

Mathematical Institute

Sparse vertex-star relaxation for high-order FEM

Pablo Brubeck, Patrick Farrell

Mathematical Institute University of Oxford

Conference on fast direct solvers October 24, 2021

Oxford Mathematics

Incompressible Neo-Hookean hyperelasticity Unstructured mesh, 1280 cells, $Q_p^d \times DQ_{p-2}$, p = 31, d = 2, 3





















p-multigrid and Fast Diagonalization in Firedrake PETSc solver options



OXFORD

p-multigrid and Fast Diagonalization in Firedrake PETSc solver options



OXFORD



Solve small problems and combine the solutions $\tilde{A}^{-1} = \sum_{j=1}^{J} R_j^{\top} A_j^{-1} R_j$.



Cell-centered patches with fixed overlap layers have decreasing overlap measure as $p \rightarrow \infty$, which deteriorates the convergence rate.



Pavarino (1994) proved that the additive Schwarz method with vertex-centered patches (generous overlap) gives *p*-independent convergence when the coarse space is of the lowest order ($V_c = Q_1$).

For separable problems, the Fast Diagonalization Method (1964) is a $O(p^{d+1})$ direct solver. Cannot diagonalize arbitrary vertex patches.

First FDM preconditioner used for static condensation by Couzy (1995).



Pavarino (1994) proved that the additive Schwarz method with vertex-centered patches (generous overlap) gives *p*-independent convergence when the coarse space is of the lowest order ($V_c = Q_1$).

For separable problems, the Fast Diagonalization Method (1964) is a $O(p^{d+1})$ direct solver. Cannot diagonalize arbitrary vertex patches.

First FDM preconditioner used for static condensation by Couzy (1995).



Pavarino (1994) proved that the additive Schwarz method with vertex-centered patches (generous overlap) gives *p*-independent convergence when the coarse space is of the lowest order ($V_c = Q_1$).

For separable problems, the Fast Diagonalization Method (1964) is a $O(p^{d+1})$ direct solver. Cannot diagonalize arbitrary vertex patches.

First FDM preconditioner used for static condensation by Couzy (1995).



Pavarino (1994) proved that the additive Schwarz method with vertex-centered patches (generous overlap) gives *p*-independent convergence when the coarse space is of the lowest order ($V_c = Q_1$).

For separable problems, the Fast Diagonalization Method (1964) is a $O(p^{d+1})$ direct solver. Cannot diagonalize arbitrary vertex patches.

First FDM preconditioner used for static condensation by Couzy (1995).



Pavarino (1994) proved that the additive Schwarz method with vertex-centered patches (generous overlap) gives *p*-independent convergence when the coarse space is of the lowest order ($V_c = Q_1$).

For separable problems, the Fast Diagonalization Method (1964) is a $O(p^{d+1})$ direct solver. Cannot diagonalize arbitrary vertex patches.

First FDM preconditioner used for static condensation by Couzy (1995).



Find $u \in V = H_0^1(\Omega)$ such that

$$\int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d} \mathbf{x} = \int_{\Omega} v f \, \mathrm{d} \mathbf{x} \quad \forall v \in V.$$

Expand $u_h = \sum_j u_j \phi_j$ and assemble the stiffness matrix and RHS

$$A_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, \mathrm{d} \mathbf{x}, \quad b_i = \int_{\Omega} \phi_i f \, \mathrm{d} \mathbf{x}.$$

For Cartesian domains, A has a special tensor product structure:

$$A = \begin{cases} B_y \otimes A_x + A_y \otimes B_x & d = 2, \\ B_z \otimes B_y \otimes A_x + B_z \otimes A_y \otimes B_x + A_z \otimes B_y \otimes B_x & d = 3. \end{cases}$$

Here $A_* = \mu_* \hat{A}$, $B_* = \hat{B}$ for * = x, y, z.



Find $u \in V = H_0^1(\Omega)$ such that

$$\int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d} \mathbf{x} = \int_{\Omega} v f \, \mathrm{d} \mathbf{x} \quad \forall v \in V.$$

Expand $u_h = \sum_j u_j \phi_j$ and assemble the stiffness matrix and RHS

$$A_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, \mathrm{d}\mathbf{x}, \quad b_i = \int_{\Omega} \phi_i f \, \mathrm{d}\mathbf{x}.$$

For Cartesian domains, A has a special tensor product structure:

$$A = \begin{cases} B_y \otimes A_x + A_y \otimes B_x & d = 2, \\ B_z \otimes B_y \otimes A_x + B_z \otimes A_y \otimes B_x + A_z \otimes B_y \otimes B_x & d = 3. \end{cases}$$

Here $A_* = \mu_* \hat{A}$, $B_* = \hat{B}$ for * = x, y, z.



Find $u \in V = H_0^1(\Omega)$ such that

$$\int_{\Omega} \nabla u \cdot \nabla v \, \mathrm{d} \mathbf{x} = \int_{\Omega} v f \, \mathrm{d} \mathbf{x} \quad \forall v \in V.$$

Expand $u_h = \sum_j u_j \phi_j$ and assemble the stiffness matrix and RHS

$$A_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, \mathrm{d}\mathbf{x}, \quad b_i = \int_{\Omega} \phi_i f \, \mathrm{d}\mathbf{x}.$$

For Cartesian domains, A has a special tensor product structure:

$$A = \begin{cases} B_y \otimes A_x + A_y \otimes B_x & d = 2, \\ B_z \otimes B_y \otimes A_x + B_z \otimes A_y \otimes B_x + A_z \otimes B_y \otimes B_x & d = 3. \end{cases}$$

Here $A_* = \mu_* \hat{A}$, $B_* = \hat{B}$ for * = x, y, z.



FDM (Lynch, Rice & Thomas, 1964)

Structured matrix inversion analogous to separation of variables. Breaks problems down into a sequence of 1D eigenvalue problems. Direct $O(p^{d+1})$ solver for Poisson in very simple geometries.

$$A_*S_* = B_*S_*\Lambda_*$$
 for $* = x, y, z$

$$A^{-1} = (S_z \otimes S_y \otimes S_x) \Lambda^{-1} (S_z \otimes S_y \otimes S_x)^\top.$$









We split the DOFs $\{I, \Gamma\}$. On the interior *I*, solve the 1D eigenproblems:

$$\hat{A}_{II}\hat{S}_{II}=\hat{B}_{II}\hat{S}_{II}\Lambda_{II}.$$

We construct the FDM basis functions

$$\hat{S} = \begin{bmatrix} \hat{S}_{II} & -B_{II}^{-1}B_{I\Gamma} \\ 0 & \mathbb{1} \end{bmatrix}$$

such that the 1D matrices

$$\hat{S}^{\top}\hat{A}\hat{S}, \quad \hat{S}^{\top}\hat{B}\hat{S}$$
 are sparse.

For Cartesian cells, the basis $\hat{S} \otimes \hat{S} \otimes \hat{S}$ sparsifies A. The blocks for the interior DOFs become diagonal.



We split the DOFs $\{I, \Gamma\}$. On the interior *I*, solve the 1D eigenproblems:

$$\hat{A}_{II}\hat{S}_{II}=\hat{B}_{II}\hat{S}_{II}\Lambda_{II}.$$

We construct the FDM basis functions

$$\hat{S} = egin{bmatrix} \hat{S}_{II} & -B_{II}^{-1}B_{I\Gamma} \ 0 & \mathbb{1} \end{bmatrix}$$

such that the 1D matrices

$$\hat{S}^{\top}\hat{A}\hat{S}, \quad \hat{S}^{\top}\hat{B}\hat{S}$$
 are sparse.

For Cartesian cells, the basis $\hat{S} \otimes \hat{S} \otimes \hat{S}$ sparsifies A. The blocks for the interior DOFs become diagonal.



We split the DOFs $\{I, \Gamma\}$. On the interior *I*, solve the 1D eigenproblems:

$$\hat{A}_{II}\hat{S}_{II}=\hat{B}_{II}\hat{S}_{II}\Lambda_{II}.$$

We construct the FDM basis functions

$$\hat{S} = \begin{bmatrix} \hat{S}_{II} & -B_{II}^{-1}B_{I\Gamma} \\ 0 & \mathbb{1} \end{bmatrix}$$

such that the 1D matrices

$$\hat{S}^{\top}\hat{A}\hat{S}, \quad \hat{S}^{\top}\hat{B}\hat{S}$$
 are sparse.

For Cartesian cells, the basis $\hat{S} \otimes \hat{S} \otimes \hat{S}$ sparsifies A. The blocks for the interior DOFs become diagonal.













Interior DOFs (I) are coupled with their projections onto the interface.





Interior DOFs (1) are coupled with their projections onto the interface.

Interface DOFs (Γ) are coupled to a line of interior DOFs and to the only vertex DOF.





The Cholesky factorization is also sparse. Fill-in only on the interface block.





The solve phase has $O(p^{d+1})$ cost. Factorization needs $O(p^{3(d-1)})$ operations.





Comparing the FDM-patch with a low-order *h*-FEM preconditioner



The Cholesky factor requires $O(p^{2(d-1)})$ storage, but $4 \times$ less than *h*-FEM.





The new basis does not sparsify A, but we may build a preconditioner P by replacing the geometry such that each individual cell is approximated by a Cartesian one.



Cartesian approximation of a vertex patch with 5 cells (p = 4).

We construct an auxiliary form that is separable and sparse. The auxiliary problem is spectrally equivalent independent of p.



$$\begin{aligned} \mathbf{a}(u,v) &= \sum_{K \in \mathcal{T}_h} \int_{\hat{K}} \hat{\nabla} u \circ F_K \cdot \hat{G}_K \hat{\nabla} v \circ F_K \, \mathrm{d}\hat{\mathbf{x}}, \\ b(u,v) &= \sum_{K \in \mathcal{T}_h} \int_{\hat{K}} \hat{\nabla} u \circ F_K \cdot \hat{\mu}_K \hat{\nabla} v \circ F_K \, \mathrm{d}\hat{\mathbf{x}}. \end{aligned}$$

Here, $F_{\mathcal{K}}: \hat{\mathcal{K}} \to \mathcal{K}$ is the coordinate mapping, $\hat{G}_{\mathcal{K}}$ is the Jacobian-weighted metric, and $\hat{\mu}_{\mathcal{K}}$ is a constant approximation to diag $(\hat{G}_{\mathcal{K}})$.

Spectrally equivalent separable form (B. & Farrell, 2021)

$$\min_{K} c_{K} \leq \frac{a(v,v)}{b(v,v)} \leq \max_{K} C_{K} \quad \forall v \in V,$$

where

$$\sigma(\hat{\mu}_{\mathcal{K}}^{-1/2}\hat{G}_{\mathcal{K}}\hat{\mu}_{\mathcal{K}}^{-1/2})\in [c_{\mathcal{K}}, C_{\mathcal{K}}].$$

We construct an auxiliary form that is separable and sparse. The auxiliary problem is spectrally equivalent independent of p.



$$a(u,v) = \sum_{K \in \mathcal{T}_h} \int_{\hat{K}} \hat{\nabla} u \circ F_K \cdot \hat{G}_K \hat{\nabla} v \circ F_K \, \mathrm{d}\hat{\mathbf{x}},$$
$$b(u,v) = \sum_{K \in \mathcal{T}_h} \int_{\hat{K}} \hat{\nabla} v \circ F_K \, \mathrm{d}\hat{\mathbf{x}},$$

$$b(u,v) = \sum_{K \in \mathcal{T}_h} \int_{\hat{K}} \hat{\nabla} u \circ F_K \cdot \hat{\mu}_K \hat{\nabla} v \circ F_K \, \mathrm{d}\hat{\mathbf{x}}.$$

Here, $F_K : \hat{K} \to K$ is the coordinate mapping, \hat{G}_K is the Jacobian-weighted metric, and $\hat{\mu}_K$ is a constant approximation to diag (\hat{G}_K) .

Spectrally equivalent separable form (B. & Farrell, 2021)

$$\min_{K} c_{K} \leq \frac{a(v,v)}{b(v,v)} \leq \max_{K} C_{K} \quad \forall v \in V,$$

where

$$\sigma(\hat{\boldsymbol{\mu}}_{\boldsymbol{K}}^{-1/2}\hat{\boldsymbol{G}}_{\boldsymbol{K}}\hat{\boldsymbol{\mu}}_{\boldsymbol{K}}^{-1/2})\in[\boldsymbol{c}_{\boldsymbol{K}},\boldsymbol{C}_{\boldsymbol{K}}].$$



	n	Cartesian $\kappa(P^{-1}A)$ Iter		non-Cartesian $\kappa(P^{-1}A)$ Iter		
	Ρ		1101.			
2D	3	1.58	7	2.13	10	
	7	1.59	7	3.07	13	
	15	1.59	6	3.81	14	
	31	1.58	6	3.53	14	
3D	3	2.98	13	3.79	16	
	7	2.91	12	5.04	19	
	15	2.85	11	5.68	19	





$$-\nabla \cdot (a(\mathbf{x})\nabla u) = f \quad \rightarrow \quad -\nabla^2 \tilde{u} + q(\mathbf{x})\tilde{u} = a^{-1/2}f$$

where $\tilde{u} = a^{1/2}u$ and $q(\mathbf{x}) = a^{-1/2}\nabla^2(a^{1/2})$.

a(x) varying across 7 orders of magnitude, 32×32 mesh, rel. tol. = 10^{-16}





$$\begin{aligned} -\nabla \cdot \mu (\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}) + \nabla p &= f, \\ \nabla \cdot \mathbf{u} - \lambda^{-1} p &= 0. \end{aligned} \qquad \begin{bmatrix} A & B^{\top} \\ B & -C \end{bmatrix} \begin{bmatrix} \underline{\mathbf{u}} \\ p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}} \\ 0 \end{bmatrix} , \end{aligned}$$

$$\beta_0^2 \leq \frac{(\underline{q}, BA^{-1}B^{\top}\underline{q})}{(\underline{q}, M_p\underline{q})} \leq \beta_1^2 \quad \forall \underline{q} \in \mathbb{R}^{n_p} \setminus \{0\}.$$

- For the standard [H¹]^d × L²-conforming space [Q_p]^d × DQ_{p-2}, β₀ ≤ Cp^{(1-d)/2}.
- For the H(div) × L²-conforming space RT_p × DQ_{p-1}, β₀ is independent of p.



$$\begin{aligned} -\nabla \cdot \mu (\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}) + \nabla p &= f, \\ \nabla \cdot \mathbf{u} - \lambda^{-1} p &= 0. \end{aligned} \qquad \begin{bmatrix} A & B^{\top} \\ B & -C \end{bmatrix} \begin{bmatrix} \underline{\mathbf{u}} \\ p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}} \\ 0 \end{bmatrix}$$

$$\beta_0^2 \leq \frac{(\underline{q}, BA^{-1}B^{\top}\underline{q})}{(\underline{q}, M_p\underline{q})} \leq \beta_1^2 \quad \forall \underline{q} \in \mathbb{R}^{n_p} \setminus \{\mathbf{0}\}.$$

- For the standard [H¹]^d × L²-conforming space [Q_p]^d × DQ_{p-2}, β₀ ≤ Cp^{(1-d)/2}.
- For the H(div) × L²-conforming space RT_p × DQ_{p-1}, β₀ is independent of p.



$$\begin{aligned} -\nabla \cdot \mu (\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}) + \nabla p &= f, \\ \nabla \cdot \mathbf{u} - \lambda^{-1} p &= 0. \end{aligned} \qquad \begin{bmatrix} A & B^{\top} \\ B & -C \end{bmatrix} \begin{bmatrix} \underline{\mathbf{u}} \\ p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}} \\ 0 \end{bmatrix}$$

$$\beta_0^2 \leq \frac{(\underline{q}, BA^{-1}B^{\top}\underline{q})}{(\underline{q}, M_p\underline{q})} \leq \beta_1^2 \quad \forall \underline{q} \in \mathbb{R}^{n_p} \setminus \{0\}.$$

- For the standard $[H^1]^d \times L^2$ -conforming space $[Q_p]^d \times DQ_{p-2}$, $\beta_0 \leq Cp^{(1-d)/2}$.
- For the H(div) × L²-conforming space RT_p × DQ_{p-1}, β₀ is independent of p.



$$\begin{aligned} -\nabla \cdot \mu (\nabla \mathbf{u} + \nabla \mathbf{u}^{\top}) + \nabla p &= f, \\ \nabla \cdot \mathbf{u} - \lambda^{-1} p &= 0. \end{aligned} \qquad \begin{bmatrix} A & B^{\top} \\ B & -C \end{bmatrix} \begin{bmatrix} \underline{\mathbf{u}} \\ p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{f}} \\ 0 \end{bmatrix}$$

$$\beta_0^2 \leq \frac{(\underline{q}, BA^{-1}B^{\top}\underline{q})}{(\underline{q}, M_p\underline{q})} \leq \beta_1^2 \quad \forall \underline{q} \in \mathbb{R}^{n_p} \setminus \{\mathbf{0}\}.$$

- For the standard $[H^1]^d \times L^2$ -conforming space $[Q_p]^d \times DQ_{p-2}$, $\beta_0 \leq Cp^{(1-d)/2}$.
- For the $H(\text{div}) \times L^2$ -conforming space $\text{RT}_p \times \text{DQ}_{p-1}$, β_0 is independent of p.

Linear elasticity: block-preconditioned MINRES iterations $\Omega = [0, 1]^d$, Dirchlet BCs, constant forcing, 8^d cells.



Mixed FEM $[Q_p]^d \times DQ_{p-2}$.

		Lamé parameter, λ				
	р	10 ⁰	101	10 ²	10 ³	∞
2D	3	20	27	30	30	30
	7	21	32	37	37	37
	15	22	36	42	42	42
	31	22	37	45	47	47
3D	3	30	48	57	58	58
	7	32	57	73	76	76
	15	33	59	85	89	89

Problem

This discretization is not *p*-robust.



- Functions in *H*(div) have continuity only on the normal component along facets. IP-DG imposes weak continuity of the tangential components.
- The FDM basis can also sparsify the additional surface integral terms.
- DG patch problems have more DOFs and are less sparse than CG.
- Tricky implementation: orientation, anisotropic polynomial degrees.





- Functions in *H*(div) have continuity only on the normal component along facets. IP-DG imposes weak continuity of the tangential components.
- The FDM basis can also sparsify the additional surface integral terms.
- DG patch problems have more DOFs and are less sparse than CG.
- Tricky implementation: orientation, anisotropic polynomial degrees.





- Functions in *H*(div) have continuity only on the normal component along facets. IP-DG imposes weak continuity of the tangential components.
- The FDM basis can also sparsify the additional surface integral terms.
- DG patch problems have more DOFs and are less sparse than CG.
- Tricky implementation: orientation, anisotropic polynomial degrees.





- Functions in *H*(div) have continuity only on the normal component along facets. IP-DG imposes weak continuity of the tangential components.
- The FDM basis can also sparsify the additional surface integral terms.
- DG patch problems have more DOFs and are less sparse than CG.
- Tricky implementation: orientation, anisotropic polynomial degrees.





Mixed FEM $RT_{p} \times DQ_{p-1}$, symmetric interior-penalty DG.

		Lamé parameter, λ				
	р	10 ⁰	10 ¹	10 ²	10 ³	∞
2D	3	19	28	31	32	32
	7	20	31	35	35	35
	15	21	33	38	38	38
	31	23	35	40	41	41
3D	3	24	40	49	51	51
	7	26	44	55	56	56
	15	29	48	59	60	60

Solution

This discretization is (more) *p*-robust.



- Constructed a sparse preconditioner by partially diagonalizing the 1D matrices.
- Effective and fast relaxation method.
- We can solve problems with very high *p*.

Ongoing work:

• Extension to H(div) problems on unstructured meshes.