Electronic Transactions on Numerical Analysis.
Volume 54, pp. 89–107, 2021.

ISSN 1068–9613.
DOI: 10.1553/etna_vol54s89

ETNA
Kent State University and
Johann Radon Institute (RICAM)

MULTILEVEL SCHWARZ PRECONDITIONERS FOR SINGULARLY PERTURBED SYMMETRIC REACTION-DIFFUSION SYSTEMS*

JOSÉ PABLO LUCERO LORCA† AND GUIDO KANSCHAT‡

Abstract. We present robust and highly parallel multilevel non-overlapping Schwarz preconditioners to solve an interior penalty discontinuous Galerkin finite element discretization of a system of steady-state, singularly perturbed reaction-diffusion equations with a singular reaction operator using a GMRES solver. We provide proofs of convergence for the two-level setting and the multigrid V-cycle as well as numerical results for a wide range of regimes.

Key words. multilevel, Schwarz, preconditioner, multigrid, reaction, diffusion, discontinuous, Galerkin

AMS subject classifications. 65N55, 65N30, 65J10, 65F08

1. Introduction.

In this paper, we present an analysis and numerical experiments of two-level Schwarz preconditioners and their multilevel versions for an interior penalty discontinuous Galerkin finite element discretization of a system of reaction-diffusion equations not requiring special mesh structures for resolving boundary layers. Our focus is on the singularly perturbed case, where the reaction system has an inertial subspace. We use a massively parallel smoother as in [10], and therefore we provide new convergence estimates for elliptic and reaction-diffusion systems including quadrilateral and hexahedral meshes. The estimates are robust with respect to the parameters of the system, and the experiments confirm the efficiency of the method.

Reaction-diffusion systems arise in a variety of physical, chemical, and biological contexts. Due to the conservation of mass, these systems are all characterized by an inertial subspace (an inertial manifold in the nonlinear case) on which the system reduces to an almost reaction-free diffusion equation. Nevertheless, the contributions orthogonal to this subspace are still important in applications and often cannot be neglected. Thus, numerical methods have to deal with long-ranged couplings in the inertial subspace as well as a short-ranged behavior in its complement in an efficient way. One particular area where these models are widely used is radiation transport, where the reaction-diffusion equation is an approximation of Boltzmann's linear transport equation that becomes relevant in the so-called diffusive regimes, which are characterized by small mean-free paths compared to the size of the domain. In these regimes, the transport equation is nearly singular, and its solution in the interior of the computational domain is close to the solution of a reaction-diffusion equation [17].

We employ the interior penalty discontinuous Galerkin (IP-DG) method to discretize the singularly perturbed reaction-diffusion system in the steady state. IP-DG methods [2, 3, 4, 18, 22] are particularly interesting to solve reaction-diffusion equations since oscillations at boundary and interior layers (the Gibbs phenomenon) are much less notable than with standard conforming finite elements for singularly perturbed problems [15]. Thus, they produce better approximations if such layers are not resolved. Using this discretization, the reaction operator

*Received November 16, 2018. Accepted November 16, 2020. Published online on December 10, 2020. Recommended by S. Brenner. G. K. is supported in part by DFG through priority program 1648 SPPEXA and by the special program "Numerical Analysis of Complex PDE Models in the Sciences" of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI) in Vienna. Computations were run in part on the bwForDev cluster at Heidelberg University.

†Department of Applied Mathematics, University of Colorado at Boulder (pablo.lucero@colorado.edu).

‡Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University (kanschat@uni-heidelberg.de.)

ETNA
Kent State University and
Johann Radon Institute (RICAM)

involves only volume integrals with no coupling between cells. Therefore, we expect that IP-DG is particularly well suited for Schwarz methods since contributions of the reaction term are included inside the local solvers.

We solve the discrete problem with a GMRES solver using multilevel preconditioners with nonoverlapping Schwarz smoothers, effectively solving a full reaction-diffusion problem in each cell; see [14]. Convergence estimates for such methods applied to pure diffusion problems have been developed in [12]. There, it is assumed that the subdomains defining the decomposition of the fine space are unions of coarse cells. It is more efficient though to employ subdomains based on fine cells, and recently, an analysis that covers this case [10] has been developed. However, its application does not extend to quadrilaterals and hexahedra since the proof uses P_1-nonconforming interpolant and enriching operators for simplices; see [7]. We provide an extension for quadrilaterals and hexahedra without such restrictions.

Subspace correction methods for singularly perturbed reaction-diffusion equations have been studied in [5, 16]. Both articles rely on strictly positive reaction terms and use Shishkin-type meshes for the robust discretization of boundary layers. This technique is extended to singularly perturbed reaction-diffusion systems in [20]. Also there, the authors assume a strictly positive definite reaction system, thus being able to make the assumption of an exponential boundary layer but excluding the presence of an inertial subspace.

As stated above, inertial subspaces can be important in applications. Since they do not exhibit boundary layers, Shishkin-type meshes will not be adapted to all solution components. We also do not want to necessarily have to resolve boundary layers, albeit not only the inertial part may be of importance. Thus, we do not solve the limit problem and propose a method which is robust in the sense that its iteration counts are uniformly bounded with respect to the reaction parameters and the mesh size.

Our main results are the proof of the stable decomposition shown in Lemma 3.8 to obtain convergence estimates for two-level preconditioners and the multigrid V-cycle preconditioner estimate in Theorem 3.15. The paper is structured as follows: in Section 2 we introduce the continuous problem and the IP-DG discretization. In Section 3 we develop two-level Schwarz and multigrid preconditioners and prove convergence estimates. Finally, in Section 4 we demonstrate the efficiency of the proposed methods by experimental results.

2. Model problem. We consider the following system of G steady-state reaction-diffusion equations with a singularly perturbed reaction term

$$-\nabla\cdot(\eta_g\nabla u_g)+\frac{1}{\varepsilon}\sum_{g'=1}^{G}\sigma_{gg'}\,(u_g-u_{g'})=S_g \qquad \text{in } \Omega, \text{ with } g=1,\dots,G, \tag{2.1}$$

where g is the *group* index identifying each reacting *substance*, η_g is the diffusion coefficient for each group g, ε is a perturbation parameter defining the relative size of the reaction with respect to the diffusion term, Ω is a convex polyhedral domain in $\mathbb{R}^d$, with $d=2,3$, and S_g is a known source. The equation is provided with the boundary conditions

$$u_g=0 \qquad \text{on } \Gamma, \text{ with } g=1,\dots,G,$$

where Γ is the boundary of Ω.

We assume $\eta_g,\sigma_{gg'}\in C^\infty(\Omega)$ and $\sigma_{gg'}\geq 0$, for all $g,g'=1,\dots,G$, and that there exists $C>0$ such that $\eta_g\geq C$ in Ω. Furthermore, we assume that the reaction matrix is symmetric and singular since

$$\sigma_{gg}=-\sum_{g'\neq g}\sigma_{gg'}, \qquad \forall g=1,\dots,G.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We introduce the Hilbert spaces

$$\mathcal{V} = \left(H_0^1(\Omega)\right)^G, \qquad \mathcal{H} = \left(L^2(\Omega)\right)^G,$$

where $H_0^1(\Omega)$ is the standard Sobolev space with zero boundary trace. They are provided with inner products

$$(u,v)_{\mathcal{V}} = \sum_{g=1}^{G} \left(\eta_g \nabla u_g, \nabla v_g\right)_{L^2(\Omega)}, \qquad (u,v)_{\mathcal{H}} = \sum_{g=1}^{G} \left(u_g, v_g\right)_{L^2(\Omega)},$$

and norms

$$\|u\|_{\mathcal{V}}^2 = (u,u)_{\mathcal{V}}, \qquad \|u\|_{\mathcal{H}}^2 = (u,v)_{\mathcal{H}}.$$

The weak form of problem (2.1) is: find $u \in \mathcal{V}$ such that

$$\mathscr{A}(u,v) = (f,v)_{\mathcal{H}},$$

where $f \in \mathcal{H}$ and the bilinear form is

$$\begin{aligned}\mathscr{A}(u,v) &= \sum_{g=1}^{G} \int_\Omega \eta_g \nabla u_g \cdot \nabla v_g dx + \frac{1}{\varepsilon} \sum_{g=1}^{G} \sum_{g'=1}^{G} \int_\Omega \sigma_{gg'} \left(u_g - u_{g'}\right) v_g dx \\ &= \left(\boldsymbol{D}\nabla u, \nabla v\right)_{\mathcal{H}} + \frac{1}{\varepsilon} \left(\boldsymbol{\Sigma} u, v\right)_{\mathcal{H}} = (u,v)_{\mathcal{V}} + \frac{1}{\varepsilon} \left(\boldsymbol{\Sigma} u, v\right)_{\mathcal{H}}.\end{aligned}$$

The second line uses the vector notation

$$\begin{aligned} u &= (u_1, \ldots, u_G)^\intercal, & v &= (v_1, \ldots, v_G)^\intercal, \\ \boldsymbol{D} &= \operatorname{diag}(\eta_1, \ldots, \eta_G), & \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma_{11} & \ldots & -\sigma_{G1} \\ \ldots & \ldots & \ldots \\ -\sigma_{1G} & \ldots & \sigma_{GG} \end{pmatrix}. \end{aligned}$$

According to our assumptions, the reaction matrix $\boldsymbol{\Sigma}$ is a symmetric, weakly diagonally dominant singular M-matrix[1] with zero column and row sum. By the Perron-Frobenius theorem, this implies that $\boldsymbol{\Sigma}$ is singular with rank less than G, and by the Geršgorin circle theorem, all eigenvalues are nonnegative.

Physically, the properties of $\boldsymbol{\Sigma}$ ensure substance conservation and the absence of sinks inside the domain. In a radiation transport context, this implies that the system can have no particle absorption, and particles only disappear when they reach the boundary. The presence of absorption would imply that all eigenvalues are positive and $\boldsymbol{\Sigma}$ would be invertible.

Under the assumptions on the parameters of equation (2.1), the bilinear form $\mathscr{A}(u,v)$ is continuous and $\mathcal{V}$-coercive, i.e., there exist constants $\gamma_{\mathscr{A}}, C_{\mathscr{A}} > 0$ such that

$$\mathscr{A}(u,u) \geq \gamma_{\mathscr{A}} \|u\|_{\mathcal{V}}^2, \qquad \mathscr{A}(u,v) \leq C_{\mathscr{A}} \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}.$$

Here we remark that even though $\gamma_{\mathscr{A}}$ is independent of ε, $C_{\mathscr{A}}$ is not. From Lax-Milgram's theorem, the variational problem admits a unique solution in $\mathcal{V}$.

[1] We use the term *singular* M-matrix, following the terminology in [13, p. 119], to denote a matrix that can be expressed as $\boldsymbol{A} = s\boldsymbol{I} - \boldsymbol{B}$, where all the elements in $\boldsymbol{B}$ are nonnegative, s is equal to the maximum of the moduli of the eigenvalues of $\boldsymbol{B}$, and $\boldsymbol{I}$ is an identity matrix.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

2.1. Discrete problem. We apply an IP-DG discretization to the bilinear form $\mathscr{A}(\cdot,\cdot)$; cf. [2]. Let $\mathbb{T}_h$ be a subdivision of the domain Ω into quadrilaterals or hexahedra κ such that each cell κ is described by a d-linear mapping Φ_κ from the reference cell $\widehat{\kappa} = [0,1]^d$ onto itself. Conformity of the faces of the mesh cells is not required, but we assume local quasi-uniformity and shape regularity in the sense that the Jacobians of Φ_κ and their inverses are uniformly bounded.

Let $\mathbb{Q}_p$ be the space of tensor product polynomials of degree up to p in each coordinate direction. Then, define the mapped space $\mathbb{Q}_p(\kappa)$ on the cell κ as the pull-back of functions under Φ_κ. The vector-valued, discontinuous function space V_h is defined as

$$V_h = \left\{ v \in \mathcal{H} \middle| v_{|\kappa} \in \mathbb{Q}_p^G(\kappa) \right\}.$$

Let $\mathbb{F}_h^I$ be the set of all interior faces of the mesh and $\mathbb{F}_h^B$ the set of all boundary faces. Let $\kappa_+, \kappa_- \in \mathbb{T}_h$ be two mesh cells with a joint face $F \in \mathbb{F}_h^I$, and let u_+ and u_- be the traces of functions u on F from κ_+ and κ_-, respectively. On an interior face F, we define the *averaging* operator as

$$\{\!\{u\}\!\} = \frac{u_+ + u_-}{\sqrt{2}}.$$

On the boundary, there is only a single value, and we set $\{\!\{u\}\!\} = u$.

We introduce the following definition of mesh integrals

$$\int_{\mathbb{T}_h} u \, dx = \sum_{\kappa \in \mathbb{T}_h} \int_\kappa u \, dx,$$

and integrals over $\mathbb{F}_h^I$ and $\mathbb{F}_h^B$ are defined accordingly. The interior penalty (IP) bilinear form for the scalar Laplacian, as described in [2], reads

$$\alpha_h(u,v) = \int_{\mathbb{T}_h} \nabla u \cdot \nabla v \, dx - \int_{\mathbb{F}_h^I \cup \mathbb{F}_h^B} \left(\{\!\{u\boldsymbol{n}\}\!\} \cdot \{\!\{\nabla v\}\!\} + \{\!\{\nabla u\}\!\} \cdot \{\!\{v\boldsymbol{n}\}\!\} \right) ds + \int_{\mathbb{F}_h^I \cup \mathbb{F}_h^B} \frac{\delta_{\text{IP}}}{h} \{\!\{u\boldsymbol{n}\}\!\} \cdot \{\!\{v\boldsymbol{n}\}\!\} \, ds,$$

where h is the minimum cell diameter adjacent to the face, $u\boldsymbol{n} = (u_1\boldsymbol{n}, u_2\boldsymbol{n}, \dots, u_G\boldsymbol{n})^\mathsf{T}$, and $\nabla u = (\nabla u_1, \nabla u_2, \dots, \nabla u_G)^\mathsf{T}$. We have replaced the jump operator used in [2] by the equivalent expression: $\sqrt{2}\{\!\{u\boldsymbol{n}\}\!\} = u_+\boldsymbol{n}_+ + u_-\boldsymbol{n}_-$. Coercivity and continuity are proven in [2] under the assumption that δ_{IP} is sufficiently large. We will assume in the following that this holds true.

We then define the discrete bilinear form, including the diffusion coefficients, as follows:

$$\begin{aligned} a_h(u,v) = \int_{\mathbb{T}_h} \boldsymbol{D}\nabla u \cdot \nabla v \, dx + \int_{\mathbb{F}_h^I \cup \mathbb{F}_h^B} 4\frac{\delta_{\text{IP}}}{h} \{\!\{\boldsymbol{D}(u\boldsymbol{n})\}\!\} \cdot \{\!\{v\boldsymbol{n}\}\!\} ds \\ - \int_{\mathbb{F}_h^I \cup \mathbb{F}_h^B} 2 \left(\{\!\{u\boldsymbol{n}\}\!\} \cdot \{\!\{\boldsymbol{D}\nabla v\}\!\} + \{\!\{\boldsymbol{D}\nabla u\}\!\} \cdot \{\!\{v\boldsymbol{n}\}\!\} \right) ds. \end{aligned} \tag{2.2}$$

Under the assumptions made in the previous sections and for δ_{IP} sufficiently large, $a_h(u,v)$ is coercive and continuous.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Using (2.2), our IP-DG formulation for the singularly perturbed reaction diffusion problem reads: Find $u \in V$ such that

$$\mathcal{A}_h(u,v) \equiv a_h(u,v) + \frac{1}{\varepsilon}\int_{\mathbb{T}_h} \boldsymbol{\Sigma} u \cdot v dx = \int_{\mathbb{T}_h} S \cdot v dx, \qquad \forall v \in V_h, \tag{2.3}$$

where S is a right-hand side or *source* term. We observe that, given the non-negativeness of $\boldsymbol{\Sigma}$, the coercivity constant for our problem coincides with the Laplacian case while the continuity constant is now dependent on ε. In order to obtain a robust solver, we precondition the problem to be able to bound the spectral radius of the preconditioned system independently of ε.

Finally, using a standard basis for the local finite element spaces on each cell and by concatenating, we obtain the linear system

$$\boldsymbol{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{u}$ and $\mathbf{f}$ are the coefficient vectors of the representation of u and f, respectively, in terms of the chosen basis.

3. Preconditioners. In this section we provide details of our solver and the preconditioner choice, as well as the technical tools needed for the numerical analysis of the preconditioned system.

It is known that the convergence of the preconditioned conjugate gradient method for symmetric real operators depends on the condition number of the preconditioned matrix only. Thus, if we find a preconditioner such that this condition number is independent of h and of the parameters of the equation, then the number of iterations required for convergence to a certain error is independent of them as well. We will estimate the condition number of the additive Schwarz method by estimating the smallest and largest eigenvalues c_{ad} and C_{ad} as

$$c_{\text{ad}} = \inf_{v\neq 0} \frac{\mathcal{A}_h(\mathcal{P}_{\text{ad}} v, v)}{\|v\|^2_{\mathcal{A}_h}} \qquad \text{and} \qquad C_{\text{ad}} = \sup_{v\neq 0} \frac{\mathcal{A}_h(\mathcal{P}_{\text{ad}} v, v)}{\|v\|^2_{\mathcal{A}_h}}.$$

For the rest of the preconditioners, we will estimate the norm of the error propagation operator of a preconditioned Richardson iteration.

3.1. Schwarz preconditioners. We choose Schwarz preconditioners, for which there is a well-known framework and theory for symmetric positive definite problems; see [8, 12, 19, 21]. The following sections provide the definitions needed to prove the convergence estimates in an abstract formulation.

Let V_j, for $j = 0, 1, 2, \ldots, J$, be Hilbert spaces with corresponding norms $\|\cdot\|_{V_j}$, where V_0 is used to denote the so-called *coarse space* in a domain decomposition context. For $j = 0, 1, 2, \ldots, J$, let

$$\mathcal{R}_j^{\intercal} : V_j \to V_h$$

denote *prolongation operators* for which there holds

$$\mathcal{R}_j^{\intercal} V_j \subset V \quad \text{and} \quad V = \sum_{j=0}^{J} \mathcal{R}_j^{\intercal} V_j, \qquad \text{for } j = 0, 1, 2, \ldots, J.$$

Here $\mathcal{R}_j^{\intercal} V_j$ is the range of the linear operator $\mathcal{R}_j^{\intercal}$. Associated with each local space V_j, for $j = 1, 2, \ldots, J$, we introduce the local discrete bilinear forms $\mathcal{A}_j(\cdot,\cdot)$, defined on $V_j \times V_j$, as the restriction of the global discrete bilinear form $\mathcal{A}_h(\cdot,\cdot)$ to $V_j \times V_j$, with $\|v_j\|^2_{\mathcal{A}_j} = \mathcal{A}_j(v_j, v_j)$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

For the coarse space V_0, we use the rediscretization of the problem on the coarse mesh, namely a bilinear form $\mathcal{A}_0(\cdot,\cdot)$ with a penalty parameter inversely proportional to the diameter of the coarse cells H, instead of the inherited coarse space obtained by the restriction to $V_0 \times V_0$. For any fixed $v \in V_0$, we define a *projection-like* operator $\widetilde{\mathcal{P}}_0 v \in V_0$ by

$$\mathcal{A}_0(\widetilde{\mathcal{P}}_0 v, w_0) := \mathcal{A}_h(v, \mathcal{R}_0^\intercal w_0), \qquad \forall w_0 \in V_0,$$

and the composite operator as $\mathcal{P}_0 := \mathcal{R}_0^\intercal \widetilde{\mathcal{P}}_0$.

The convergence analysis of our method follows the standard framework for subspace correction methods (see for instance [19, 21]), which is based on three main assumptions:

ASSUMPTION 3.1 (Stable decomposition). *The spaces $\{V_j\}$ are said to provide a stable decomposition if there exists a constant C_V such that each $v \in V_h$ admits a decomposition*

$$v = \sum_{j=0}^{J} \mathcal{R}_j^\intercal v_j,$$

with $v_j \in V_j$, such that

$$\sum_{j=0}^{J} \|v_j\|_{\mathcal{A}_j}^2 \leq C_V \|v\|_{\mathcal{A}_h}^2,$$

where $\|v\|_{\mathcal{A}_h}^2 = \mathcal{A}_h(v,v)$ and $\|v\|_{\mathcal{A}_j}^2$ defined accordingly.

If $v \in$ range $(\mathcal{I} - \mathcal{P}_0)$, then $v \in V_h$ admits a stable decomposition without including the coarse space as follows (see [21, p. 49]):

$$\sum_{j=1}^{J} \|v_j\|_{\mathcal{A}_j}^2 \leq C_V \|v\|_{\mathcal{A}_h}^2.$$

ASSUMPTION 3.2 (Strengthened Cauchy-Schwarz inequality). *There exist constants $\theta_j \in [0,1]$, for $i,j = 0,1,2,\ldots,J$, such that*

$$\mathcal{A}_h(\mathcal{R}_i^\intercal v_i, \mathcal{R}_j^\intercal v_j) \leq \theta_{ij} \mathcal{A}_h(\mathcal{R}_i^\intercal v_i, \mathcal{R}_i^\intercal v_i)^{\frac{1}{2}} \mathcal{A}_h(\mathcal{R}_j^\intercal v_j, \mathcal{R}_j^\intercal v_j)^{\frac{1}{2}}, \qquad \forall v_i \in V_i, v_j \in V_j.$$

We denote the spectral radius of $\Theta = \{\theta_{ij}\}$ by $\rho(\Theta)$.

ASSUMPTION 3.3 (Local stability). *There exists $\omega \in [1,2)$ such that*

$$\mathcal{A}_h(\mathcal{R}_j^\intercal v_j, \mathcal{R}_j^\intercal v_j) \leq \omega \mathcal{A}_j(v_j, v_j), \qquad \forall v_j \in V_j.$$

We now introduce a set of *projection-like* operators $\widetilde{\mathcal{P}}_j : V_h \to V_j$, for $j = 0,1,2,\ldots,J$. These projection-like operators will serve as the building blocks for the construction of Schwarz methods. For any fixed $v \in V_h$, define $\widetilde{\mathcal{P}}_j v \in V_j$ by

$$\mathcal{A}_j(\widetilde{\mathcal{P}}_j v, w_j) := \mathcal{A}_h(v, \mathcal{R}_j^\intercal w_j), \qquad \forall w_j \in V_j.$$

We note that the well-posedness of the global problem ensures that $\widetilde{\mathcal{P}}_j$ is well defined for $j = 0,1,2,\ldots,J$. To map the elements of V_j into the global discrete space V_h, we employ the *prolongation operator* $\mathcal{R}_j^\intercal$ and define the composite operator

$$\mathcal{P}_j := \mathcal{R}_j^\intercal \circ \widetilde{\mathcal{P}}_j, \qquad \text{for } j = 0,1,2,\ldots,J.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Trivially, we have $\mathcal{P}_j : V_h \to V_h$, for $j = 0, 1, 2, \ldots, J$. After these preparations, we can write the operator $\mathcal{A}_h$ preconditioned with the additive Schwarz method as

$$\mathcal{P}_{\text{ad}} := \mathcal{P}_0 + \mathcal{P}_1 + \mathcal{P}_2 + \cdots + \mathcal{P}_J.$$

To facilitate the comprehension of the method with respect to its implementation, we write the additive operator in a more explicit form. We use the operator notation for the bilinear forms $\mathcal{A}_h$ and $\mathcal{A}_j$ to obtain the following expression for the local projections:

$$\mathcal{A}_j \widetilde{\mathcal{P}}_j v := \mathcal{R}_j \mathcal{A}_h v, \qquad \forall v \in V_h.$$

Thus,

$$\widetilde{\mathcal{P}}_j = \mathcal{A}_j^{-1} \mathcal{R}_j^{\intercal} \mathcal{A}_h, \qquad \text{and} \qquad \mathcal{P}_j = \mathcal{R}_j^{\intercal} \mathcal{A}_j^{-1} \mathcal{R}_j \mathcal{A}_h.$$

Finally, our additive Schwarz preconditioned system reads

$$\mathcal{P}_{\text{ad}} = \mathcal{R}_0^{\intercal} \mathcal{A}_0^{-1} \mathcal{R}_0 \mathcal{A}_h + \sum_{j=1}^{J} \mathcal{R}_j^{\intercal} \mathcal{A}_j^{-1} \mathcal{R}_j \mathcal{A}_h.$$

While the additive version applies all subspace corrections at once and adds them at the end, the multiplicative version applies them successively. It can be defined easily by the *error propagation* operator

$$\mathcal{E}_{\text{mu}} = (\mathcal{I} - \mathcal{P}_N) \circ (\mathcal{I} - \mathcal{P}_{N-1}) \circ \cdots \circ (\mathcal{I} - \mathcal{P}_0),$$

where $\mathcal{I}$ denotes the identity operator on V. Using $\mathcal{E}_{\text{mu}}$, we define the multiplicative Schwarz preconditioner as

$$\mathcal{P}_{\text{mu}} = \mathcal{I} - \mathcal{E}_{\text{mu}},$$

where $\mathcal{I}$ denotes the identity operator on V_h.

Finally, we consider the symmetric hybrid version, which is additive with respect to the subdomain spaces, but applies the coarse grid correction in a multiplicative way:

$$\mathcal{P}_{\text{hy}} = \mathcal{I} - \left(\mathcal{I} - \sum_{i=1}^{N} \mathcal{P}_i \right) (\mathcal{I} - \mathcal{P}_0) \left(\mathcal{I} - \sum_{i=1}^{N} \mathcal{P}_i \right).$$

In the following, we prove convergence estimates for the operators $\mathcal{P}_{\text{ad}}$, $\mathcal{P}_{\text{hy}}$, and $\mathcal{P}_{\text{mu}}$. For $\mathcal{P}_{\text{ad}}$, we estimate the condition number, for $\mathcal{P}_{\text{mu}}$ we bound the error operator of a preconditioned Richardson iteration, and for $\mathcal{P}_{\text{hy}}$ we defer the proof to Section 3.3, where we study multigrid preconditioners, of which $\mathcal{P}_{\text{hy}}$ is a special case. We use the general abstract convergence theory of Schwarz methods given in [21, §2]. We quote the convergence results below.

THEOREM 3.4. *Let the Assumptions 3.1, 3.2, and 3.3 hold. Then the following bounds hold for the additive Schwarz preconditioned system:*

$$c_{ad} \geq \frac{1}{C_V}, \qquad C_{ad} \leq \omega(\rho(\Theta) + 1),$$

where c_{ad} and C_{ad} are the smallest and largest eigenvalues of the preconditioned system, respectively.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. See [21, §2.3]. ∎

THEOREM 3.5. *Let the Assumptions 3.1, 3.2, and 3.3 hold. Then the following bounds hold for the hybrid Schwarz preconditioned system:*

$$|\mathcal{A}_h\left([\mathcal{I} - \mathcal{P}_{hy}]\, v, v\right)| \leq c_{MG}\mathcal{A}_h(v,v), \qquad \forall v \in V_h,$$

where c_{MG} is a constant independent of h and ε.

Proof. We defer this proof to Section 3.3 as it is a special case of a multigrid preconditioner, and as such its convergence estimate is given in Theorem 3.15. ∎

The multiplicative operator is not symmetric, and we will consider a simple Richardson iteration applied to the corresponding preconditioned system and provide an upper bound for the norm of the error propagation operator.

THEOREM 3.6. *Let the Assumptions 3.1, 3.2, and 3.3 hold. Then the following bounds hold for the multiplicative Schwarz preconditioned system:*

$$\|\mathcal{E}_{mu}\| \leq 1 - \frac{2-\omega}{\left(2\max\{1,\omega^2\}\rho^2(\Theta)+1\right)C_V} \leq 1.$$

Proof. See [21, §2.3]. ∎

3.2. Application to the discrete problem. In this section we define the Schwarz method for the discrete problem in equation (2.3) and verify that Assumptions 3.1, 3.2, and 3.3 apply.

After enumerating the cells $\kappa_j \in \mathbb{T}_h$, for $j = 1, \ldots, J$, we choose the local spaces $V_j = V(\kappa_j) = \mathbb{Q}_p^G(\kappa_j)$, together with the coarse space V_0 defined on $\mathbb{T}_H$. We remark that we are using a nonoverlapping subdivision in order to define the direct decomposition $V_h = \bigoplus_{j=1}^{J} \mathcal{R}_j^{\intercal} V_j$, where $\mathcal{R}_j^{\intercal}\colon V_j \to V_h$ is the simple injection. Similarly, for $v \in V_h$, $\mathcal{R}_j v(x) = v(x)$ if $x \in \kappa_j$ and zero otherwise. In the following, we list three standard results from [12] that we need for our proof.

For any $v \in V_D = \prod_{K\in\mathbb{T}_H} \mathcal{V}(K)$, there holds the trace inequality (see [12, Lemma 3.1]),

$$\|v\|^2_{\mathcal{H}(\partial D)} \leq c\left[\frac{1}{H}\|v\|^2_{\mathcal{H}(D)} + H\|v\|^2_{\mathcal{V}(D)}\right]. \tag{3.1}$$

Suppose D is a convex domain. For any $v \in V_D$, let $\overline{u} = \frac{1}{\operatorname{meas}(D)}\int_D v dx$ be the average value of v over D. Then we have a Poincaré inequality as follows (see [12, Lemma 3.2]):

$$\|v - \overline{v}\|_{\mathcal{H}(D)} \leq c\,\operatorname{diam}(D)\|u\|_{\mathcal{V}(D)} \qquad \text{on } D.$$

In particular, if $D \in \mathbb{T}_H$, then

$$\|v - \overline{v}\|_{\mathcal{H}(D)} \leq cH\|u\|_{\mathcal{V}(D)} \qquad \text{on } D. \tag{3.2}$$

Let $v, w \in V_h$, and let $v_j, w_j \in V_j$, $j = 1, \ldots, J$, be given (uniquely) by $v = \sum_{j=1}^{J} v_j$, $w = \sum_{j=1}^{J} w_j$. Then the following identity holds (see [12, Lemma 3.3]):

$$a_h(v,w) = \sum_{j=1}^{J} a_j(v_j, w_j) + I(v,w), \tag{3.3}$$

where $I(\cdot,\cdot) : V_h \times V_h \to \mathbb{R}$ comprises all terms located outside the block-diagonal of the bilinear from $a_h(v,w)$, connecting different subdomains.

We then obtain the following interface estimate for cell-wise subdomains:

ETNA
Kent State University and
Johann Radon Institute (RICAM)

LEMMA 3.7. *There exists a constant c such that*

$$|I(v,v)| \leq c \left[\frac{1}{h^2} \sum_{K \in \mathbb{T}_h} \|v\|^2_{\mathcal{H}(K)} + a_h(v,v) \right].$$

Proof. We extend the result of [12, Lemma 4.3]. The following estimate from [12, Equation (4.20)] holds when using cell-wise subdomains:

$$|I(v,v)| \leq c \left(a_h(v,v) + \frac{1}{h} \sum_{F \in \left(\mathbb{F}_h^I \cup \mathbb{F}_h^B\right)} \|v\|^2_{\mathcal{H}(F)} \right),$$

where $\|\cdot\|_F$ is the L^2-inner product on the faces of the cell K of the fine mesh.

Using the trace inequality $\|v\|^2_{\mathcal{H}(F)} = c \left[\frac{1}{h} \|v\|^2_{\mathcal{H}(K)} + h \|\nabla v\|^2_{\mathcal{H}(K)} \right]$ from [12, Equation (3.9)], we obtain

$$|I(v,v)| \leq c \left(a_h(v,v) + \frac{1}{h} \sum_{K \in \mathbb{T}_h} \left[\frac{1}{h} \|v\|^2_{\mathcal{H}(K)} + h \|\nabla v\|^2_{\mathcal{H}(K)} \right] \right).$$

The result follows by observing that $\sum_{K \in \mathbb{T}_h} \|\nabla v\|^2_{\mathcal{H}(K)} \leq c\, a_h(v,v)$. ∎

Finally, we concentrate on a stable decomposition. The convergence theory from [12] requires that the subdomains used for the Schwarz method are at least the same size as the cells in the coarse mesh. Recently, an extension has been published in [10] to include the case of cell-wise subdomains, however, the proof uses P_1-nonconforming interpolant and enriching operators for simplices [7].

We achieve a stable decomposition by a close examination of the proof in [12], which holds for simplices, quadrilaterals, and hexahedra. In particular, it does not require auxiliary spaces with continuity assumptions like for Crouzeix-Raviart elements.

LEMMA 3.8. *Every $v \in V_h$ admits a decomposition of the form $v = \sum_{j=0}^{J} \mathcal{R}_j^\intercal V_j$, $v_j \in V_j$, $j = 0, \ldots, J$, which satisfies the bound*

$$\sum_{i=0}^{J} a_j\left(v_j, v_j\right) \leq C_{V,\Delta}\, a\left(v,v\right),$$

with $C_{V,\Delta} = \mathcal{O}\left(\frac{H^2}{h^2}\right)$, where h and H denote the cell diameters used in the fine and coarse meshes, respectively.

Proof. Let $v_0 \in V_0$ be the piecewise constant average of v on the coarse mesh $\mathbb{T}_H$, and let $w = v - \mathcal{R}_0^\intercal v_0$. We decompose w into nonoverlapping cell-wise subdomains as

$$w = \sum_{j=1}^{J} v_j,$$

where $v_1, \ldots, v_J$ are uniquely determined. From equation (3.3) we have

$$a_h(w,w) = \sum_{j=1}^{J} a_j(v_j, v_j) + I(w,w),$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

or equivalently,

$$a_h(v-\mathcal{R}_0^\intercal v_0, v-\mathcal{R}_0^\intercal v_0) = \sum_{j=1}^{J} a_j(v_j,v_j) + I(v-\mathcal{R}_0^\intercal v_0, v-\mathcal{R}_0^\intercal v_0).$$

Reordering and estimating the interface term by its absolute value, we obtain

$$\sum_{j=1}^{J} a_j(v_j,v_j) \leq a_h(v-\mathcal{R}_0^\intercal v_0, v-\mathcal{R}_0^\intercal v_0) + \left|I(v-\mathcal{R}_0^\intercal v_0, v-\mathcal{R}_0^\intercal v_0)\right|.$$

Using Lemma 3.7 we have

$$\begin{aligned}\sum_{j=1}^{J} a_j(v_j,v_j) \leq & c\left(a\left(v-\mathcal{R}_0^\intercal v_0, v-\mathcal{R}_0^\intercal v_0\right) + \frac{1}{h^2}\sum_{K\in\mathbb{T}_h} \|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{H}(K)}\right)\\ \leq & c\Bigg(\left(a_h(v,v)^{1/2} + a\left(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0\right)^{1/2}\right)^2\\ & \qquad + \frac{1}{h^2}\sum_{D\in\mathbb{T}_H} \|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{H}(D)}\Bigg),\end{aligned}$$

where we used Minkowsky's inequality and regrouped the L^2-inner products. We expand the first term and use equation (3.2) to obtain

$$\begin{aligned}\sum_{j=1}^{J} a_j(v_j,v_j) \leq & c\Big(a_h(v,v) + 2a_h(v,v)^{1/2}a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0)^{1/2}\\ & \qquad + a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) + \frac{H^2}{h^2}\|v\|^2_{\mathcal{V}}\Big)\\ \leq & c\left(2a_h(v,v) + 2a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) + \frac{H^2}{h^2}a_h(v,v)\right),\end{aligned}$$

where we used Young's inequality and the coercivity of $a_h(\cdot,\cdot)$. Finally, including the coarse space, we achieve

$$\sum_{j=0}^{J} a_j(v_j,v_j) \leq c\left(a_0(v_0,v_0) + a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) + \frac{H^2}{h^2}a_h(v,v)\right).$$

It remains to bound $a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0)$ in a such way that the estimate is independent of the usage of cell-wise or larger subdomains and that a constant $\mathcal{O}(\frac{H}{h})$ is achieved, as we show below. Since v_0 is piecewise constant on $\mathbb{T}_H$ and hence also on $\mathbb{T}_h$,

$$a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) = \delta_{\text{IP}} \sum_{F\in\mathbb{F}_h^I} \frac{1}{h}\|\mathcal{R}_0^\intercal v_0^+ - \mathcal{R}_0^\intercal v_0^-\|^2_{\mathcal{H}(F)} + \delta_{\text{IP}} \sum_{F\in\mathbb{F}_h^B} \frac{1}{h}\|\mathcal{R}_0^\intercal v_0^+\|^2_{\mathcal{H}(F)}, \tag{3.4}$$

where we observe that

$$a_0(v_0,v_0) = \frac{h}{H}a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0). \tag{3.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Adding and subtracting v in equation (3.4) gives

$$a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) \leq c\delta_{\text{IP}} \Bigg(\sum_{F\in\mathbb{F}_h^I} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+ - (v-\mathcal{R}_0^\intercal v_0)^-\|^2_{\mathcal{H}(F)} + \sum_{F\in\mathbb{F}_h^B} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+\|^2_{\mathcal{H}(F)} + \sum_{F\in\mathbb{F}_h^I} \frac{1}{h}\|v^+ - v^-\|^2_{\mathcal{H}(F)} + \sum_{F\in\mathbb{F}_h^B} \frac{1}{h}\|v^+\|^2_{\mathcal{H}(F)} \Bigg).$$

The last two terms are obviously bounded by $a_h(v,v)$. Also, since u_0 is piecewise constant on each $D \in \mathbb{T}_H$,

$$\|(v-\mathcal{R}_0^\intercal v_0)^+ - (v-\mathcal{R}_0^\intercal v_0)^-\|_{\mathcal{H}(F)} = \|v^+ - v^-\|_{\mathcal{H}(F)}$$

whenever F is in the interior of some $D \in \mathbb{T}_H$. Thus,

$$\begin{aligned}
&\sum_{F\in\mathbb{F}_h^I} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+ - (v-\mathcal{R}_0^\intercal v_0)^-\|^2_{\mathcal{H}(F)} + \sum_{F\in\mathbb{F}_h^B} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+\|^2_{\mathcal{H}(F)} \\
&\quad = \sum_{D\in\mathbb{T}_H} \Bigg(\sum_{F\subset D} \|v^+ - v^-\|_{\mathcal{H}(F)} + \sum_{F\in\partial D} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+ - (v-\mathcal{R}_0^\intercal v_0)^-\|^2_{\mathcal{H}(F)} + \sum_{F\subset\partial D\in\mathbb{F}_h^B} \frac{1}{h}\|(v-\mathcal{R}_0^\intercal v_0)^+\|^2_{\mathcal{H}(F)} \Bigg) \\
&\quad \leq c a_h(v,v) + c \sum_{D\in\mathbb{T}_H} \frac{1}{h}\|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{H}(\partial D)}.
\end{aligned}$$

Now using the trace inequality in equation (3.1), we obtain

$$\sum_{D\in\mathbb{T}_H} \frac{1}{h}\|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{H}(\partial D)} \leq c \sum_{D\in\mathbb{T}_H} \frac{1}{h}\left[\frac{1}{H}\|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{H}(D)} + H\|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{V}(D)}\right].$$

Also note that $\|v-\mathcal{R}_0^\intercal v_0\|^2_{\mathcal{V}(D)} = \|v\|^2_{\mathcal{V}(D)}$. Hence, applying the approximation result from equation (3.2) to $\|v-v_0\|_{\mathcal{H}(D)}$, we find

$$a_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) \leq c\frac{H}{h} a_h(v,v).$$

Therefore, using this result for equation (3.5), we see that $a_0(v_0,v_0) \leq c a_h(v,v)$, and hence, the result is achieved. ∎

LEMMA 3.9 (Stable decomposition). *The spaces V_j provide a stable decomposition of V with respect to the bilinear form $\mathcal{A}_h(\cdot,\cdot)$ in the sense of Assumption 3.1.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. Let $C_{V,\Delta}$ be the stable decomposition constant for the Laplacian, as deduced in Lemma 3.8. We then have

$$\begin{aligned}\sum_{i=0}^{J} \mathcal{A}_j\left(v_j, v_j\right) &= \sum_{i=0}^{J} \left\{ a_j\left(v_j, v_j\right) + \frac{1}{\varepsilon}\left(\mathbf{\Sigma} v_j, v_i\right)_{\mathcal{H}} \right\} \\ &= \sum_{i=0}^{J} a_j\left(v_j, v_j\right) + \frac{1}{\varepsilon}\left(\mathbf{\Sigma} v, v\right)_{\mathcal{H}} \\ &\leq C_{V,\Delta} a\left(v, v\right) + \frac{1}{\varepsilon}\left(\mathbf{\Sigma} v, v\right)_{\mathcal{H}} \\ &\leq \max\left\{C_{V,\Delta}, 1\right\} \mathcal{A}_h\left(v, v\right).\end{aligned}$$

It follows that the V_j-decomposition for our reaction-diffusion problem is *energy stable* with $C_V = C_{V,\Delta} = \mathcal{O}\left(\frac{H^2}{h^2}\right)$, where H and h are the largest and smallest cell diameters, respectively. ∎

LEMMA 3.10. *There exists a strengthened Cauchy-Schwarz inequality in the sense of Assumption 3.2.*

Proof. (See [12, §4.2].) Verifying this inequality involves deriving a bound for the spectral radius $\rho(\Theta)$ of the $J \times J$ matrix $\Theta = [\theta_{ij}]_{j=0}^{J}$. That such a value exists is a consequence of the Cauchy-Schwarz inequality. The important point, however, is to establish a small bound for ρ. To do so, we observe that $a_h(\mathcal{R}_i^\intercal v_i, \mathcal{R}_j^\intercal v_j) = 0$ if the supports of v_i and v_j do not share a face f_{ij}. For the remaining cases, we take $\theta_{ij} = 1$. It follows at once from Gershgorin's circle theorem that

$$\rho(\Theta) \leq \max_m \operatorname{card}\left\{k | f_{mk} \neq 0 \text{ almost everywhere}\right\} + 1, \qquad f_{mk} \in \mathbb{F}_h^I \cup \mathbb{F}_h^B,$$

i.e., the spectral radius $\rho(\Theta)$ is bounded by 1 plus the maximum number of adjacent subdomains that a given subdomain can have. In practice, this number is 4 in 2D and 6 in 3D. Even for "unusual" subdomain partitions, this number is not expected to be large. ∎

LEMMA 3.11 (Local Stability). *There holds that*

$$\mathcal{A}_h(\mathcal{R}_j^\intercal v_j, \mathcal{R}_j^\intercal v_j) \leq \omega \mathcal{A}_j(v_j, v_j), \qquad \forall v_j \in V_j,$$

where $\omega = \alpha \frac{H}{h}$ *for* $\alpha < 1$.

Proof. In the case of exact local solvers $\omega = 1$, as in our case, the coarse bilinear form uses a penalty parameter depending on the cell diameter of the coarse mesh. Inspecting the bilinear form (2.2), we observe that for our coarse space bilinear form, it holds that

$$\mathcal{A}_h(\mathcal{R}_0^\intercal v_0, \mathcal{R}_0^\intercal v_0) \leq \frac{H}{h} \mathcal{A}_0(v_0, v_0),$$

and hence, our local stability constant would be $\omega = \frac{H}{h}$. However, this would violate Assumption 3.3. To remedy this, we scale the bilinear forms with a relaxation parameter α in order to accomplish the required upper bound. We can always introduce such a relaxation parameter, but we are not free to scale the local bilinear forms arbitrarily in order to decrease C_V from Lemma 3.9; a small value of ω means that the corrections of the error are small. In such a case, C_V will necessarily be large; see [19, p. 155] and [21, p. 41]. Finally, we remark that this is only required for our proofs, but in practice such a relaxation parameter is not necessary. ∎

ETNA
Kent State University and
Johann Radon Institute (RICAM)

3.3. Multigrid V-cycle preconditioner. The preconditioners developed in the previous section are easily implemented as smoothers for multigrid preconditioners. In this section we provide convergence estimates for the multigrid V-cycle.

Let $\{\mathbb{T}\}_{\ell=0,\dots,L}$ be a hierarchy of meshes of quadrilateral and hexahedral cells in two and three dimensions, respectively. In view of multilevel methods, the index ℓ refers to the mesh level defined as follows: let a coarse mesh $\mathbb{T}_0$ be given. Then, the mesh hierarchy is defined recursively such that the cells of $\mathbb{T}_{\ell+1}$ are obtained by splitting each cell of $\mathbb{T}_\ell$ into 2^d children by connecting the edge and face midpoints (refinement). These meshes are nested in the sense that every cell of $\mathbb{T}_\ell$ is equal to the union of its four (respectively eight) children. We define the mesh size h_ℓ as the maximum of the diameters of the cells of $\mathbb{T}_\ell$. Due to this refinement process, we have $h_\ell \approx 2^{-1} h_{\ell-1}$.

From the nestedness of mesh cells, the finite element spaces associated with these meshes are nested as well:

$$V_0 \subset V_1 \subset \cdots \subset V_L.$$

We introduce the L^2-projections $\mathcal{Q}_{\ell-1}$ and the embedding operators $\mathcal{Q}_{\ell-1}^{\intercal}$

$$\begin{aligned} \mathcal{Q}_{\ell-1} &: V_\ell \to V_{\ell-1}, \\ \mathcal{Q}_{\ell-1}^{\intercal} &: V_{\ell-1} \to V_\ell, \end{aligned}$$

such that

$$(\mathcal{Q}_{\ell-1} v_\ell, w_{\ell-1})_{\mathcal{H}} = \left(v_\ell, \mathcal{Q}_{\ell-1}^{\intercal} w_{\ell-1}\right)_{\mathcal{H}}, \qquad \forall v_{\ell-1} \in V_{\ell-1}, w_{\ell-1} \in V_{\ell-1}.$$

Let $\mathcal{A}_\ell(\cdot,\cdot)$ be the bilinear form defined in equation (2.3) on the mesh $\mathbb{T}_\ell$. We define the operator $\mathcal{A}_\ell : V_\ell \longrightarrow V_\ell$ such that $\mathcal{A}_\ell(u_\ell, v_\ell) = (\mathcal{A}_\ell u_\ell, v_\ell)_{\mathcal{H}}$. For the rest of the paper, we will redefine the operators $\mathcal{P}$ used in the 2-level analysis as follows: $\mathcal{P}_{\ell-1}$ is what is used to be the coarse grid solver $\mathcal{P}_0$, while $\mathcal{P}_{\ell,j}$ represent the projections onto the subdomain spaces $V_j = V_{\ell,j}$ on the mesh level ℓ. It holds that $\mathcal{A}_{\ell-1}\mathcal{P}_{\ell-1} = \mathcal{Q}_{\ell-1}\mathcal{A}_\ell$.

Let $\mathcal{B}_\ell$ be a smoother defined as the preconditioning operator for the preconditioned systems presented in Section 3.1 without including the coarse space, i.e.,

$$\mathcal{B}_{\ell,\text{ad}} = \sum_{i=1}^{N_\ell} \mathcal{P}_{\ell,i}\mathcal{A}_\ell^{-1} = \sum_{i=1}^{N_\ell} \mathcal{R}_{\ell,i}^{\intercal}\mathcal{A}_{\ell,i}^{-1}\mathcal{R}_{\ell,i} \qquad \text{and} \qquad \mathcal{B}_{\ell,\text{mu}} = \left(\mathcal{I} - \prod_{i=N_\ell}^{1} \mathcal{P}_{\ell,i}\right)\mathcal{A}_\ell^{-1}.$$

We define the multigrid preconditioner $\mathcal{M}_L$ by induction. Let $\mathcal{M}_0 = \mathcal{A}_0^{-1}$. For $1 \le \ell \le L$, we define the action $\mathcal{M}_\ell g$ of $\mathcal{M}_\ell$ on a vector $g \in V_\ell$ in terms of $\mathcal{M}_{\ell-1}$ as follows:

1. Let $x_0 = 0$.
2. Define x_i, for $i = 1, \dots, m$, by m pre-smoothing steps

$$x_i = x_{i-1} + \mathcal{B}_\ell\left(g - \mathcal{A}_\ell x_{i-1}\right).$$

3. Define y_0 by a coarse grid correction

$$y_0 = x_m + \mathcal{Q}_{\ell-1}^{\intercal}\mathcal{M}_{\ell-1}\mathcal{Q}_{\ell-1}\left(g - \mathcal{A}_\ell x_m\right).$$

4. Define y_i, for $i = 1, \dots, m$, by the m post-smoothing steps

$$y_i = y_{i-1} + \mathcal{B}_\ell\left(g - \mathcal{A}_\ell x_{i-1}\right).$$

5. Let $\mathcal{M}_\ell g = y_m$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Our analysis of the multigrid algorithm follows [11] since we have noninherited forms. There, convergence is proven in an abstract framework under the following three assumptions:

ASSUMPTION 3.12 (Stability). *There is a constant $C_Q > 0$ such that for all levels $\ell = 2, \dots, L$ and all $v_\ell \in V_\ell$,*

$$\mathcal{A}_\ell \left(\left[\mathcal{I}_\ell - \mathcal{Q}_{\ell-1}^\intercal \mathcal{P}_{\ell-1} \right] v_\ell, \left[\mathcal{I}_\ell - \mathcal{Q}_{\ell-1}^\intercal \mathcal{P}_{\ell-1} \right] v_\ell \right) \leq C_Q \mathcal{A}_\ell(v_\ell, v_\ell).$$

ASSUMPTION 3.13 (Regularity-approximation property). *There is a constant $C_1 > 0$, such that for all levels $\ell = 2, \dots, L$ and all $v_\ell \in V_\ell$,*

$$\mathcal{A}_\ell \left(\left[\mathcal{I}_\ell - \mathcal{Q}_{\ell-1}^\intercal \mathcal{P}_{\ell-1} \right] v_\ell, v_\ell \right) \leq C_1 \frac{\|\mathcal{A}_\ell v_\ell\|_{L^2}^2}{\Lambda_\ell},$$

where Λ_ℓ is the maximum eigenvalue of $\mathcal{A}_\ell$.

ASSUMPTION 3.14 (Smoothing property). *There is a constant $C_R > 0$ such that for all levels $\ell = 2, \dots, L$ and all $v_\ell \in V_\ell$,*

$$\frac{\|v_\ell\|_{L^2}^2}{\Lambda_\ell} \leq C_R(\overline{R} v_\ell, v_\ell),$$

where $\overline{R} = (\mathcal{I} - \mathcal{K}_\ell^2)\mathcal{A}_\ell^{-1}$ and $\mathcal{K}_\ell = \mathcal{I} - \mathcal{B}_\ell \mathcal{A}_\ell$.

From [11] we obtain the estimate for the error propagation operator defined as $\mathcal{I} - \mathcal{M}_\ell \mathcal{A}_\ell$.

THEOREM 3.15. *Let Assumptions 3.12, 3.13, and 3.14 hold. Furthermore, assume $m > 2C_1C_R$. Then, for all $\ell \geq 0$, there holds*

$$|\mathcal{A}_\ell \left(\left[\mathcal{I} - \mathcal{M}_\ell \mathcal{A}_\ell \right] v_\ell, v_\ell \right)| \leq c_{MG} \mathcal{A}_\ell(v_\ell, v_\ell), \qquad \forall v_\ell \in V_\ell,$$

with

$$c_{MG} = \frac{C_1 C_R}{m + C_1 C_R}$$

for the two-level method, i.e., $\mathcal{P}_{hy}$, and

$$c_{MG} = \frac{C_1 C_R}{m - C_1 C_R}$$

for $L > 2$.

We refer to [11] for the proof in an abstract setting; we show below that the assumptions apply to our method. Assumption 3.13 is proven in [1, Theorem 9] and Assumption 3.14 in [6, Theorem 5.1]. To prove Assumption 3.12, we use Lemma 3.11 as follows:

$$\begin{aligned}
&\mathcal{A}_\ell \left(\mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell, \mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell \right) \leq 2\mathcal{A}_{\ell-1} \left(\mathcal{P}_{\ell-1} v_\ell, \mathcal{P}_{\ell-1} v_\ell \right) \\
&\mathcal{A}_\ell \left(\mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell, \mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell \right) \leq 2\mathcal{A}_\ell \left(v_\ell, \mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell \right) \\
&\mathcal{A}_\ell \left(\mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell, \mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell \right) - 2\mathcal{A}_\ell \left(v_\ell, \mathcal{Q}_\ell^\intercal \mathcal{P}_{\ell-1} v_\ell \right) + \mathcal{A}_\ell \left(v_\ell, v_\ell \right) \leq \mathcal{A}_\ell \left(v_\ell, v_\ell \right),
\end{aligned}$$

and we deduce that

$$\mathcal{A}_\ell \left(\left[\mathcal{I} - \mathcal{Q}_{\ell-1}^\intercal \mathcal{P}_{\ell-1} \right] v_\ell, \left[\mathcal{I} - \mathcal{Q}_{\ell-1}^\intercal \mathcal{P}_{\ell-1} \right] v_\ell \right) \leq \mathcal{A}_\ell \left(v_\ell, v_\ell \right).$$

Hence, Assumption 3.12 holds with $C_Q = 1$.

We note that the preceding theorem requires $m > 1$ for $L > 2$, but as we will see in the next section, $m = 1$ suffices for our setting. For completeness, we provide the results for $m > 1$ as well.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

4. Numerical experiments. Noting that some of the smoothers and preconditioners that we use are not symmetric, we apply a GMRES solver for all our calculations.

All our experiments are performed on a unit square $\Omega = (0,1) \times (0,1)$ with Dirichlet boundary conditions. As our work is centered on the behavior of the system with respect to the reaction-term, we set all diffusion coefficients to one. Each mesh $\mathbb{T}_\ell$ consists of $2^\ell \times 2^\ell$ cells such that each cell is a square with side length $h_\ell = 2^{-\ell}$. We use bilinear elements and $\delta_0 = 2$ and $h_\ell / h_{\ell-1} = 1/2$ for all our experiments. When using multigrid V-cycles, the coarsest mesh consists of a single cell.

4.1. Poisson's equation. As a baseline for further experiments, we present the results for Poisson's equation using different preconditioners; see Table 4.1. We observe that all preconditioners achieve a flat iteration count, albeit with a different number of iterations at the very fine levels. The two-level additive Schwarz method, for instance, requires almost double the number of iterations than the multigrid method with additive Schwarz preconditioners.

TABLE 4.1
GMRES iterations for a DG discretization of Poisson's equation using tensor product polynomials of degree 1 and a unit source to reduce the residual by 10^{-8} for $\Sigma = 0$. Here, U denotes an unpreconditioned solver, 2AS, 2HS, 2MS are two-level additive, hybrid, and multiplicative Schwarz solvers, respectively, and MGAS, MGMS are multigrid solver with additive and multiplicative Schwarz smoothers, respectively.

levels	U	2AS	2HS	2MS	MGAS	MGMS
2	3	3	3	4	3	4
3	10	10	6	6	6	6
4	22	18	9	7	10	7
5	43	24	11	7	12	8
6	85	26	11	7	13	8
7	> 100	25	11	7	14	8
8	> 100	25	11	7	14	8

4.2. Two groups. In case of a two group problem, because of the conservation condition of a zero column sum and symmetry, all reaction matrices are multiples of

$$\boldsymbol{\Sigma} = \frac{1}{\varepsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

We display the iteration results in Table 4.2.

We observe that the iteration count flattens for all considered methods, with numbers very similar to the pure Laplacian problem, which indicates that the reaction operator does not affect the results presented in the previous section. The fact that the results do not improve is explained by the reaction operator having a non-trivial kernel, where we effectively solve for the Laplacian.

4.3. Multigroup. We devise a reaction matrix with a *contrast* between the coefficients in different groups that is inversely proportional to various powers of ε as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{1,1} & -1 & -\varepsilon^{-1} & -\varepsilon^{-2} & -\varepsilon^{-3} & \dots \\ -1 & \Sigma_{2,2} & -1 & -1 & -1 & \dots \\ -\varepsilon^{-1} & -1 & \Sigma_{3,3} & -\varepsilon^{-1} & -\varepsilon^{-2} & \dots \\ -\varepsilon^{-2} & -1 & -\varepsilon^{-1} & \Sigma_{4,4} & -\varepsilon^{-1} & \dots \\ -\varepsilon^{-3} & -1 & -\varepsilon^{-2} & -\varepsilon^{-1} & \Sigma_{5,5} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.2
GMRES iterations using a source $(1,0)$ *or* $(0,1)$ *to reduce the residual by* 10^{-8}. *Here "max" is the maximum number of iterations for different* ε.

levels \ ε	MGAS 1.0	MGAS 10^{-1}	MGAS 10^{-2}	MGAS 10^{-3}	MGAS 10^{-4}	MGMS max	2AS max	2HS max	2MS max
2	5	5	4	4	4	4	6	5	4
3	8	8	6	6	6	6	14	8	6
4	10	10	10	10	10	7	22	10	7
5	12	12	12	12	12	8	25	11	7
6	13	13	13	13	13	8	25	11	7
7	14	14	14	14	14	8	25	11	7
8	14	14	14	14	14	8	25	11	7
9	14	14	14	14	14	8	25	11	7

where

$$\Sigma_{g,g} = -\sum_{g' \neq g} \Sigma_{g,g'} = 1 + \varepsilon^{-1} + \varepsilon^{-2} + \varepsilon^{-3} + \dots$$

We remark that the elements in the diagonal are such that the matrix has zero column sum. We use the top left 5×5-block of this matrix as the reaction matrix in the following tests.

The results are presented in Table 4.3; the numerical tests were performed for the sources $(1,0,1,0,1)$, $(0,1,0,1,0)$, $(0,1,1,1,0)$, and $(1,0,0,0,1)$, and we report the maximum iteration count encountered. In this experiment, the columns are shown only up to $\varepsilon = 0.01$ to avoid floating point underflow problems. Note, that this involves values of $\varepsilon^{-3} = 10^{-6}$. It can be observed, that the iteration count flattens as in the other cases, and the performance of the method is unaffected by the increase in the number of groups or their different scaling.

TABLE 4.3
GMRES iterations to reduce the residual by 10^{-8} *for a 5-groups calculation, where "max" is the maximum number of iterations over the values of* ε *in the left columns.*

levels \ ε	MGAS 1.0	MGAS 0.1	MGAS 0.01	MGMS max	2AS max	2HS max	2MS max
2	5	5	4	4	9	5	4
3	8	7	6	6	15	8	6
4	10	10	10	7	22	10	7
5	12	12	12	8	25	11	7
6	13	13	13	8	26	11	7
7	14	14	14	8	25	11	7
8	14	14	14	8	25	11	7
9	14	14	14	8	25	11	7

We also display the results for the use of more than one pre- and postsmoothing steps in Table 4.4. We observe an improvement in the iteration count, always flattening, that becomes less significant as the number of smoothing iterations increases, which suggests that there is a sweet spot to be found with regards to the computational cost.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.4
GMRES iterations to reduce the residual by 10^{-8} for a 5-groups calculation, with different number of smoothings per level, where "max" is the maximum number of iterations for different ε.

levels \ ε	2 smoothings			4 smoothings			8 smoothings		
	1.0	0.1	0.01	1.0	0.1	0.01	1.0	0.1	0.01
2	4	3	3	3	2	2	2	2	2
3	5	5	5	4	4	4	3	3	3
4	7	7	7	5	5	5	4	4	4
5	8	8	8	6	6	6	5	5	5
6	9	9	9	7	7	7	6	6	6
7	9	9	9	7	7	7	7	7	7
8	9	9	9	7	7	7	6	7	7
9	9	9	9	6	7	7	6	7	7

4.4. Space dependent reaction matrix. We modify the matrix used in the previous section by scaling it with the following function depending only on the space variables:

$$f_i(x,y) = \begin{cases} (x,y) \in \mathbf{\Omega}_i & \sin^2(2\pi x)\sin^2(2\pi y), \\ (x,y) \notin \mathbf{\Omega}_i & 0. \end{cases}$$

Here $\mathbf{\Omega}_i$, with $i = 0,1,2,3$, are the four quadrants of the square domain. Note that these results in reaction- and diffusion-dominated regions and inertial subspaces in the group space depending on the spatial coordinates,

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{1,1} & -f_0 & -\varepsilon^{-1}f_1 & -\varepsilon^{-2}f_2 & -\varepsilon^{-3}f_3 & \dots \\ -f_0 & \Sigma_{2,2} & -f_0 & -f_0 & -f_0 & \dots \\ -\varepsilon^{-1}f_1 & -f_0 & \Sigma_{3,3} & -\varepsilon^{-1}f_1 & -\varepsilon^{-2}f_2 & \dots \\ -\varepsilon^{-2}f_2 & -f_0 & -\varepsilon^{-1}f_1 & \Sigma_{4,4} & -\varepsilon^{-1}f_1 & \dots \\ -\varepsilon^{-3}f_3 & -f_0 & -\varepsilon^{-2}f_2 & -\varepsilon^{-1}f_1 & \Sigma_{5,5} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The results are shown in Table 4.5 for different source terms as in the previous section. In this case, the columns are given only up to $\varepsilon = 0.01$ to avoid floating point underflow. We see that once again we achieve a flat iteration count, with a slightly larger absolute value for the finest meshes. The reaction term does not affect the convergence of the method even when the reaction coefficients vary in space as well as between the groups.

5. Conclusions. We have introduced a domain decomposition smoother based on the solution of the complete reaction-diffusion system on each cell of the mesh in the fashion of additive or multiplicative nonoverlapping Schwarz methods. We prove that these smoothers produce two-level and multilevel preconditioners that perform robustly with respect to the mesh size and the parameters of the equation. Our numerical experiments confirm the robustness and show that the obtained iteration counts are indeed low and thus that the methods are very efficient.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.5
GMRES iterations to reduce the residual by 10^{-8} for a 5-groups calculation, where "max" is the maximum number of iterations for different ε.

	MGAS			MGMS	2AS	2HS	2MS
levels \ ε	1.0	0.1	0.01	max	max	max	max
2	6	7	6	4	19	7	4
3	9	10	9	6	22	10	6
4	11	12	12	7	25	11	7
5	13	13	13	8	27	12	8
6	13	14	14	8	28	12	8
7	14	14	15	8	28	13	8
8	14	15	15	9	27	12	8
9	14	15	15	9	27	12	8
10	15	15	15	9	27	12	8
11	15	15	15	9	27	12	8
12	15	15	15	9	27	12	8

REFERENCES

[1] P. F. ANTONIETTI AND G. PENNESI, *V-cycle multigrid algorithms for discontinuous Galerkin methods on non-nested polytopic meshes*, J. Sci. Comput., 78 (2019), pp. 625–652.
[2] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.
[3] D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. MARINI, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2001/02), pp. 1749–1779.
[4] G. A. BAKER, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 137 (1977), pp. 45–59.
[5] I. BOGLAEV, *On a domain decomposition algorithm for a singularly perturbed reaction-diffusion problem*, J. Comput. Appl. Math., 98 (1998), pp. 213–232.
[6] J. H. BRAMBLE, *Multigrid Methods*, Longman, Harlow, 1993.
[7] S. BRENNER, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal., 41 (2003), pp. 306–324.
[8] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 2002.
[9] R. DAUTRAY AND J.-L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 2. Functional and Variational Methods*, Springer, Berlin, 1985.
[10] M. DRYJA AND P. KRZYŻANOWSKI, *A massively parallel nonoverlapping additive Schwarz method for discontinuous Galerkin discretization of elliptic problems*, Numer. Math., 132 (2016), pp. 347–367.
[11] H.-Y. DUAN, S.-Q. GAO, R. C. E. TAN, AND S. ZHANG, *A generalized BPX multigrid framework covering nonnested V-cycle methods*, Math. Comp, 76 (2007), pp. 137–152.
[12] X. FENG AND O. KARAKASHIAN, *Two-level non-overlapping Schwarz methods for a discontinuous Galerkin method*, SIAM J. Numer. Anal., 39 (2001), pp. 1343–1365.
[13] R. A. HORN, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[14] G. KANSCHAT AND J. P. LUCERO LORCA, *A weakly penalized discontinuous Galerkin method for radiation in dense, scattering media*, Comput. Methods Appl. Math., 16 (2016), pp. 563–577.
[15] A. J. LEW AND G. C. BUSCAGLIA, *A discontinuous-Galerkin-based immersed boundary method*, Internat. J. Numer. Methods Engrg., 76 (2008), pp. 427–454.
[16] H. MACMULLEN, J. MILLER, E. O'RIORDAN, AND G. SHISHKIN, *A second-order parameter-uniform overlapping Schwarz method for reaction–diffusion problems with boundary layers*, J. Comput. Appl. Math., 130 (2001), pp. 231–244.
[17] T. A. MANTEUFFEL AND K. J. RESSEL, *Least-squares finite-element solution of the neutron transport equation in diffusive regimes*, SIAM J. Numer. Anal., 35 (1998), pp. 806–835.
[18] J. NITSCHE, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, 36 (1970), pp. 9–15.
[19] B. F. SMITH, P. E. BJØRSTAD, AND W. D. GROPP, *Domain Decomposition*, Cambridge University Press, Cambridge, 1996.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[20] M. STEPHENS AND N. MADDEN, *A parameter-uniform Schwarz method for a coupled system of reaction-diffusion equations*, J. Comput. Appl. Math., 230 (2009), pp. 360–370.
[21] A. TOSELLI AND O. B. WIDLUND, *Domain Decomposition Methods–Algorithms and Theory*, Springer, Berlin, 2005.
[22] M. F. WHEELER, *An elliptic collocation finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.