
Developments in Overlapping Schwarz Preconditioning of High-Order Nodal Discontinuous Galerkin Discretizations

L. N. Olson¹, J. S. Hesthaven¹, and L. C. Wilcox¹

Division of Applied Mathematics, Brown University, 182 George Street, Box F,
Providence, RI 02912, USA. Luke.Olson@brown.edu, Jan.Hesthaven@brown.edu,
lucasw@dam.brown.edu

Summary. Recent progress has been made to more robustly handle the increased complexity of high-order schemes by focusing on the local nature of the discretization. This locality is particularly true for many Discontinuous Galerkin formulations and is the focus of this paper. The contributions of this paper are twofold. First, novel observations regarding various flux representations in the discontinuous Galerkin formulation are highlighted in the context of overlapping Schwarz methods. Second, we conduct additional experiments using high-order elements for the indefinite Helmholtz equation to expose the impact of overlap.

1 Introduction

We consider the Helmholtz equation

$$-\nabla \cdot \nabla u(\mathbf{x}) - \omega^2 u(\mathbf{x}) = f(\mathbf{x}) \quad \text{in } \Omega, \quad (1a)$$

$$u(\mathbf{x}) = g(\mathbf{x}) \quad \text{on } \Gamma. \quad (1b)$$

Although the form presented in (1) is evidently straightforward, it does still expose a number of difficulties that we discuss in this paper. The problem turns cumbersome quickly as the wave number increases since the resulting system of equations becomes indefinite. Identifying the key components to efficiently solving this wave problem will likely carry over into more complicated situations, such as Maxwell's equations.

The approach taken in this paper is an overlapping Schwarz-type method. The method presented is motivated by efforts from a variety of researchers who have outlined a number situations where Schwarz methods have proved to be effective: indefinite problems, discontinuous Galerkin discretizations, and high-order elements [4, 2, 3, 8, 9]. Based on this previously detailed success, we study the performance of a preconditioned additive Schwarz method that utilizes element overlap to maintain efficient performance as the order

of the discontinuous spectral element method increases and as indefiniteness becomes more prominent.

2 DG

The LDG formulation we follow yields several advantageous properties in the resulting linear system of equations. The global mass matrix is block diagonal, allowing cheap inversion, while symmetry is preserved in the global discretization matrix.

We begin by considering an admissible, shape regular triangulation \mathcal{K} of $\Omega \in \mathbb{R}^2$ and let $h_\kappa = 1/2 \cdot \text{diam}(\kappa)$, for $\kappa \in \mathcal{K}$. The numerical approximation u_h on element $\kappa \in \mathcal{K}_h$ is composed of Lagrange interpolating polynomials $L_j(\mathbf{x})$ at selected degrees of freedom \mathbf{x}_j within κ . In 1-D, we describe these locations as the Gauss-Lobatto-Legendre (GLL) quadrature points. Similarly, for our 2-D reference triangle, $\hat{\kappa}$, we choose a distribution of nodes governed by electrostatics [6]. $N_\kappa = \frac{(n+1)(n+2)}{2}$ points are needed to ensure an order n resolution in the local polynomial approximation on element κ . Figure 1 shows an example on the reference element. Finally, we define $\mathcal{P}_n(\kappa)$, the local spectral element space where we seek an approximation.

The standard LDG formulation [1] is described first by introducing a slack variable $\mathbf{q} = \nabla u$. The first-order system for (1) on an arbitrary element κ is

$$-\nabla \cdot \mathbf{q} - \omega^2 u = f \quad \text{in } \kappa, \quad (2a)$$

$$\mathbf{q} - \nabla u = 0 \quad \text{in } \kappa. \quad (2b)$$

Multiplying each equation by scalar and vector test functions $\phi(\mathbf{x})$ and $\boldsymbol{\psi}(\mathbf{x})$, respectively, and integrating by parts yields the weak formulation. The local traces of u and \mathbf{q} are replaced by approximations—*numerical fluxes*— u^* and \mathbf{q}^* , respectively. With this substitution and integrating by parts again, the associated (and slightly stronger) weak discrete problem is: Find $(u_{h,n}, \mathbf{q}_{h,n})$ such that

$$-\int_\kappa \nabla \cdot \mathbf{q}_{h,n} \phi_n \, d\mathbf{x} - \omega^2 \int_\kappa u_{h,n} \phi_n \, d\mathbf{x} = \int_\kappa f_{h,n} \phi_n \, d\mathbf{x} + \int_{\partial\kappa} \mathbf{n}_k \cdot (\mathbf{q}^* - \mathbf{q}_{h,n}) \phi_n \, d\mathbf{x}, \quad (3a)$$

$$\int_\kappa \mathbf{q}_{h,n} \cdot \boldsymbol{\psi}_n \, d\mathbf{x} - \int_\kappa \nabla u_{h,n} \cdot \boldsymbol{\psi}_n \, d\mathbf{x} = \int_{\partial\kappa} (u^* - u_{h,n}) \mathbf{n}_k \cdot \boldsymbol{\psi} \, d\mathbf{x}, \quad (3b)$$

for all $\kappa \in \mathcal{K}_h$ and $(\phi_n, \boldsymbol{\psi}_n)$. The function spaces are the local spectral element spaces defined using Lagrange interpolation above.

Defining the numerical flux is what separates different discontinuous Galerkin approaches [1] and is the most distinguishing feature of a formulation since the interelement connectivity is solely defined by the representation of the numerical flux on each edge. This choice directly impacts the approximation properties as well as the stability of the method. Moreover, the resulting

(global) linear system of equations will perhaps exhibit symmetry and varying sparsity patterns depending on how the trace is approximated along each edge of each element in the tessellation. For a given element κ , define u^- to be the value of u interior to the element and define u^+ to be the value of u in the adjacent, neighboring element. For a scalar function u and vector function \mathbf{q} , the *jump* and the *average* between neighboring elements are respectively defined as $\llbracket u \rrbracket = u^- \mathbf{n}^- + u^+ \mathbf{n}^+$, $\{u\} = \frac{1}{2}(u^- + u^+)$, $\llbracket \mathbf{q} \rrbracket = \mathbf{q}^- \cdot \mathbf{n}_{k^-} + \mathbf{q}^+ \cdot \mathbf{n}_{k^+}$, $\{\mathbf{q}\} = \frac{1}{2}(\mathbf{q}^- + \mathbf{q}^+)$. For $\kappa \in \mathcal{K}$ with $\partial\kappa \in \Gamma_{\text{bdy}}$, these values adjusted by extending the solution to a ghost element.

By defining the numerical fluxes u^* and \mathbf{q}^* independent of ∇u , we will be able to formulate the weak problem (3) independent of the slack variable $\mathbf{q}(\mathbf{x})$. In general, the numerical fluxes for the LDG method are defined as [1]

$$u^* = \{u_{n,h}\} + \beta \cdot \llbracket u_{n,h} \rrbracket \quad \mathbf{q}^* = \{\mathbf{q}_{n,h}\} - \beta \llbracket \mathbf{q}_{n,h} \rrbracket - \eta_k \llbracket u_{n,h} \rrbracket. \quad (4)$$

The sign on β is specifically opposite to ensure symmetry of the associated stiffness matrix [1]. Adhering to this form of a numerical flux is beneficial since the method is consistent and locally conservative. Further, if $\eta_k > 0$ the method is considered stable [1]. Setting $\beta = 0$ yields a *central* flux for u^* and a stabilized central flux for \mathbf{q}^* , while using $\beta = 0.5\mathbf{n}^-$ results in an upwinding scheme. The impact computationally is addressed in Section 4.

The numerical flux u^* is independent of $\mathbf{q}_{h,n}$ allowing us to write the discrete system completely independent of the slack variable \mathbf{q} -c.f. lifting operators in [1]. As we sum the weak problem over all elements $\kappa \in \mathcal{K}$ we will need the following global matrices: S^x , S^y , and M , which are stiffness and mass matrices and $F_{u^*}^{x,y}$ and $F_{q^*}^{x,y}$, which couple nodes in adjacent elements. Introducing global data vectors $\tilde{\mathbf{q}}^x$, $\tilde{\mathbf{q}}^y$, and $\tilde{\mathbf{u}}$ and summing the weak problem (3) over all elements $\kappa \in \mathcal{K}$, we arrive at the following

$$-S^x \tilde{\mathbf{q}}^x - S^y \tilde{\mathbf{q}}^y - \omega^2 M \tilde{\mathbf{u}} = M \mathbf{f} + F_{q^*}^x \tilde{\mathbf{q}}^x + F_{q^*}^y \tilde{\mathbf{q}}^y - \tau F_{q^*}^T \tilde{\mathbf{u}}, \quad (5)$$

$$M \tilde{\mathbf{q}}^x - S^x \tilde{\mathbf{u}} = F_{u^*}^x \tilde{\mathbf{u}}, \quad (6)$$

$$M \tilde{\mathbf{q}}^y - S^y \tilde{\mathbf{u}} = F_{u^*}^y \tilde{\mathbf{u}}. \quad (7)$$

Solving for the slack variable $\tilde{\mathbf{q}}^{x,y}$ in equations (6) and (7), and substituting into (5) eliminates the dependence on $\tilde{\mathbf{q}}$. The system, written in compact form is then

$$(-S + F - \omega^2 M) \tilde{\mathbf{u}} = M \mathbf{f}, \quad (8)$$

where $S = S^x M^{-1} S^x + S^y M^{-1} S^y$ and $F = F_{q^*}^x M^{-1} S^x + F_{q^*}^x M^{-1} F_{u^*}^x + F_{q^*}^y M^{-1} S^y + F_{q^*}^y M^{-1} F_{u^*}^y - \tau F_{u^*}^T$. The operator S is clearly negative semi-definite, while for $\tau > 0.0$, the composite operator $S - F$ is strictly negative definite. A full eigenspectrum analysis and the impact on the preconditioner is unknown. However, it suffices to say that for moderate ω , indefinite and near singular matrices should be expected.

3 PAS

Extensive work by Cai et al. [4, 2, 3] and Elman [5] conclude that standard Krylov based iterative methods handle a moderate number of flipped eigenvalues quite well for this indefinite problem. We will also use this class of methods and, in particular, choose the Generalized Minimum Residual method (GMRES). GMRES can be applied to indefinite systems and, more importantly, the preconditioned implementation permits indefinite preconditioning matrices. This will be beneficial in the case of the preconditioned additive Schwarz (PAS) method. It noteworthy that BiCGStab yielded slightly improved results in our tests, but the observed trends remained the same.

Our implementation is a culmination of approaches, which includes overlapping subdomains and a coarse grid solution phase with the ability to handle non-nested coarse grids. It is important to note that a global coarse solve does not improve the convergence process if the grid is not rich enough to fully resolve a wave. There are a couple notable features about our approach. First, given a coarse grid tessellation, Ω^H , and subdomain $\Omega_s^h \subset \Omega^h$, we define the restriction operator based on standard finite element interpolation as $R_{0ij}^T = \phi_i(\mathbf{x}_j)$. Here, $\phi_i(\mathbf{x})$ is a coarse grid basis function (bilinear in our case) and \mathbf{x}_j is a node in Ω_s on the fine grid. $R_{0ij} = 0$ if \mathbf{x}_j is not in the underlying footprint of ϕ_i and is thus still sparse, although not in comparison to the injection operators used in the subdomain solves. To efficiently implement this process, let V be the Vandermonde matrix built from our orthogonal set of polynomials: $V_{i,j} = p_j(\mathbf{x}_i)$. With this we can transfer between modal and nodal representations easily with $\mathbf{f} = V\hat{\mathbf{f}}$ and $\hat{\mathbf{f}} = V^{-1}\mathbf{f}$ and since V^{-1} can be built locally in preprocessing. The advantage is clear when we look at more general interpolation in this respect. Let V_{cc} be the coarse basis/coarse nodes Vandermonde matrix and V_{cf} be the coarse basis/fine nodes Vandermonde matrix. Then $P_0 = V_{cf}V_{cc}^{-1} \equiv R_0^T$ defines the equivalent interpolation operator at the expense of only a few calculations. Second, in order to ensure proper interpolation of constant solutions, we incorporate a row equilibration technique, by rescaling each row of R_0 by the row sum:

$$R_{0ij} \leftarrow \frac{1}{\sum_j R_{0ij}} R_{0ij}. \quad (9)$$

The composite preconditioning matrix is then defined to be $M^{-1} = R_0^T A_0^{-1} R_0 + \sum_{s=1}^S R_s^T A_s^{-1} R_s$.

Overlap is also introduced in our algorithm. This increases communication, but, as we present in the next section, is an essential component particularly for high-order approximations and as the matrix increases in indefiniteness and size. We define $\delta = 0$ to be the case with no geometric overlap, keeping in mind the nature of the discontinuous discretization, where degrees of freedom in neighboring elements may share a geometric location, resulting in some resemblance of overlap. By increasing δ , we simply mean that each subdomain

is padded by δ layers of elements. At first glance, this may seem extreme, since Fischer and Lottes [9] extend only by strips of nodes into the adjacent elements. However, the class of problems we address is altogether different, requiring large numbers of elements, and requiring only moderate polynomial degrees, making overlap overhead costs small as the mesh is further refined. Moreover, layers of nodes within an electrostatic distribution are not readily available either in the element itself or in the reference element, where they have a straightforward formation in the case of tensor-based element.

4 Numerics

Using the central flux in the DG method is more correctly termed the Brezzi method [1]. Due to the ease of implementation, this formulation has grown in popularity, also benefiting from slightly improved conditioning over a *bona fide* LDG method where $\beta = 0.5\mathbf{n}^-$. Unfortunately, if $\beta = 0$, the data from elements κ^+ is needed to describe equations (3) in element κ^- as well as data from the neighbors of κ^+ , which we label κ^{++} . Thus the influence on one element extends two layers beyond a given element. The noncompact stencil is also prevalent for $\beta \neq 0$, unless $\beta = 0.5\mathbf{n}^-$, which corresponds to upwind flux. This is considered the LDG method since fortuitous cancellation of the terms eliminates the extension to neighboring elements, resulting in a stencil width of only one layer. Figure 1 articulates this effect. A more detailed explanation of the effects on discretization error and the eigenspectrum can be found in [7], although convergence of the iterative solution process is not addressed.

Also shown in Figure 1 is the so-called Interior Penalty method (IP). Here, a local gradient is used in the definition of the flux, which also results in a compact stencil. The IP method offers a straightforward implementation, however the poor conditioning of this approach requires careful attention. Table 1 illustrates a typical situation. The results are presented for the definite case ($\omega = 0.0$) on a grid with $h \approx 1/8$. A single level additive Schwarz scheme is used to precondition the GMRES acceleration. The first column reiterates the fact that the Brezzi approach ($\beta = 0.0$) has slightly better conditioning than the LDG implementation ($\beta = 0.5\mathbf{n}^-$), while the IP system suffers from a very poor spectrum. Column 2 also provides insight, showing that while the LDG scheme is slightly more ill-conditioned, the local type preconditioning scheme is more effective due to the compact stencil. The Brezzi operator responds similarly under preconditioning, but due to the wide stencil, the relative improvement is not as drastic. The preconditioning also has significant influence on the IP method, but due to the poor conditioning, it is difficult to fully quantify the effect of PAS. We will follow the Brezzi method throughout the rest of the paper since it is a widely used formulation of DG and since we expect the preconditioning results to be on the pessimistic side. A more comprehensive study of the various DG methods and preconditioning, similar to Table 1, is an ongoing research effort.

Table 1. GMRES iterations for Brezzi, LDG, and IP formulations with and without preconditioning.

N	Brezzi		LDG		IP	
	w/o PAS	w/ PAS	w/o PAS	w/ PAS	w/o PAS	w/ PAS
2	73	21	121	21	355	57
4	167	28	252	29	1291	151
6	316	30	456	32	> 2000	294
8	534	38	713	36	> 2000	568

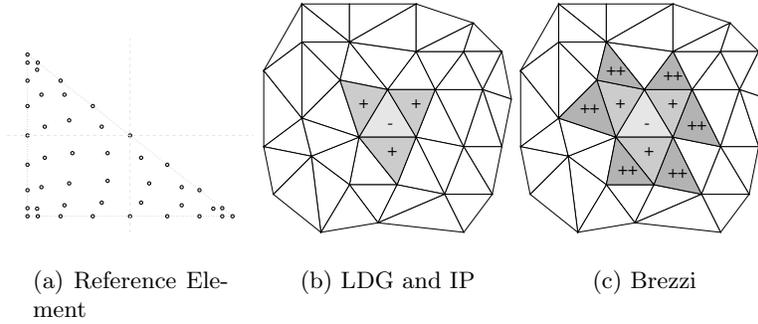


Fig. 1. Stencil width relative to element κ^- .

Our test problem is basic, yet still exposes a principle difficulty: indefiniteness and high-order discretizations. We consider a smooth, solution $u(x, y) = \sin(2\pi\omega x) \sin(2\pi\omega y)$.

Comparing the iterations in Table 2 indicates that a coarse grid is beneficial for high-order discretizations. GMRES iterations are reduced for each polynomial order when using a richer coarse grid. It is interesting to further note that the relative improvement is consistent as the order is increased. Overlap, however, has a much larger impact on the convergence of the preconditioned iterative method as indicated in Table 2.

Table 2. GMRES iterations with $h_f \approx 1/8$, $\omega = 1.0$: adding overlap

h_c	$\delta = 0$										$\delta = 1$												
	order n																						
0	1	2	3	4	5	6	7	8	9	10	11												
	26	38	49	60	71	82	93	105	116	128	140												
1/4	22	32	39	50	58	67	72	81	88	100	108	→	22	22	23	24	24	25	25	26	26	27	28
1/8	14	25	30	36	43	47	55	60	66	73	79												

As the frequency ω increases, more degrees of freedom are needed to fully resolve the solution. When the problem is viewed on a coarser grid, the dis-

cretization lacks resolution and the solution found on the coarse grid no longer resembles an accurate approximation to the fine grid solution. Thus the two-level error correction becomes ineffective and possibly pollutes the fine grid solution. Figure 2 shows that the iteration counts remain bounded as the polynomial order is increased for each selected ω . The iterations increase as the frequency is increased, but this is expected as more low eigenvalues are shifted to the positive half-plane. As expected, coarse solves do not improve solution for large wave numbers, however there is significant improvement as we introduce overlap, particularly for the case of the highly indefinite problem, $\omega = 50$.

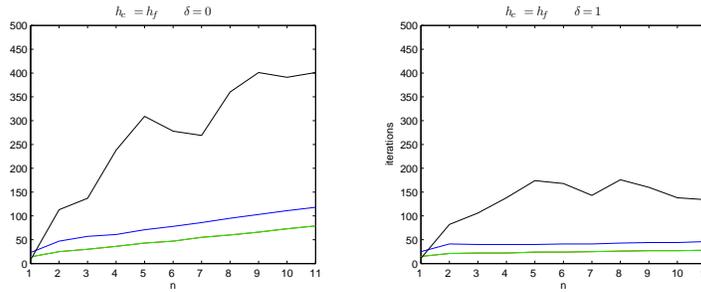


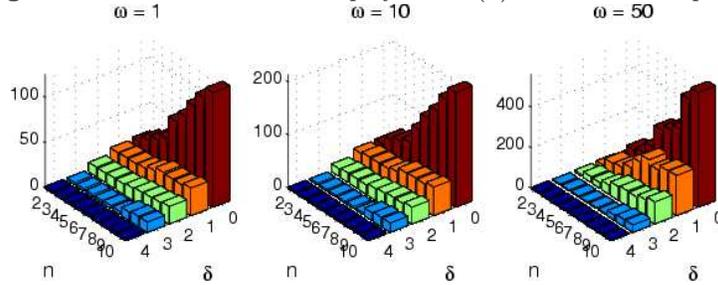
Fig. 2. GMRES iterations versus polynomial order: Comparing overlap impact for $\omega = 1.0, 10.0, 50.0$.

A more definitive exposé is to test problems where the discretization is not under or over resolved. Referring to dispersion analysis, using around several degrees of freedom per wavelength (in 1-D) is generally considered well resolved. Table 3 confirms the importance of overlap. The relative improvement as n increases is attributed to the fact that larger subdomain solves are being used. The trend in overlap continues only so far. Figure 3 illustrates that performance is improved as the overlap is increased, however the relative impact becomes less.

Table 3. GMRES iterations: $h_f \approx 1/4$, no coarse grid

n	ω	No PAS $\delta = 0$ $\delta = 1$		
		avg. iterations		
1	0 ... 7	48	23	18
2	6 ... 10	106	43	27
3	9 ... 13	170	57	30
4	12 ... 16	271	72	36
5	15 ... 20	392	106	48
6	19 ... 23	534	151	67
7	22 ... 26	705	193	72

Fig. 3. GMRES iterations versus polynomial (n) order and overlap (δ).



References

1. D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779 (electronic), 2001/02.
2. X.-C. Cai. A family of overlapping Schwarz algorithms for nonsymmetric and indefinite elliptic problems. In *Domain-based parallelism and problem decomposition methods in computational science and engineering*, pages 1–19. SIAM, Philadelphia, PA, 1995.
3. X.-C. Cai, M. A. Casarin, F. W. Elliott, Jr., and O. B. Widlund. Overlapping Schwarz algorithms for solving Helmholtz’s equation. In *Domain decomposition methods, 10 (Boulder, CO, 1997)*, volume 218 of *Contemp. Math.*, pages 391–399. Amer. Math. Soc., Providence, RI, 1998.
4. X.-C. Cai and O. B. Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13(1):243–258, 1992.
5. H. C. Elman, O. G. Ernst, and D. P. O’Leary. A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations. *SIAM J. Sci. Comput.*, 23(4):1291–1315 (electronic), 2001.
6. J. S. Hesthaven. From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex. *SIAM J. Numer. Anal.*, 35(2):655–676 (electronic), 1998.
7. R. M. Kirby. *Toward dynamic spectral/hp refinement: algorithms and applications to flow-structure interactions*. PhD thesis, Brown University, May 2003.
8. C. Lasser and A. Toselli. Overlapping preconditioners for discontinuous Galerkin approximations of second order problems. In *Domain decomposition methods in science and engineering (Lyon, 2000)*, Theory Eng. Appl. Comput. Methods, pages 78–84. Internat. Center Numer. Methods Eng. (CIMNE), Barcelona, 2002.
9. J. W. Lottes and P. F. Fischer. Hybrid multigrid/schwarz algorithms for the spectral element method. Technical Report ANL/MCS-P1052-0403, Mathematics and Computer Science Division, Argonne National Laboratory, April 2003.