$. With this in mind, we can derive the jacobian

$$\begin{aligned} D\mathcal{F}_{\mathrm{h}}(u) = & \; P_0 DT_0(w)\, Dw + \sum_{i=1}^{N} P_i DT_i(v)\, Dv + P_0 DT_0(u) \\ = & \; I - (I - Q_0(w_0))\left(I - \sum_{i=1}^{N} Q_i(v_i)\right)(I - Q_0(u_0)). \end{aligned} \tag{18}$$

Here, we have again $u_0 = u - P_0 T_0(u)$ and $v_i = v - P_i T_i(v), i = 1, ..., N$ as well as $w_0 = w - P_0 T_0(w)$; Q_i is defined in (11).

We refer to this method as H-ASPEN, where "H" indicates a multiplicative coupling where the coarse operator is applied twice, i.e., before and after the local corrections. Let us finally remark that the computational cost for each Newton step of H-RASPEN is slightly higher, since the nonlinear coarse problem has to be solved twice.

4.5. Inexact and restricted variants. All four approaches can be extended to corresponding inexact or restricted variants in a straight-forward way.

In particular, when using u as the linearization point for all nonlinear Schwarz operators, we obtain ASPIN variants of our nonlinear two-level Schwarz methods

$$\begin{aligned}
D\mathcal{F}_{\mathrm{add}}(u) &\approx \sum_{i=0}^{N} Q_i(u), \\
D\mathcal{F}_{\mathrm{h},1}(u) &\approx I - \left(I - \sum_{i=1}^{N} Q_i(u) \right) (I - Q_0(u)), \\
D\mathcal{F}_{\mathrm{h},2}(u) &\approx I - (I - Q_0(u)) \left(I - \sum_{i=1}^{N} Q_i(u) \right), \text{ and} \\
D\mathcal{F}_{\mathrm{h}}(u) &\approx I - (I - Q_0(u)) \left(I - \sum_{i=1}^{N} Q_i(u) \right) (I - Q_0(u)).
\end{aligned}$$

In particular, these variants are equivalent to applying the corresponding additive or hybrid linear two-level Schwarz preconditioners to the tangent $DF(u)$.

Furthermore, using the restricted prolongation operators $\widetilde{P}_i$ instead of P_i, $i = 1, ..., N$, to add the local corrections in the nonlinear Schwarz preconditioners leads directly to the corresponding RASPEN or RASPIN variants; cf. section 3.

5. Differences to existing two-level methods. In the literature, two different existing nonlinear two-level Schwarz preconditioners can be found. First, an additive variant is described in [45, 6] and second, a multiplicative approach is chosen in [15]. Additionally, in [34], a linear second level is introduced, which we do not consider here.

5.1. Additive nonlinear coarse space. In this subsection, we describe the additive approach to implement a nonlinear coarse problem chosen in [45, 6] and discuss the similarities and differences to our approach. First, we assume that we have a nonlinear coarse problem

$$F_0(u_0^*) = 0 \tag{19}$$

with the unique solution u_0^* and $F_0 : V_0 \to V_0$. The nonlinear function F_0 can be obtained by a Galerkin approach as before or, e.g., by a coarser finite element discretization of the nonlinear partial differential equation.

Finally, in [45, 6], the coarse correction $C_0(u)$ is implicitly defined by the equation

$$F_0(C_0(u)) = R_0 F(u) \tag{20}$$

and the nonlinear operator by

$$\mathcal{F}_{\mathrm{CKM}}(u) := P_0 C_0(u) - P_0 C_0(u^*) + \sum_{i=1}^{N} P_i T_i(u). \tag{21}$$

Here, u^* is the solution of the original nonlinear problem (1). Let us note that with (20) we have $C_0(u^*) = u_0^*$, which can be obtained by solving (19). A linearization using Newton's method leads to

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\mathrm{CKM}}(u^{(k)}) \right)^{-1} \mathcal{F}_{\mathrm{CKM}} \left(u^{(k)} \right),$$

with the derivative

(22)
$$\begin{aligned} D\mathcal{F}_{\mathrm{CKM}}(u) &= \sum_{i=1}^{N} P_i DT_i(u) + P_0 DC_0(u) \\ &= \sum_{i=1}^{N} P_i \left(R_i DF(u_i) P_i\right)^{-1} R_i DF(u_i) + P_0 \left(DF_0(C_0(u))\right)^{-1} R_0 DF(u). \end{aligned}$$

The derivative $DC_0(u)$ is thereby obtained by a derivation of (20) and we have $u_i = u - P_i T_i(u),\ i = 0, ..., N$, as before.

In contrast to our approaches, also coarse functions F_0 which are not obtained by a Galerkin approach can be used. Therefore, this approach is more general in that sense. Especially considering a coarse problem arising from a discretization on a coarse triangulation can save compute time, since the assembly of the problem is less costly. Nonetheless, we intend to use nonlinear coarse problems which can also be used for unstructured meshes and domain decompositions; see section 6. Therefore, we construct coarse spaces based on a Galerkin approach. A disadvantage of the approach from [45, 6] is that (19) has to be solved in advance to compute $C_0(u^*)$, which is not necessary in our approaches. This is cheap for coarse spaces based on a coarse triangulation, but costly for coarse spaces based on a Galerkin approach. A variant, where $P_0 C_0(u^*)$ is used as initial value for Newton's iteration is beneficial, but in principal applicable to all nonlinear solvers.

5.2. Multiplicative nonlinear coarse space. Another approach to implement a coarse space in a multiplicative fashion is presented in [15]. The authors chose to use an FAS update and the method is thus denoted by FAS-(R)ASPEN. Let us therefore define a scaled restriction operator $R_{0,D}$, which plays the same role as R_0, but in the residual or dual spaces of V and, respectively, V_0. The nonlinear coarse function F_0 can be defined as before. The FAS coarse correction $\widetilde{C}(u)$ for u is then defined by

$$F_0(\widetilde{C}_0(u) + R_{0,D}u) = F_0(R_{0,D}u) - R_0 F(u); \tag{23}$$

see [15] for details. The FAS correction is applied multiplicatively before the local corrections are computed, which is similar to H1-(R)ASPEN described above. The nonlinear operator is thus defined by

$$\mathcal{F}_{\mathrm{FAS}}(u) := P_0 \widetilde{C}_0(u) + \sum_{i=1}^{N} P_i T_i(u + P_0 \widetilde{C}_0(u)). \tag{24}$$

A linearization of $\mathcal{F}_{\mathrm{FAS}}(u) = 0$ with Newton's method leads to

$$u^{(k+1)} = u^{(k)} - \left(D\mathcal{F}_{\mathrm{FAS}}(u^{(k)})\right)^{-1} \mathcal{F}_{\mathrm{FAS}}\left(u^{(k)}\right),$$

with the derivative

$$D\mathcal{F}_{\mathrm{FAS}}(u) = P_0 D\widetilde{C}_0(u) + \sum_{i=1}^{N} P_i DT_i(u + P_0 \widetilde{C}_0(u)) \left(I + P_0 D\widetilde{C}_0(u)\right). \tag{25}$$

Here, we have

$$\sum_{i=1}^{N} P_i DT_i(u + P_0 \widetilde{C}_0(u)) = \sum_{i=1}^{N} P_i \left(R_i DF(u_{0,i}) P_i\right)^{-1} R_i DF(u_{0,i})$$

with $u_{0,i} = u + P_0\widetilde{C}_0(u) + T_i(u)$ and

$$D\widetilde{C}_0(u) = -R_{0,D} + \Big(DF_0(R_{0,D}u + \widetilde{C}_0(u))\Big)^{-1}\Big(DF_0(R_{0,D}u)R_{0,D} - R_0DF(u)\Big).$$

Here, the derivative of the coarse correction is obtained by deriving (23). This method can be denoted by FAS-ASPEN and replacing P_i, $i = 1, ..., N$, by $\widetilde{P}_i$ in (24) leads to FAS-RASPEN as introduced in [15].

Though similar to H1-RASPEN and in contrast to the two-level methods defined in section 4, in FAS-RASPEN the scaled restriction operator $R_{0,D}$ has to be defined, which has, in our experience, a large impact on the convergence of the method.

Additionally, in our framework, one can vary the ordering of coarse and local corrections.

6. Coarse spaces for unstructured grids. An essential ingredient of two-level Schwarz methods is a suitable coarse space V_0 that enables fast global transport of information and therefore numerical scalability. If a coarse triangulation is available, classical Lagrangian basis functions are a natural choice for the coarse basis; see, e.g., [51] for linear Schwarz preconditioning. However, for many realistic applications, where unstructured grids and domain decompositions are used, a coarse triangulation is typically not available and, in addition, difficult to obtain. On the other hand, heterogeneous problems might require an additional treatment of the heterogeneities by the coarse space; see, e.g., [1, 23, 3] for multiscale coarse spaces and, e.g., [21, 49, 16, 22, 28, 18, 26, 27] for adaptive coarse spaces.

Several coarse spaces for linear Schwarz methods have been proposed which can be constructed based on unstructured domain decompositions, without the need for an additional coarse triangulation; see, e.g., [14, 11, 10, 12, 13, 48, 3, 16, 22, 28, 18, 26, 27, 24, 25]. Most of those approaches make use of energy-minimizing extensions based on the differential operator of the PDE. Here, we will construct a coarse space of MsFEM (Multiscale Finite Element Method) [32] type, i.e., a coarse space spanned by energy-minimizing nodal basis functions, and use it to compute the nonlinear coarse correction (8). Therefore, our coarse spaces are also related to, e.g., the approaches in [11, 3, 14, 8].

In order to obtain a scalable two-level Schwarz preconditioner for linear problems, it is necessary that the coarse space is able to represent the null space of the global Neumann operator on all subdomains which do not touch the Dirichlet boundary; cf. [51]. The construction of our coarse spaces is also guided by this principle. The property of representing the null space can be achieved by constructing a coarse basis that forms a partition of unity on the corresponding subdomains and multiplying it with the null space. For simplicity, we restrict ourselves to scalar PDEs where the null space consists only of constant functions. Therefore, the partition of unity already gives us a scalable coarse space.

We obtain the partition of unity by first constructing a corresponding partition of unity on the interior interface

$$\Gamma' := \bigcup_{\bar{\Omega}_i \cap \partial\Omega_D = \emptyset} \partial\Omega_i$$

of the nonoverlapping domain decomposition $\{\Omega_i\}_{i=1,...,N}$, i.e., on the boundary of all subdomains which do not touch the Dirichlet boundary $\partial\Omega_D$. Then, we extend this interface partition of unity to the interior of the nonoverlapping subdomains resulting in a partition of unity on all subdomains which do not touch the Dirichlet boundary.

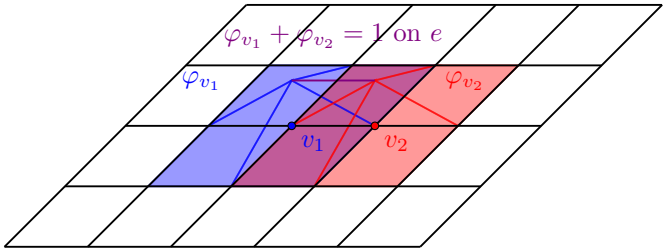

FIG. 2. *Graphical representation of the interface partition of unity properties (26) and (29) for the domain Ω decomposed into 5×5 subdomains. The nodal values are chosen based on the Kronecker delta property and edge values are chosen to sum to 1; here, the values are linear on the edges.*

First, let the domain decomposition interface

$$\Gamma := \bigcup_{i=1,\dots,N} \partial\Omega_i$$

be decomposed into edges and vertices and $\mathcal{E}$ and $\mathcal{V}$ be the sets of all edges and vertices, respectively. Now, for each $\nu \in \mathcal{V}$, we construct a function η_ν defined on $\bigcup_{i=1,\dots,N} \partial\Omega_i$, such that

$$\begin{aligned} \eta_\nu(\nu') &= \delta_{\nu,\nu'} && \forall \nu, \nu' \in \mathcal{V}, \\ \sum_{\nu \in \mathcal{V}} \eta_\nu &= 1 && \text{on } \Gamma', \text{ and} \\ \eta_\nu &= 0 && \text{on } \partial\Omega_D. \end{aligned} \tag{26}$$

where $\delta_{\nu,\nu'}$ is the Kronecker delta symbol.

Then, we compute extensions of these functions to the interior. Therefore, we rely on an energy semi-norm $|\cdot|_{a,\Omega}$, which is induced by a bilinear form $a(\cdot,\cdot)$, with

$$a(v,v) = |v|^2_{a,\Omega} = 0 \Leftrightarrow v \equiv c,$$

with $c \in \mathbb{R}$. Here, we first explain the construction of the basis functions for a generic semi-norm $|\cdot|_{a,\Omega}$. In subsection 6.1, we propose two specific choices, which are also used in our numerical experiments in section 8; other choices for $|\cdot|_{a,\Omega}$ are possible as well.

The coarse basis function φ_ν corresponding to the vertex $\nu \in \mathcal{V}$ is computed by solving the energy-minimization problem

$$\varphi_\nu = \operatorname*{arg\,min}_{\substack{v \in V \\ v|_\Gamma = \eta_\nu \text{ on } \Gamma}} |v|^2_{a,\Omega}\,. \tag{27}$$

Note that this is equivalent to solving N independent problems

$$\varphi_\nu|_{\bar{\Omega}_i} = \operatorname*{arg\,min}_{\substack{v_i \in V(\Omega_i) \\ v_i|_{\partial\Omega_{i,I}} = \eta_\nu \text{ on } \partial\Omega_{i,I}}} |v_i|^2_{a,\Omega_i}\,, \tag{28}$$

where $\partial\Omega_{i,I} := \partial\Omega_i \setminus \partial\Omega$, for $i = 1, \dots, N$, and $V(\Omega_i)$ is the local finite element space associated with Ω_i. In order to obtain basis functions with local support, we require

that

$$\eta_\nu|_e = 0 \quad \forall e \in \mathcal{E} \text{ with } \bar{e} \cap \nu = \emptyset \tag{29}$$

such that (27) is equivalent to solving (28) for only those subdomains Ω_i with $v \cap \bar{\Omega}_i \neq \emptyset$; on all other subdomains φ_ν will vanish. We will introduce two choices for the interface values in subsection 6.2. With that, the computation of each coarse basis functions only requires the solution of a few local energy-minimization problems and can therefore be performed efficiently in parallel; cf., e.g., [29] for the parallel implementation of a similar coarse space.

6.1. Energy semi-norms. In order to compute φ_ν from η_ν, we have to solve local minimization problems (28). Constructing an energy-minimizing extension corresponds to solving the linear problem: find $\varphi_\nu \in \{w \in V : w|_\Gamma = \eta_\nu\}$, such that

$$a_{\Omega_i}(\varphi_\nu, v_i) = 0 \qquad \forall v_i \in V(\Omega_i)$$

for all $i = 1, ..., N$. Therefore, we partition all degrees of freedom into interior (I) and (Γ) interface degrees of freedom, such that the finite element matrix corresponding to $a(\cdot,\cdot)$ can be written as

$$\begin{bmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{bmatrix}.$$

Then, the extensions can be computed by

$$\varphi_\nu = \begin{bmatrix} -A_{II}^{-1} A_{I\Gamma} \\ I \end{bmatrix} \eta_\nu. \tag{30}$$

In our nonlinear setting, the bilinear form corresponding to the current linearization changes in each Newton iteration. Instead of recomputing the energy-minimizing extension in each Newton iteration, we select a constant coarse basis, computed before the Newton iteration. In particular, we propose two different discrete semi-norms resulting from linear problems related to our nonlinear problem (2).

Linear Laplacian energy. As a first approach, we assemble the finite element matrix corresponding to the linear Laplacian, i.e., to the bilinear form

$$a(u, v) = \int_\Omega \nabla u \, \nabla v \, dx;$$

see also (2) with $p = 2$. These extensions are certainly not energy minimizing for (2) with $p > 2$. However, since the null spaces are equal for $p = 2$ and $p > 2$, we obtain a reasonable partition of unity.

One draw-back of this approach for general nonlinear problems is that it is based on the availability of a suitable linear surrogate problem. In the next paragraph, we suggest a bilinear form that is derived from (1) and already available without additional computational work.

Energy of the first Newton linearization. In a second approach, we use the tangential matrix

$$DF(u^{(0)})$$

from our first Newton iteration, i.e., evaluated in our initial guess $u^{(0)}$, to compute the extensions. This is advantageous because the corresponding matrix is always assembled within the Newton iteration. Furthermore, no linear surrogate problem has

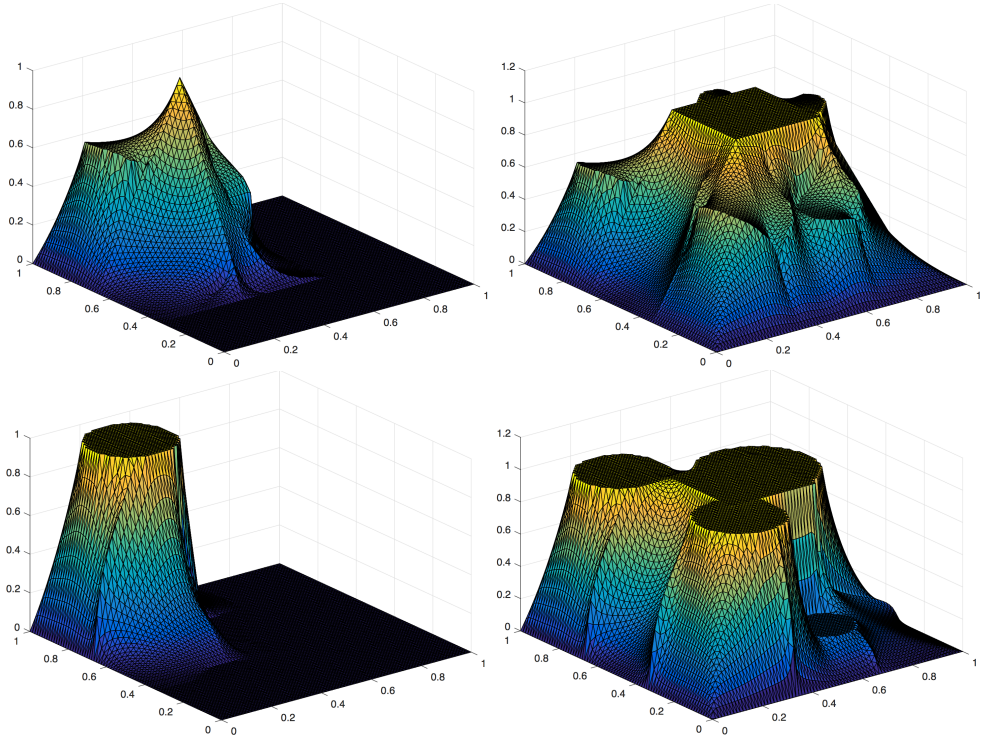

FIG. 3. *Domain decomposition into 9 subdomains with coefficient distribution depicted in Figure 1 with* $\alpha = 10^6$. **Top Left:** *MsFEM-D basis function in one vertex;* **Top Right:** *Sum of all four MsFEM-D basis functions;* **Bottom Left:** *MsFEM-E basis function in one vertex;* **Bottom Right:** *Sum of all four MsFEM-E basis functions.*

to be found or designed in order to compute the extensions. Therefore, this approach is more general compared to the previous one. However, the energy-minimality property is, in general, not fulfilled for later Newton iterations. Let us remark that the extensions could be updated using $DF(u^{(k)})$ in certain Newton steps if the heterogeneities or nonlinearities changed drastically during the Newton iteration. For our model problems, this was not necessary.

In both approaches, we compute the extensions by (30). Since the matrix A_{II} is block-diagonal, the extensions can be computed concurrently for all subdomains; also cf. (28).

6.2. Interface values. The interface properties (26) and (29) are fulfilled for many classical nodal discretizations; cf. Figure 2 for a graphical representation for rectangular subdomains and piecewise bilinear (Q1) basis functions. For heterogeneous model problems (e.g., Figure 1), multi scale basis functions can significantly differ from Lagrangian basis functions while maintaining these properties; cf. Figure 3.

For unstructured grids and domain decompositions, we propose the following two definitions of interface values; since the vertex values are determined by the properties (26), it is only left to define the edge values.

Distance based edge values. As a first option, we consider the edge values introduced in [30]. Therefore, let ν be a vertex, e an adjacent edge, and ν' the other vertex

adjacent to the edge e. Then, we set the interface values to

$$\eta_\nu(x) = \frac{||x-\nu||}{||x-\nu'|| + ||x-\nu||}$$

for any $x \in e$; cf., e.g., [30]. Based on this formula, we compute the values η_ν on each edge adjacent to ν. On all edges which are not adjacent to ν, we set η_ν to zero. We denote this variant as MsFEM-D, where "D" indicates that the computation of the edge values is based on distances.

It is easy to see that this choice is not optimal for heterogeneous problem. On straight edges, the values are linear, independent of variations in the coefficient function; see Figure 3. This is improved in our second variant.

Energy minimal edge values. In contrast to the previous variant, where the edge values are based on distances to the adjacent vertices, we also consider a variant which incorporates the coefficient function. In particular, we propose to use the vertex basis functions of the OS-ACMS coarse space introduced in [28]. The OS-ACMS coarse space is an adaptive coarse space for overlapping Schwarz methods consisting of energy-minimizing basis functions. In particular, the extensions into the interior are computed in the same way as described in subsection 6.1. In addition to the vertex basis functions the OS-ACMS coarse space also contains edge basis functions, which are constructed from the solutions of local generalized eigenvalue problems. However, it turns out that the vertex basis functions alone are already robust for many heterogeneous coefficient distributions; cf., e.g., the heterogeneous problem in Figure 1.

In order to compute the edge values of the OS-ACMS coarse vertex basis function η_ν, let

$$\Omega_e = \overline{\Omega_{e,1} \cup \Omega_{e,2}}$$

where e is an edge and $\Omega_{e,1}$ and $\Omega_{e,2}$ are the two subdomains adjacent to e. Then, we compute functions

$$v_{\nu,e} = \operatorname*{arg\,min}_{\substack{v \in V(\Omega_e) \\ v(\nu') = \delta_{\nu\nu'}}} |v|^2_{a,\Omega_e}$$

for all edges e and define the edge values of η_ν as the edge values of the $v_{\nu,e}$, i.e.,

$$\eta_\nu|_e = v_{\nu,e}|_e$$

We denote this variant as MsFEM-E, where "E" indicates that the computation of the edge values is based on energy minimization.

Obviously, the second variant is more expensive than the first one, however, we can observe that the edge values account for variations in the coefficient functions; see Figure 3. As a consequence, the MsFEM-E coarse space is more robust (in the sense of linear and nonlinear convergence) for heterogeneous problems; see subsection 8.2.

7. Algorithmic description. In this section, we additionally provide an algorithmic point of view of the different methods. We concentrate on the ASPEN variants, since RASPEN, ASPIN, or RASPIN share the same algorithmic structure. We also provide a comparison with classical Newton-Krylov-Schwarz approaches; see Algorithm 1.

Algorithm 1 Newton-Krylov-Schwarz

Init $u^{(0)}$
Compute $F(u^{(0)})$
Loop on $k = 0, 1, 2, ...$ until $||F(u^{(k)})||/||F(u^{(0)})|| < tol_{\text{outer}}$
  Compute $DF(u^{(k)})$
  Solve iteratively $M^{-1}DF(u^{(k)})\delta u^{(k)} = M^{-1}F(u^{(k)})$
  /* Here, M^{-1} can be any one- or two-level Schwarz preconditioner; We use GMRES to solve the system up to a certain tolerance $\Rightarrow$ inexact Newton method*/
  Update $u^{(k+1)} = u^{(k)} - \lambda\,\delta u^{(k)}$
  /*λ is the step-length computed, e.g, by a linesearch approach /*
  Compute $F(u^{(k+1)})$
EndLoop

The usage of different linear preconditioners M^{-1} in Algorithm 1 changes the linear but not the nonlinear convergence behavior. We consider different linear one- and two-level Schwarz preconditioners corresponding to their nonlinear relatives described above, e.g., hybrid or additive Schwarz methods with $\mathbb{P}1$, MsFEM-D, or MsFEM-E coarse spaces.

7.1. The ASPEN algorithm. Let now $\mathcal{F}_X$ be a nonlinear function corresponding to an arbitrary one- or two-level ASPEN method, e.g., $\mathcal{F}_X := \mathcal{F}_{\text{add}}$ or $\mathcal{F}_X := \mathcal{F}_{\text{h},2}$. The ASPEN method then writes as presented in Algorithm 2.

Algorithm 2 ASPEN

Init $u^{(0)}$
Compute $F(u^{(0)})$
Loop on $k = 0, 1, 2, ...$ until $||F(u^{(k)})||/||F(u^{(0)})|| < tol_{\text{outer}}$
  Compute $g^{(k)} := \mathcal{F}_X(u^{(k)})$
  Compute $D\mathcal{F}_X(u^{(k)})$
  Solve iteratively $D\mathcal{F}_X(u^{(k)})\delta u^{(k)} = g^{(k)}$
  /* Here, $D\mathcal{F}_X(u^{(k)})$ has already the structure of a Schwarz-preconditioned operator; depending on the nonlinear ASPEN variant; We use GMRES to solve the system up to a certain tolerance $\Rightarrow$ inexact Newton methods*/
  Update $u^{(k+1)} = u^{(k)} - \lambda\,\delta u^{(k)}$
  /*λ is the step-length computed, e.g, by a linesearch approach /*
  Compute $F(u^{(k+1)})$
EndLoop

Let us remark that we can also formulate a stopping criterion based on $\mathcal{F}_X$ instead of F. We choose the latter option for a better comparability between the different methods and use the tolerance $tol_{\text{outer}} = 10^{-6}$ in all computations in section 8. In contrast to Newton-Krylov approaches, the evaluation of $\mathcal{F}_X(u^{(k)})$ requires the solution of nonlinear problems, more precisely, the computation of $T_i(u)$, $i = 0, ..., N$. Regardless if a local correction ($i > 0$) or the coarse correction ($i = 0$) has to be computed, this is done with Newton's method and we use $tol_{\text{inner/coarse}} = 10^{-3}$ in our computations; see Algorithm 3.

Algorithm 3 Compute correction $g_i := T_i(u^{(k)})$

Init $g_i^{(0)} = 0$
Compute $F_i^{(0)} := R_i F(u^{(k)} - P_i g_i^{(0)})$
Loop on $l = 0, 1, 2, ...$ until $||F_i^{(l)}||/||F_i^{(0)}|| < tol_{\text{inner/coarse}}$
  Compute $DF_i^{(l)} := R_i DF(u^{(k)} - P_i g_i^{(l)}) P_i$
  Solve directly $DF_i^{(l)} \delta g_i^{(l)} = F_i^{(l)}$
  /* Sparse direct solvers are used here ⇒ exact Newton method*/
  Update $g_i^{(l+1)} = g_i^{(l)} + \lambda \, \delta g_i^{(l)}$
  /*λ is the step-length computed, e.g, by a linesearch approach /*
  Compute $F_i^{(l+1)} := R_i F(u^{(k)} - P_i g_i^{(l+1)})$
EndLoop
Set $g_i := g_i^{(l+1)}$

7.2. Evaluation of nonlinear functions. Now, we have stated all ingredients to specify the evaluation of $\mathcal{F}_X(u^{(k)})$; see Algorithms 4 to 8. The main difference of the competing methods, i.e., the ordering of local and coarse corrections, can be easily observed in the algorithmic notation.

Algorithm 4 Evaluation of one-level ASPEN function $g^{(k)} := \mathcal{F}_A(u^{(k)})$

Compute local corrections $g_i := T_i(u^{(k)}),\ i = 1, ..., N$ /*with Algorithm 3*/
Set $g^{(k)} := \sum_{i=1}^{N} P_i g_i$

Algorithm 5 Evaluation of two-level additive ASPEN function $g^{(k)} := \mathcal{F}_{\text{add}}(u^{(k)})$

Compute local corrections $g_i := T_i(u^{(k)}),\ i = 1, ..., N$ /*with Algorithm 3*/
Compute coarse correction $g_0 := T_0(u^{(k)})$ /*with Algorithm 3*/
Set $g^{(k)} := \sum_{i=1}^{N} P_i g_i + P_0 g_0$

Let us again remark that we obtain the corresponding RASPEN methods by replacing $P_i,\ i = 1, ..., N$, by $\widetilde{P}_i,\ i = 1, ..., N$, in Algorithms 4 to 8. The corresponding ASPIN or RASPIN methods are finally obtained by replacing $D\mathcal{F}_X$ in Algorithm 2 by the appropriate approximations suggested in section 4.

7.3. Globalization and inexact solution of the tangential system. We always use an inexact Newton method in the outer loop and solve our preconditioned tangential system using GMRES up to a certain tolerance tol_{GMRES}. In contrast, we use an exact Newton method for the local and coarse corrections, i.e., we solve the tangent problems using sparse direct solves. Without any additional globalization strategy, i.e., by fixing $\lambda = 1$ in Algorithms 1 to 3, global convergence of Newton's methods is in both cases not guaranteed. Controlling the step-length instead is beneficial.

A successful approach to increase the convergence radius of Newton's method for the solution of an arbitrary nonlinear problem $G(x) = 0$ is the globally convergent Inexact Newton Backtracking (INB) approach; see [19]. For a given forcing term η, the descent condition

$$||G(x^{(k)} - \lambda \delta x^{(k)})|| \leq (1 - c\,(1 - \eta))\, ||G(x^{(k)})|| \tag{31}$$

Algorithm 6 Evaluation of two-level hybrid ASPEN function $g^{(k)} := \mathcal{F}_{\mathrm{h},1}(u^{(k)})$ (coarse correction first)

Compute coarse correction $g_0 := T_0(u^{(k)})$ /*with Algorithm 3*/
Set $\tilde{u} := u^{(k)} - P_0 g_0$
Compute local corrections $g_i := T_i(\tilde{u}),\ i = 1, ..., N$ /*with Algorithm 3*/
Set $g^{(k)} := \sum_{i=1}^{N} P_i g_i + P_0 g_0$

Algorithm 7 Evaluation of two-level hybrid ASPEN function $g^{(k)} := \mathcal{F}_{\mathrm{h},2}(u^{(k)})$ (coarse correction second)

Compute local corrections $g_i := T_i(u^{(k)}),\ i = 1, ..., N$ /*with Algorithm 3*/
Set $\tilde{u} := u^{(k)} - \sum_{i=1}^{N} P_i g_i$
Compute coarse correction $g_0 := T_0(\tilde{u})$ /*with Algorithm 3*/
Set $g^{(k)} := \sum_{i=1}^{N} P_i g_i + P_0 g_0$

has to be fulfilled for λ using a backtracking approach. Here, the constant c is usually small, e.g., $c = 10^{-4}$, $x^{(k)}$ is the current Newton iterate, and $\delta x^{(k)}$ is the Newton update. The forcing term is initialized with $\eta = tol_{\mathrm{GMRES}}$ and modified dependent on λ during the backtracking procedure; see [19] for details. In our computations, we always choose $tol_{\mathrm{GMRES}} = 10^{-10}$ if no globalization strategy is used and $tol_{\mathrm{GMRES}} = 10^{-4}$ if INB is used. The latter choice avoids an oversolving of the linear system and potentially saves GMRES iterations; more elaborate choices of the forcing terms can be found in [20]. To avoid ineffectively small step-lengths, we choose $\lambda \in [0.1, 1]$.

The INB approach can be used without any modification in Algorithm 1 and also in the computation of the local or coarse corrections (Algorithm 3), even though the forcing term is close to zero in the latter case. In contrast to that, enforcing (31) in Algorithm 2, i.e., enforcing

$$||\mathcal{F}_X(u^{(k)} - \lambda \delta u^{(k)})|| \leq (1 - c\,(1 - \eta))\,||\mathcal{F}_X(u^{(k)})||, \tag{32}$$

requires many evaluations of $\mathcal{F}_X(u^{(k)} - \lambda \delta u^{(k)})$ for different λ, which results in many additional inner and coarse iterations. We propose to replace $\mathcal{F}_X$ by the original nonlinear function F in (32) as long as $\delta u^{(k)}$ is a descent direction with respect to the energy $\frac{1}{2}||F(\cdot)||^2$. If $\delta u^{(k)}$ is not a descent direction, we suggest to completely reject the ASPEN update $\delta u^{(k)}$ and use a Newton-Krylov step instead. This was not necessary in any of our computations.

Since we are essentially interested in the globalization and convergence properties of the different nonlinear coarse spaces and methods themselves, we do not use INB in general. These properties can be polluted using variable step lengths. Nonetheless, for some of the considered algorithms applied to some of our model problems using INB or an alternative globalization approach is necessary for convergence and we present some results using INB in subsection 8.2. Of course, in practice, we always recommend to use INB or to include an alternative globalization strategy.

8. Numerical results. In this section, we consider the different nonlinear Schwarz methods described in section 4 using standard $\mathbb{P}1$ coarse spaces and the coarse spaces described in section 6. We compare the numbers of outer Newton iterations, inner Newton iterations, coarse Newton iterations, and linear GMRES iterations. We always provide the sum over all outer Newton steps for the inner Newton iterations,

Algorithm 8 Evaluation of symmetric two-level hybrid ASPEN function $g^{(k)} := \mathcal{F}_{\mathrm{h}}(u^{(k)})$

Compute coarse correction $g_0 := T_0(u^{(k)})$ /*with Algorithm 3*/
Set $\tilde{u} := u^{(k)} - P_0 g_0$
Compute local corrections $g_i := T_i(\tilde{u}),\ i = 1, ..., N$ /*with Algorithm 3*/
Set $\hat{u} := \tilde{u} - \sum_{i=1}^{N} P_i g_i$
Compute coarse correction $\tilde{g}_0 := T_0(\hat{u})$ /*with Algorithm 3*/
Set $g^{(k)} := \sum_{i=1}^{N} P_i g_i + P_0 g_0 + P_0 \tilde{g}_0$

coarse Newton iterations, and linear GMRES iterations and additionally average over all subdomains for the inner iterations.

8.1. Homogeneous model problem. For a homogeneous coefficient distribution, distance based and the energy minimizing edge values for our MsFEM type coarse spaces yield comparable results; however, the computation of the energy minimizing edge values is computationally more demanding. Therefore, we compare only $\mathbb{P}1$ and MsFEM-D coarse spaces in this subsection.

In particular, we cover the following aspects in our numerical investigations for the homogeneous model problem: we compare the different coarse spaces for the suggested two-level RASPEN methods (see Table 1), we investigate the work distribution in the different suggested one- and two-level RASPEN methods (see Figure 4), we compare the one- and two-level RASPEN methods to the associated Newton-Krylov-Schwarz approaches for regular and METIS decompositions (see Table 5), and briefly investigate the use of inexact (RASPIN) methods for regular and METIS domain decompositions (see Table 2).

Different coarse spaces. In Table 1, we present a comparison of the $\mathbb{P}1$ coarse space and a MsFEM type coarse space with (distance based edge values) using the two different energy-minimizing extensions. In general, for the homogeneous model problem, all three coarse spaces show a similar behavior with respect to all metrics. In all cases, the multiplicative approaches are superior and, for this simple test case, H1-RASPEN shows the best performance. Let us remark that the $DK(u^{(0)})$ extension is our favorite variant since it can be used for METIS domain decompositions and does not require the availability of a linear surrogate matrix K^{lin}.

Work distribution. As mentioned before, we average the numbers of inner Newton iterations over the subdomains in order to estimate the average local work on each subdomain. However, in a parallel implementation, also the work distribution is important, which is dominated by the distribution of inner Newton iterations in nonlinear DD methods. To investigate this, we visualize the number of inner iterations for each subdomain for the different RASPEN approaches in Figure 4 for a regular domain decomposition into 25 subdomains. Here, H-RASPEN and H1-RAPSEN yield the best work distribution. Let us remark that the work imbalance in nonlinear DD methods can also be exploited to save energy. This was discussed in [42] for nonlinear FETI-DP methods as well as for ASPIN, where the nonlinear DD approaches always had a lower power consumption compared with corresponding Newton-Krylov-DD methods.

Regular and METIS domain decompositions. In Table 5, we consider all one- and two-level Newton-Krylov-Schwarz and RASPEN approaches described in this paper using MsFEM type coarse spaces. We do not consider $\mathbb{P}1$ coarse spaces for METIS

TABLE 1

Comparison of different coarse spaces ($\mathbb{P}1$ and MsFEM-D with different extensions described in subsection 6.1) for two-level RAPSEN methods for regular domain decompositions; best results for the largest experiments are marked in bold; **outer it.** *gives the number of global Newton iterations;* **inner it.** *gives the number of local Newton iterations summed up over the outer Newton iterations (average over subdomains);* **coarse it.** *gives the number of nonlinear iterations on the coarse level summed up over the outer Newton iterations;* **GMRES it.** *gives the number of GMRES iterations summed up over the outer Newton iterations.*

p-Laplace homogeneous $p = 4$; $H/h = 16$ for regular domains; overlap $\delta = 1$;									
		A-RASPEN				H-RASPEN			
N	**coarse space**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**
9	$\mathbb{P}1$	6	31.3	24	98	4	15.2	28	55
	$DK(u^{(0)})$-MsFEM-D	6	33.4	27	93	4	17.1	29	52
	K^{lin}-MsFEM-D	6	13.6	24	90	3	13.8	21	38
25	$\mathbb{P}1$	5	26.3	24	109	4	13.4	29	74
	$DK(u^{(0)})$-MsFEM-D	6	29.7	29	122	4	13.3	30	72
	K^{lin}-MsFEM-D	6	28.5	26	119	3	11.3	22	52
49	$\mathbb{P}1$	6	29.5	28	150	4	12.1	28	82
	$DK(u^{(0)})$-MsFEM-D	6	29.2	28	137	4	12.6	29	80
	K^{lin}-MsFEM-D	6	29.7	29	135	**3**	**10.2**	23	**56**
		H1-RASPEN				H2-RASPEN			
9	$\mathbb{P}1$	4	15.2	17	55	5	27.5	17	78
	$DK(u^{(0)})$-MsFEM-D	4	17.1	18	51	5	27.8	15	74
	K^{lin}-MsFEM-D	3	13.8	13	39	5	27.2	15	75
25	$\mathbb{P}1$	4	13.4	18	72	5	25.8	17	101
	$DK(u^{(0)})$-MsFEM-D	4	13.3	19	70	5	25.8	18	96
	K^{lin}-MsFEM-D	3	11.3	14	53	5	25.9	17	97
49	$\mathbb{P}1$	4	12.1	18	79	5	25.2	18	110
	$DK(u^{(0)})$-MsFEM-D	4	12.6	20	78	5	25.3	17	109
	K^{lin}-MsFEM-D	**3**	**10.2**	**15**	**56**	5	25.3	18	110

domain decompositions because they would require additional coarse triangulations, which are often not available in practice. Especially the hybrid two-level approaches can reduce the number of outer Newton iterations drastically for both types of domain decompositions. All two-level methods also enable numerical scalability in the linear iterations, which is not the case in one-level RASPEN or one-level Newton-Krylov-Schwarz. Also the nonlinear convergence of the two-level RASPEN methods is superior compared to one-level RASPEN. For the sake of clarity, we will neglect the "A" (additive) and "H2" (hybrid–coarse correction second) variants and concentrate only on the "H" (hybrid–symmetric) and "H1" (hybrid–coarse correction first) variants in the following sections.

RASPEN vs. RASPIN methods. We compare one- and two-level RASPEN and RASPIN methods for regular and METIS domain decompositions. Whereas the RASPIN method yields better convergence compared to the RASPEN method for the one-level case, the convergence behavior is comparable for the two-level variants; see Table 2.

To summarize, considering our metrics, i.e. nonlinear and linear convergence as well as load balance, H1-RASPEN and H-RASPEN perform best. Even for irregular decompositions, our Galerkin product based approaches combined with suitable coarse spaces perform equally well.

8.2. Heterogeneous model problem. For the heterogeneous model problem, i.e., $\alpha = 10^3$ or $\alpha = 10^6$ in the yellow circles in Figure 1 (left) and $\alpha = 1$ elsewhere, we compare Netwon-Krylov-Schwarz and nonlinear one- and two-level Schwarz methods.

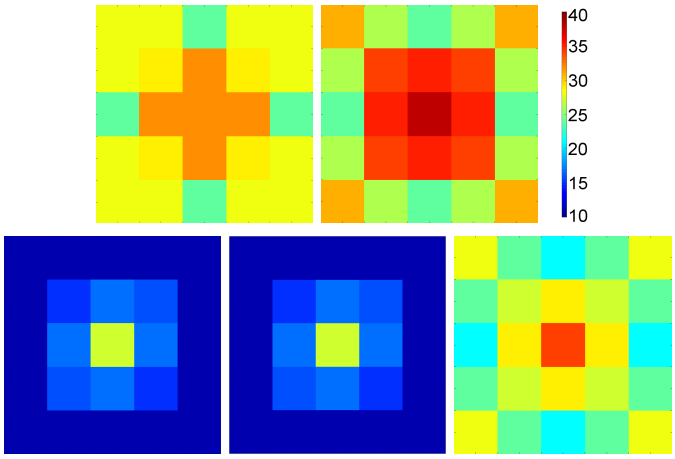

FIG. 4. *Number of inner Newton iterations per subdomain summed over all outer Newton iterations. Considered problem: Homogeneous 4-Laplace with MsFEM-D coarse space and $DK(u^{(0)})$ extension; $H/h = 16$ and overlap $\delta = 1$.* **From top left to bottom right:** *one-level RASPEN, A-RASPEN, H-RASPEN, H1-RASPEN, and H2-RASPEN.*

Here, an appropriate coarse space is necessary to obtain good linear convergence and, in case of nonlinear Schwarz methods, also a fast nonlinear convergence. Choosing the MsFEM-E coarse space results in very fast convergence of H1-RASPEN, outperforming the corresponding Newton-Krylov approach. On the other hand, one-level RASPEN or H1-RASPEN with a $\mathbb{P}1$ or MsFEM-D coarse spaces do not converge within 20 Newton iterations; see Table 3 for the results. Additionally, we provide the convergence history of the outer Newton iteration for four exemplary methods (RASPEN, H1-RASPEN with MsFEM-D and MsFEM-E coarse spaces, and Newton-Krylov-H1-RAS with MsFEM-E coarse space) in Figure 5 (top row).

Since a globalization strategy can have a huge impact on the convergence behavior, we repeated the same tests using globally convergent INB as described in subsection 7.3. The corresponding results are presented in Table 4 and Figure 5 (bottom row). Now, all considered methods converge within 20 iterations and the number of inner as well as coarse iterations is reduced. Additionally, due to the larger stopping tolerance $tol_{\text{GMRES}} = 10^{-4}$ also the number of linear iterations is reduced drastically. Again, we observe that the choice of an appropriate coarse space is critical for fast convergence of the two-level nonlinear Schwarz method. The H1-RASPEN approach with MsFEM-E coarse space clearly outperforms all other approaches with respect to linear as well as nonlinear convergence. Notably, for the considered heterogeneous model problem, the convergence behavior is even nearly independent of the coefficient jump as well as the globalization strategy; cf. Table 3 and Table 4; only the number of GMRES iterations varies due to the different stopping tolerance.

REFERENCES

[1] J. Aarnes and T. Y. Hou. Multiscale domain decomposition methods for elliptic problems with high aspect ratios. *Acta Math. Appl. Sin. Engl. Ser.*, 18(1):63–76, 2002.

[2] P. Brune, M. Knepley, B. Smith, and X. Tu. Composing scalable nonlinear algebraic solvers. *SIAM Review*, 57(4):535–565, 2015.

[3] M. Buck, O. Iliev, and H. Andrä. Multiscale finite elements for linear elasticity: Oscillatory boundary conditions. In *Domain Decomposition Methods in Science and Engineering XXI*,

TABLE 2

Comparison of one- and two-level RASPEN and corresponding RASPIN methods; best results for the largest experiment are marked in bold; **outer it.** *gives the number of global Newton iterations;* **inner it.** *gives the number of local Newton iterations summed up over the outer Newton iterations (average over subdomains);* **coarse it.** *gives the number of nonlinear iterations on the second level summed up over the outer Newton iterations;* **GMRES it.** *gives the number of GMRES iterations summed up over the outer Newton iterations.*

p-Laplace homogeneous; MsFEM-D coarse space with $DK(u^{(0)})$ extension $p = 4$; $H/h = 16$ for regular domains; overlap $\delta = 1$;									
		Regular				METIS			
N	**solver**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**
16	RASPIN	6	26.3	-	125	7	29.2	-	244
	RASPEN	11	56.3	-	247	8	43.4	-	291
	H-RASPIN	9	37.0	77	148	6	21.3	44	123
	H-RASPEN	7	26.0	65	118	5	19.9	43	105
	H1-RASPIN	6	21.6	31	97	5	18.6	25	102
	H1-RASPEN	7	26.0	39	119	5	19.9	27	104
25	RASPIN	7	27.9	-	194	7	27.7	-	289
	RASPEN	6	28.3	-	172	7	33.4	-	290
	H-RASPIN	5	16.7	35	91	4	14.7	32	82
	H-RASPEN	4	13.3	30	72	4	14.7	34	82
	H1RASPIN	5	16.8	23	88	5	16.8	25	101
	H1-RASPEN	4	13.3	19	70	4	14.7	22	80
36	RASPIN	6	24.7	-	191	7	27.5	-	336
	RASPEN	12	51.0	-	396	10	46.3	-	483
	H-RASPIN	9	30.5	80	181	6	19.0	51	130
	H-RASPEN	7	23.1	69	143	7	21.9	60	158
	H1-RASPIN	6	18.4	32	115	6	19.3	32	129
	H1-RASPEN	7	23.1	41	142	7	21.8	37	155
49	RASPIN	7	27.8	-	254	8	29.6	-	452
	RASPEN	6	27.3	-	232	10	46.5	-	560
	H-RASPIN	5	15.3	35	103	**6**	18.5	49	137
	H-RASPEN	**4**	**12.6**	29	80	**6**	18.8	53	142
	H1-RASPIN	5	15.3	23	98	**6**	**18.1**	**30**	**134**
	H1-RASPEN	**4**	**12.6**	**20**	**78**	**6**	18.8	33	136

volume 98 of *LNCSE*, pages 237–245. Springer, 2014.

[4] X.-C. Cai and M. Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In *Contemporary Math*, pages 21–28. AMS, 1994.

[5] X.-C. Cai and D. E. Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SISC*, 24(1):183–200, 2002.

[6] X.-C. Cai, D. E. Keyes, and L. Marcinkowski. Non-linear additive Schwarz preconditioners and application in computational fluid dynamics. *International Journal for Numerical Methods in Fluids*, 40(12):1463–1470.

[7] X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SISC*, 21(2):792–797, 1999.

[8] J. G. Calvo. On the approximation of a virtual coarse space for domain decomposition methods in two dimensions. *Math. Models Methods Appl. Sci.*, 28(7):1267–1289, 2018.

[9] P. Cresta, O. Allix, C. Rey, and S. Guinard. Nonlinear localization strategies for domain decomposition methods: Application to post-buckling analyses. *Computer Methods in Applied Mechanics and Engineering*, 196(8):1436 – 1446, 2007. Domain Decomposition Methods: recent advances and new challenges in engineering.

[10] C. R. Dohrmann, A. Klawonn, and O. B. Widlund. Domain decomposition for less regular subdomains: overlapping Schwarz in two dimensions. *SINUM*, 46(4):2153–2168, 2008.

[11] C. R. Dohrmann, A. Klawonn, and O. B. Widlund. A family of energy minimizing coarse spaces for overlapping Schwarz preconditioners. In *Domain decomposition methods in science and engineering XVII*, volume 60 of *LNCSE*, pages 247–254. Springer, Berlin, 2008.

[12] C. R. Dohrmann and O. B. Widlund. Hybrid domain decomposition algorithms for compressible and almost incompressible elasticity. *IJNME*, 82(2):157–183, 2010.

[13] C. R. Dohrmann and O. B. Widlund. Lower dimensional coarse spaces for domain decomposi-

TABLE 3

Comparison of one- and two-level RASPEN methods and two-level Newton-Krylov-Schwarz methods with different coarse spaces applied to highly heterogeneous problems; **outer it.** *gives the number of global Newton iterations;* **inner it.** *gives the number of local Newton iterations summed up over the outer Newton iterations (average over subdomains);* **coarse it.** *gives the number of nonlinear iterations on the second level summed up over the outer Newton iterations;* **GMRES it.** *gives the number of GMRES iterations summed up over the outer Newton iterations.*

p-Laplace heterogeneous; four circles (see Figure 1 (left)) $\mathbb{P}1$, $DK(u^{(0)})$-MsFEM, and $DK(u^{(0)})$-MsFEM-E coarse space $p = 4$; $H/h = 16$; overlap $\delta = 1$; $N = 25$ square subdomains						
α	**solver**	**coarse space**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**
1e3	RASPEN	-	>20	-	-	-
	NK-H1-RAS	$\mathbb{P}1$	14	-	-	370
		$DK(u^{(0)})$-MsFEM-D	14	-	-	346
		$DK(u^{(0)})$-MsFEM-E	14	-	-	341
	H1-RASPEN	$\mathbb{P}1$	>20	-	-	-
		$DK(u^{(0)})$-MsFEM-D	>20	-	-	-
		$DK(u^{(0)})$-MsFEM-E	6	22.5	26	150
1e6	RASPEN	-	>20	-	-	-
	NK-H1-RAS	$\mathbb{P}1$	17	-	-	738
		$DK(u^{(0)})$-MsFEM-D	17	-	-	654
		$DK(u^{(0)})$-MsFEM-E	17	-	-	422
	H1-RASPEN	$\mathbb{P}1$	>20	-	-	-
		$DK(u^{(0)})$-MsFEM-D	>20	-	-	-
		$DK(u^{(0)})$-MsFEM-E	5	32.1	31	139

tion. In *Domain Decomposition Methods in Science and Engineering XXI*, pages 527–535. Springer, 2014.

[14] C. R. Dohrmann and O. B. Widlund. On the design of small coarse spaces for domain decomposition algorithms. *SISC*, 39(4):A1466–A1488, 2017.

[15] V. Dolean, M. J. Gander, W. Kheriji, F. Kwok, and R. Masson. Nonlinear preconditioning: how to use a nonlinear Schwarz method to precondition Newton's method. *SISC*, 38(6):A3357–A3380, 2016.

[16] V. Dolean, F. Nataf, R. Scheichl, and N. Spillane. Analysis of a two-level Schwarz method with coarse spaces based on local Dirichlet-to-Neumann maps. *Comput. Methods Appl. Math.*, 12(4):391–414, 2012.

[17] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT Numerical Mathematics*, 37(2):296–311, Jun 1997.

[18] E. Eikeland, L. Marcinkowski, and T. Rahman. Overlapping Schwarz methods with adaptive coarse spaces for multiscale problems in 3D. *Numerische Mathematik*, Nov 2018.

[19] S. Eisenstat and H. Walker. Globally convergent inexact Newton methods. *SIAM Journal on Optimization*, 4(2):393–422, 1994.

[20] S. C. Eisenstat and H. F. Walker. Choosing the forcing terms in an inexact Newton method. *SISC*, 17(1):16–32, 1996. Special issue on iterative methods in numerical linear algebra (Breckenridge, CO, 1994).

[21] J. Galvis and Y. Efendiev. Domain decomposition preconditioners for multiscale flows in high-contrast media. *Multiscale Model. Simul.*, 8(4):1461–1483, 2010.

[22] M. J. Gander, A. Loneland, and T. Rahman. Analysis of a new harmonically enriched multiscale coarse space for domain decomposition methods. *arXiv preprint arXiv:1512.05285*, 2015.

[23] I. G. Graham, P. O. Lechner, and R. Scheichl. Domain decomposition for multiscale PDEs. *Numer. Math.*, 106(4):589–626, 2007.

[24] A. Heinlein, C. Hochmuth, and A. Klawonn. Monolithic overlapping Schwarz domain decomposition methods with GDSW coarse spaces for saddle point problems. Technical report, Universität zu Köln, July 2018.

[25] A. Heinlein, C. Hochmuth, and A. Klawonn. Reduced dimension GDSW coarse spaces for monolithic Schwarz domain decomposition methods for incompressible fluid flow problems. Technical report, Universität zu Köln, May 2019.

[26] A. Heinlein, A. Klawonn, J. Knepper, and O. Rheinbach. An adaptive GDSW coarse space for two-level overlapping Schwarz methods in two dimensions. In *Domain Decomposition*

TABLE 4

Comparison of some one- and two-level RASPEN methods and two-level Newton-Krylov-Schwarz methods **with backtracking** *using different coarse spaces applied to highly heterogeneous problems;* **outer it.** *gives the number of global Newton iterations;* **inner it.** *gives the number of local Newton iterations summed up over the outer Newton iterations (average over subdomains);* **coarse it.** *gives the number of nonlinear iterations on the second level summed up over the outer Newton iterations;* **GMRES it.** *gives the number of GMRES iterations summed up over the outer Newton iterations.*

p-Laplace heterogeneous; four circles(see Figure 1 (left)) $\mathbb{P}1$, $DK(u^{(0)})$-MsFEM, and $DK(u^{(0)})$-MsFEM-E coarse space $p = 4$; $H/h = 16$; overlap $\delta = 1$; $N = 25$ square subdomains						
α	**solver**	**coarse space**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**
1e3	RASPEN	-	7	30.1	-	223
	NK-H1-RAS	$\mathbb{P}1$	10	-	-	143
		$DK(u^{(0)})$-MsFEM-D	10	-	-	126
		$DK(u^{(0)})$-MsFEM-E	10	-	-	94
	H1-RASPEN	$\mathbb{P}1$	7	33.4	29	105
		$DK(u^{(0)})$-MsFEM-D	7	32.4	28	99
		$DK(u^{(0)})$-MsFEM-E	5	19.6	18	49
1e6	RASPEN	-	10	53.1	-	454
	NK-H1-RAS	$\mathbb{P}1$	15	-	-	322
		$DK(u^{(0)})$-MsFEM-D	15	-	-	289
		$DK(u^{(0)})$-MsFEM-E	15	-	-	133
	H1-RASPEN	$\mathbb{P}1$	12	70.6	71	312
		$DK(u^{(0)})$-MsFEM-D	13	77.0	70	364
		$DK(u^{(0)})$-MsFEM-E	5	25.8	25	48

Methods in Science and Engineering XXIV, pages 373–382, Cham, 2018. Springer.

[27] A. Heinlein, A. Klawonn, J. Knepper, and O. Rheinbach. Adaptive GDSW coarse spaces for overlapping Schwarz methods in three dimensions. Technical report, Universität zu Köln, October 2018.

[28] A. Heinlein, A. Klawonn, J. Knepper, and O. Rheinbach. Multiscale coarse spaces for overlapping Schwarz methods based on the ACMS space in 2D. *ETNA*, 48:156–182, 2018.

[29] A. Heinlein, A. Klawonn, and O. Rheinbach. A parallel implementation of a two-level overlapping Schwarz method with energy-minimizing coarse space based on Trilinos. *SISC*, 38(6):C713–C747, 2016.

[30] A. Heinlein and M. Lanser. Coarse spaces for nonlinear Schwarz methods on unstructured grids. Technical report, Universität zu Köln, November 2018. Accepted for publication to Springer LNCSE.

[31] J. Hinojosa, O. Allix, P.-A. Guidault, and P. Cresta. Domain decomposition methods with nonlinear localization for the buckling and post-buckling analyses of large structures. *Advances in Engineering Software*, 70:13 – 24, 2014.

[32] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134(1):169 – 189, 1997.

[33] J. Huang, C. Yang, and X. Cai. A nonlinearly preconditioned inexact newton algorithm for steady state lattice Boltzmann equations. *SISC*, 38(3):A1701–A1724, 2016.

[34] F.-N. Hwang and X.-C. Cai. A class of parallel two-level nonlinear Schwarz preconditioned inexact Newton algorithms. *CMAME*, 196(8):1603 – 1611, 2007.

[35] F.-N. Hwang, Y.-C. Su, and X.-C. Cai. A parallel adaptive nonlinear elimination preconditioned inexact Newton method for transonic full potential equation. *Comput. & Fluids*, 110:96–107, 2015.

[36] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SISC*, 20(1):359–392, 1998.

[37] A. Klawonn, M. Lanser, and O. Rheinbach. Nonlinear FETI-DP and BDDC methods. *SISC*, 36(2):A737–A765, 2014.

[38] A. Klawonn, M. Lanser, and O. Rheinbach. Toward extremely scalable nonlinear domain decomposition methods for elliptic partial differential equations. *SISC*, 37(6):C667–C696, 2015.

[39] A. Klawonn, M. Lanser, and O. Rheinbach. A highly scalable implementation of inexact nonlinear FETI-DP without sparse direct solvers. In *ENUMATH 2015*, volume 112 of *LNCSE*,

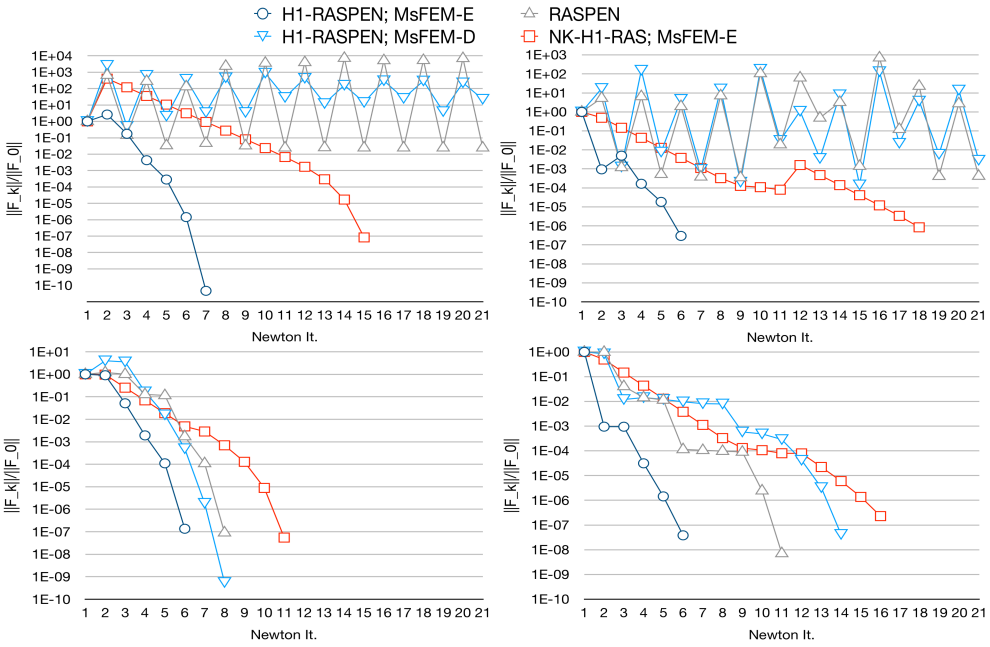

FIG. 5. *Convergence behavior of outer Newton iteration for heterogeneous problems; see Table 3 and Table 4 for corresponding results.* **Left Column:** $\alpha = 10^3$*;* **Right Column:** $\alpha = 10^6$*;* **Top Row:** *Without globalization strategy;* **Bottom Row:** *With globalization strategy, i.e., using globally convergent INB as described in subsection 7.3.*

pages 255–264. Springer Cham, 2016.

[40] A. Klawonn, M. Lanser, and O. Rheinbach. Nonlinear BDDC methods with approximate solvers. *Electron. Trans. Numer. Anal.*, 49:244–273, 2018.

[41] A. Klawonn, M. Lanser, O. Rheinbach, and M. Uran. Nonlinear FETI-DP and BDDC methods: a unified framework and parallel results. *SISC*, 39(6):C417–C451, 2017.

[42] A. Klawonn, M. Lanser, O. Rheinbach, G. Wellein, and M. Wittmann. Energy efficiency of nonlinear domain decomposition methods. Technical report, Universität zu Köln, 2018.

[43] L. Liu and D. E. Keyes. Field-split preconditioned inexact Newton algorithms. *SISC*, 37(3):A1388–A1409, 2015.

[44] S. Lui. On Schwarz alternating methods for the incompressible Navier–Stokes equations. *SISC*, 22(6):1974–1986, 2001.

[45] L. Marcinkowski and X.-C. Cai. Parallel performance of some two-level ASPIN algorithms. In *Domain decomposition methods in science and engineering*, volume 40 of *LNCSE*, pages 639–646. Springer, Berlin, 2005.

[46] C. Negrello, P. Gosselet, and C. Rey. Nonlinearly preconditioned FETI solver for substructured formulations of nonlinear problems. working paper or preprint, April 2018.

[47] J. Pebrel, C. Rey, and P. Gosselet. A nonlinear dual-domain decomposition method: Application to structural problems with damage. *International Journal for Multiscale Computational Engineering*, 6(3):251–262, 2008.

[48] M. Sarkis. Partition of unity coarse spaces and Schwarz methods with harmonic overlap. In *Recent developments in domain decomposition methods (Zürich, 2001)*, volume 23 of *LNCSE*, pages 77–94. Springer, Berlin, 2002.

[49] N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl. Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. *Numer. Math.*, 126(4):741–770, 2014.

[50] X.-C. Tai and M. Espedal. Rate of convergence of some space decomposition methods for linear and nonlinear problems. *SISC*, 35:1558–1570, 1998.

[51] A. Toselli and O. Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2005.

TABLE 5

Comparison of all considered one- and two-level RASPEN and corresponding Newton-Krylov-Schwarz variants using the MsFEM-D type coarse space with $DK(u^{(0)})$ extensions; best results for the largest experiment are marked in bold; **outer it.** *gives the number of global Newton iterations;* **inner it.** *gives the number of local Newton iterations summed up over the outer Newton iterations (average over subdomains);* **coarse it.** *gives the number of nonlinear iterations on the second level summed up over the outer Newton iterations;* **GMRES it.** *gives the number of GMRES iterations summed up over the outer Newton iterations.*

p-Laplace homogeneous; MsFEM coarse space with $DK(u^{(0)})$ extension $p = 4$; $H/h = 16$ for regular domains; overlap $\delta = 1$;									
		Regular				METIS			
N	**solver**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**	**outer it.**	**inner it. (avg.)**	**coarse it.**	**GMRES it. (sum)**
9	NK-RAS	18	-	-	272	18	-	-	438
	RASPEN	5	25.2	-	89	6	32.9	-	171
	NK-A-RAS	18	-	-	265	18	-	-	374
	A-RASPEN	6	33.4	27	93	6	34.2	27	136
	NK-H-RAS	18	-	-	246	18	-	-	332
	H-RASPEN	4	17.1	29	52	4	17.7	31	75
	NK-H1-RAS	18	-	-	247	18	-	-	337
	H1-RASPEN	4	17.1	18	51	4	17.7	19	73
	NK-H2-RAS	18	-	-	234	18	-	-	329
	H2-RASPEN	5	27.8	15	74	5	28.4	17	103
16	NK-RAS	19	-	-	403	19	-	-	638
	RASPEN	11	56.3	-	247	8	43.4	-	291
	NK-A-RAS	19	-	-	362	19	-	-	472
	A-RASPEN	8	38.1	42	149	7	35.4	37	176
	NK-H-RAS	19	-	-	327	19	-	-	415
	H-RASPEN	7	26.0	65	118	5	19.9	43	105
	NK-H1-RAS	19	-	-	327	19	-	-	418
	H1-RASPEN	7	26.0	39	119	5	19.9	27	104
	NK-H2-RAS	19	-	-	315	19	-	-	393
	H2-RASPEN	6	31.6	27	107	5	28.2	19	113
25	NK-RAS	19	-	-	488	19	-	-	753
	RASPEN	6	28.3	-	172	7	33.4	-	290
	NK-A-RAS	19	-	-	369	19	-	-	442
	A-RASPEN	6	29.7	29	122	7	33.8	35	173
	NK-H-RAS	19	-	-	345	19	-	-	394
	H-RASPEN	4	13.3	30	72	4	14.7	34	82
	NK-H1-RAS	19	-	-	346	19	-	-	398
	H1-RASPEN	4	13.3	19	70	4	14.7	22	80
	NK-H2-RAS	19	-	-	330	19	-	-	390
	H2-RASPEN	5	25.8	18	96	5	26.5	19	113
36	NK-RAS	20	-	-	650	20	-	-	967
	RASPEN	12	51.0	-	396	10	46.3	-	483
	NK-A-RAS	20	-	-	448	20	-	-	524
	A-RASPEN	7	32.2	39	150	8	35.6	46	205
	NK-H-RAS	20	-	-	418	20	-	-	472
	H-RASPEN	7	23.1	69	143	7	21.9	60	158
	NK-H1-RAS	20	-	-	417	20	-	-	472
	H1-RASPEN	7	23.1	41	142	7	21.8	37	155
	NK-H2-RAS	20	-	-	400	20	-	-	456
	H2-RASPEN	5	26.9	23	104	5	25.7	20	119
49	NK-RAS	20	-	-	691	20	-	-	1 123
	RASPEN	6	27.3	-	232	10	46.5	-	560
	NK-A-RAS	20	-	-	428	20	-	-	540
	A-RASPEN	6	29.2	28	137	7	32.9	40	186
	NK-H-RAS	20	-	-	403	20	-	-	488
	H-RASPEN	**4**	**12.6**	29	80	**6**	**18.8**	53	142
	NK-H1-RAS	20	-	-	408	20	-	-	487
	H1-RASPEN	**4**	**12.6**	20	**78**	**6**	**18.8**	33	**136**
	NK-H2-RAS	20	-	-	383	20	-	-	466
	H2-RASPEN	5	25.3	**17**	109	**6**	27.5	**23**	145