Local Fourier Analysis of Domain Decomposition and Multigrid Methods for High-Order Matrix-Free Finite Elements

Jeremy L Thompson

University of Colorado Boulder

jeremy@jeremylt.org

July 8, 2021

Funding

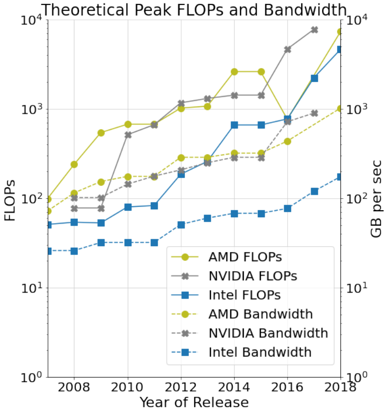
This work is supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of two U.S. Department of Energy organizations (Office of Science and the National Nuclear Security Administration) responsible for the planning and preparation of a capable exascale ecosystem, including software, applications, hardware, advanced system engineering and early testbed platforms, in support of the nation's exascale computing imperative.

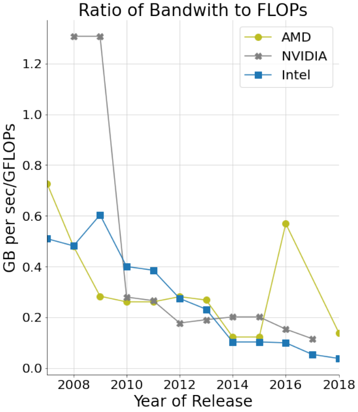
Overview

Big Picture

- High-order matrix-free representations of PDEs are better suited to modern hardware than sparse matrices
- High-order matrix-free representations require preconditioned iterative solvers
- Local Fourier Analysis (LFA) provides sharp convergence estimates for these preconditioners
- We develop LFA of p-multigrid and Balancing Domain Decomposition by Constraints (BDDC) on high-order element subdomains
- Further, we investigate LFA of p-multigrid with a BDDC smoother

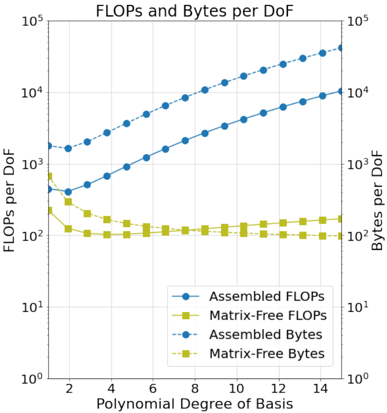
Modern Hardware

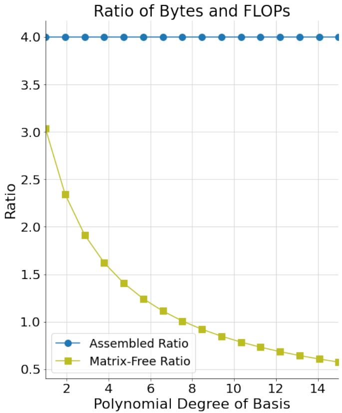




Modern hardware has lower memory bandwidth than FLOPs [4]

Benefits of Matrix-Free





Requirements for matrix-vector product with sparse matrix vs matrix-free for screened Poisson $\nabla^2 u - \alpha^2 u = f$ in 3D

Matrix-free representations using tensor product bases better match modern hardware limitations

Matrix-Free Representation

Weak form for an arbitrary second order PDE [2]:

$$
\begin{gathered}
\text{find } u \in V \text{ such that for all } v \in V \\
\langle v, u \rangle = \int_\Omega v \cdot f_0 \left(u, \nabla u \right) + \nabla v : f_1 \left(u, \nabla u \right) = 0
\end{gathered}
\tag{1}
$$

- $\cdot$ - contraction over fields
- $:$ - contraction over fields and spatial dimensions

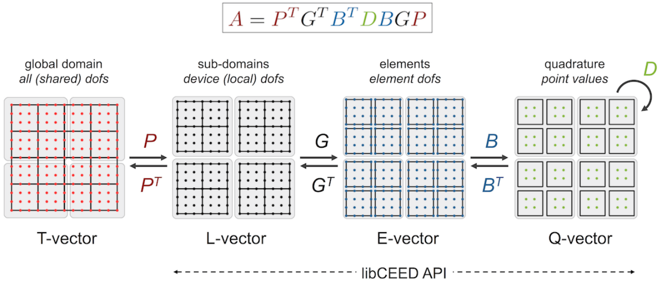
Matrix-Free Representation

Galerkin form for an arbitrary second order PDE:

$$\sum_e \mathcal{E}^T \left[\left(\mathrm{N}^e\right)^T \mathrm{W}^e \Lambda \left(f_0 \left(u^e, \nabla u^e \right) \right) + \sum_{i=0}^{d-1} \left(\mathrm{D}_i^e\right)^T \mathrm{W}^e \Lambda \left(f_1 \left(u^e, \nabla u^e \right) \right) \right] = 0 \tag{2}$$

- $\mathcal{E}$ - element assembly/restriction operator
- N^e - interpolation to quadrature points
- D_i^e - derivatives at quadrature points
- W^e - quadrature weights
- Λ - pointwise multiplication at quadrature points

libCEED Representation



[1]

- P - parallel element assembly operator
- G - local element assembly operator
- B - basis action operator
- D - weak form and geometry at quadrature points

LFA Background

Consider a scalar Toeplitz operator L_h on the infinite 1D grid G_h

$$\begin{aligned} L_h &\mathrel{\hat{=}} \left[s_\kappa \right]_h (\kappa \in V) \\ L_h w_h (x) &= \sum_{\kappa \in V} s_\kappa w_h (x + \kappa h) \end{aligned} \tag{3}$$

where

- $V \subset \mathcal{Z}$ is an index set
- $s_k \in \mathcal{R}$ are constant coefficients
- $w_h (x)$ is a l^2 function on G_h

LFA Background

If for all grid functions $\varphi(\theta, x)$

$$L_h \varphi(\theta, x) = \tilde{L}_h(\theta) \varphi(\theta, x) \tag{4}$$

then $\tilde{L}_h(\theta) = \sum_{\kappa \in V} s_\kappa e^{\imath \theta \kappa}$ is the **symbol** of L_h

Our function can be diagonalized by the standard Fourier modes

LFA Background

For a $q \times q$ system of equations, the matrix symbol is given by

$$\mathrm{L}_h = \begin{bmatrix} L_h^{1,1} & \dots & L_h^{1,q} \\ \vdots & \vdots & \vdots \\ L_h^{q,1} & \dots & L_h^{q,q} \end{bmatrix} \quad \Rightarrow \quad \tilde{\mathrm{L}}_h = \begin{bmatrix} \tilde{L}_h^{1,1} & \dots & \tilde{L}_h^{1,q} \\ \vdots & \vdots & \vdots \\ \tilde{L}_h^{q,1} & \dots & \tilde{L}_h^{q,q} \end{bmatrix} \tag{5}$$

LFA of High-Order FEM

For a scalar PDE operator on a single 1D finite element

$$\tilde{\mathrm{A}}\left(\boldsymbol{\theta}\right) = \mathrm{Q}^T \left(\mathrm{A}^e \odot \left[e^{\imath\left(\mathrm{x}_j - \mathrm{x}_i\right)\cdot\boldsymbol{\theta}/\mathrm{h}}\right]\right)\mathrm{Q} \tag{6}$$

where

$$\mathrm{A}^e = \mathrm{B}^T \mathrm{D} \mathrm{B} \tag{7}$$

$$\mathrm{Q} = \begin{bmatrix} \mathrm{I} \\ \mathrm{e}_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \end{bmatrix} \tag{8}$$

LFA of High-Order FEM

Symbol naturally extends to multiple components and higher dimensions

Multiple Components:

$$Q_n = I_n \otimes Q \tag{9}$$

Multiple Dimensions:

$$Q_{nd} = Q \otimes Q \otimes \cdots \otimes Q \tag{10}$$

Example: Scalar Poisson

$$\int \nabla v \nabla u = \int fv \tag{11}$$

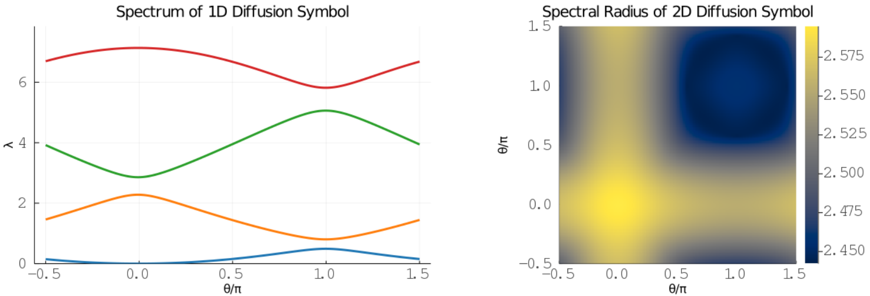
- B - given by tensor H1 Lagrange basis
- D - given by quadrature weights and product

```
# mesh
dim = 1
mesh = Mesh1D(1.0)

# basis
p = 3
ncomp = 1
basis = TensorH1LagrangeBasis(p+1, p+1, ncomp, dim)

# weak form
function diffusionweakform(du::Array{Float64}, w::Array{Float64})
    return dv = du*w[1]
end
```

Example: Scalar Poisson



Scalar Poisson problem on quartic elements

$$\text{low frequencies - } \theta \in T^{\text{low}} = [-\pi/2, \pi/2)^d$$
$$\text{high frequencies - } \theta \in T^{\text{high}} = [-\pi/2, 3\pi/2)^d \setminus T^{\text{low}}$$

LFA of High-Order Smoothers

Error propagation operator for smoothers given by

$$\mathsf{S} = \mathsf{I} - \mathsf{M}^{-1} A \tag{12}$$

with a symbol given by

$$\tilde{\mathsf{S}}\left(\omega, \boldsymbol{\theta}\right) = \mathsf{I} - \tilde{\mathsf{M}}^{-1}\left(\omega, \boldsymbol{\theta}\right) \tilde{A}\left(\boldsymbol{\theta}\right) \tag{13}$$

Jacobi Smoothing

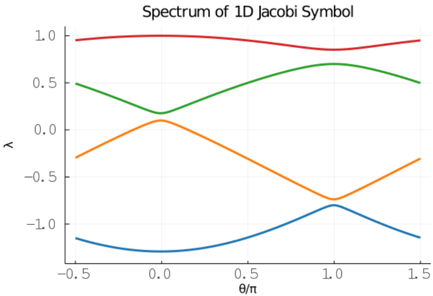
Jacobi smoothing given by

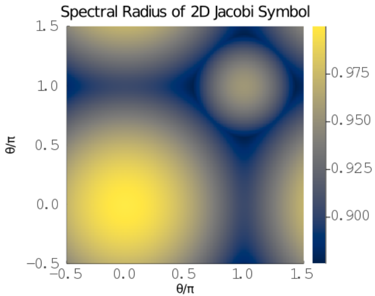
$$\mathsf{M}^{-1} = \omega \ \mathrm{diag}\left(A\right)^{-1} \tag{14}$$

with an error symbol given by

$$\tilde{\mathsf{S}}\left(\omega, \boldsymbol{\theta}\right) = \mathsf{I} - \omega \ \left(\mathsf{Q}^T \, \mathrm{diag}\left(A^e\right) \mathsf{Q}\right)^{-1} \tilde{A}\left(\boldsymbol{\theta}\right) \tag{15}$$

Example: Jacobi Smoothing





Jacobi smoothing with $\omega = 1.0$ on quartic elements

$$\text{low frequencies - } \theta \in T^{\text{low}} = [-\pi/2, \pi/2)^d$$
$$\text{high frequencies - } \theta \in T^{\text{high}} = [-\pi/2, 3\pi/2)^d \setminus T^{\text{low}}$$

Chebyshev Smoother

Error in *kth* order Chebyshev smoothing is given by

$$\begin{aligned} \mathsf{E}_0 &= \mathsf{I} \\ \mathsf{E}_1 &= \mathsf{I} - \frac{1}{\alpha} \left(\operatorname{diag} \mathsf{A}\right)^{-1} \mathsf{A} \\ \mathsf{E}_k &= \left(\left(\operatorname{diag} \mathsf{A}\right)^{-1} \mathsf{A} \mathsf{E}_{k-1} - \alpha \mathsf{E}_{k-1} - \beta_{k-2} \mathsf{E}_{k-2}\right) / \gamma_{k-1} \end{aligned} \tag{16}$$

for an operator with a spectrum on the interval $[\alpha - c, \alpha + c]$ where

$$\begin{aligned} \beta_0 &= -\frac{c^2}{2\alpha} & \gamma_0 &= -\alpha \\ \beta_k &= \frac{c}{2} \frac{T_k(\eta)}{T_{k+1}(\eta)} = \left(\frac{c}{2}\right)^2 \frac{1}{\gamma_k} & \gamma_k &= \frac{c}{2} \frac{T_{k+1}(\eta)}{T_k(\eta)} = -\left(\alpha + \beta_{k-1}\right). \end{aligned} \tag{17}$$

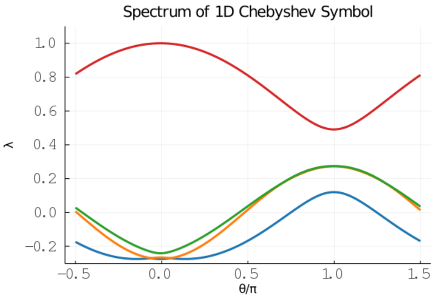
Chebyshev Smoother

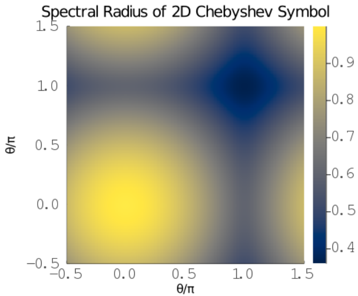
The error symbol of kth order Chebyshev smoother is given by

$$\begin{gathered}\tilde{\mathsf{E}}_0\left(\boldsymbol{\theta}\right) = \mathsf{I} \\ \tilde{\mathsf{E}}_1\left(\boldsymbol{\theta}\right) = \mathsf{I} - \frac{1}{\alpha}\tilde{\mathsf{A}}_J\tilde{\mathsf{A}}\left(\boldsymbol{\theta}\right) \\ \tilde{\mathsf{E}}_k\left(\boldsymbol{\theta}\right) = \left(\tilde{\mathsf{A}}_J\tilde{\mathsf{A}}\left(\boldsymbol{\theta}\right)\tilde{\mathsf{E}}_{k-1}\left(\boldsymbol{\theta}\right) - \alpha\tilde{\mathsf{E}}_{k-1}\left(\boldsymbol{\theta}\right) - \beta_{k-2}\tilde{\mathsf{E}}_{k-2}\left(\boldsymbol{\theta}\right)\right)/\gamma_{k-1}\end{gathered} \tag{18}$$

with $\tilde{A}_J$ being the symbol of the Jacobi preconditioner

Example: Chebyshev Smoothing





Third order Chebyshev smoothing quartic elements

$$\text{low frequencies - } \theta \in T^{\text{low}} = [-\pi/2, \pi/2)^d$$
$$\text{high frequencies - } \theta \in T^{\text{high}} = [-\pi/2, 3\pi/2)^d \setminus T^{\text{low}}$$

Two-Grid Multigrid Error

Multigrid methods target the low frequency error

$$E_{2MG} = S_f \left(I - P_{ctof} A_c^{-1} R_{ftoc} A_f \right) S_f \tag{19}$$

- A_f - fine grid PDE operator
- A_c^{-1} - coarse grid solve (low frequency error)
- S_f - fine grid smoother (high frequency error)
- P_{ctof} - coarse to fine grid prolongation operator
- R_{ftoc} - fine to coarse grid restriction operator

Grid transfer operators and coarse representation differentiate h-multigrid and p-multigrid

Two-Grid Multigrid Error

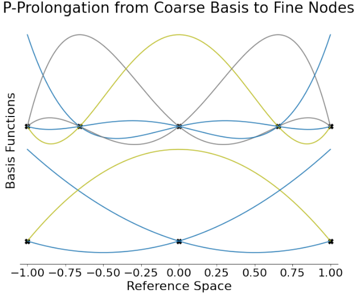
The definition of the symbol follows naturally

$$\tilde{\mathrm{E}}_{\mathrm{2MG}}(\boldsymbol{\theta}) = \tilde{\mathrm{S}}_f(\boldsymbol{\theta}, \omega) \left(\mathrm{I} - \tilde{\mathrm{P}}_{\mathrm{ctof}}(\boldsymbol{\theta}) \left(\tilde{\mathrm{A}}_c(\boldsymbol{\theta}) \right)^{-1} \tilde{\mathrm{R}}_{\mathrm{ftoc}}(\boldsymbol{\theta}) \tilde{\mathrm{A}}_f(\boldsymbol{\theta}) \right) \tilde{\mathrm{S}}_f(\boldsymbol{\theta}, \omega) \tag{20}$$

- A_f - fine grid PDE operator
- A_c^{-1} - coarse grid solve (low frequency error)
- S_f - fine grid smoother (high frequency error)
- $\mathrm{P}_{\mathrm{ctof}}$ - coarse to fine grid prolongation operator
- $\mathrm{R}_{\mathrm{ftoc}}$ - fine to coarse grid restriction operator

P-Multigrid Transfer Operators

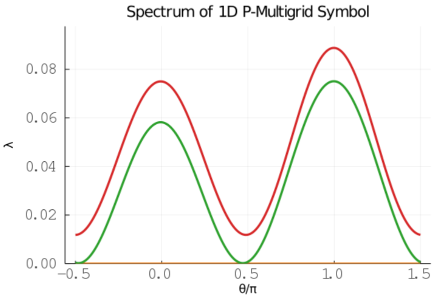
P-multigrid prolongation can be represented as an interpolation from the coarse to fine grid

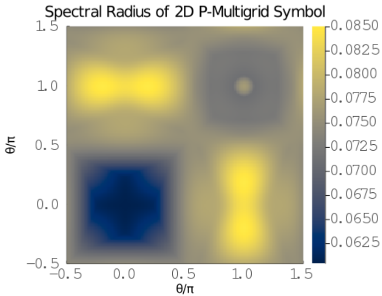


$$\begin{aligned} P_{ctof} &= P_f^T G_f^T P_e G_c P_c \\ P_e &= I D_{scale} B_{ctof} \end{aligned} \tag{21}$$

D scales for node multiplicity

Example: P-Multigrid





p-multigrid with third order Chebyshev on quadratic to linear elements

$$\text{low frequencies - } \theta \in T^{\text{low}} = [-\pi/2, \pi/2)^d$$
$$\text{high frequencies - } \theta \in T^{\text{high}} = [-\pi/2, 3\pi/2)^d \setminus T^{\text{low}}$$

Validation: P-Multigrid

p_{fine} to p_{coarse}	LFA	libCEED
$p = 2$ to $p = 1$	0.312	0.301
$p = 4$ to $p = 2$	1.436	1.402
$p = 4$ to $p = 1$	1.436	1.401
$p = 8$ to $p = 4$	1.989	1.885
$p = 8$ to $p = 2$	1.989	1.874
$p = 8$ to $p = 1$	1.989	1.875

LFA and experimental two-grid convergence factors with Jacobi smoothing for 3D Laplacian with $\omega = 1.0$

3D manufactured solution on the domain $[-3, 3]^3$ with Dirichlet boundaries:

$$f(x, y, z) = xyz \sin(\pi x) \sin(\pi (1.23 + 0.5y)) \sin(\pi (2.34 + 0.25z)) \tag{22}$$

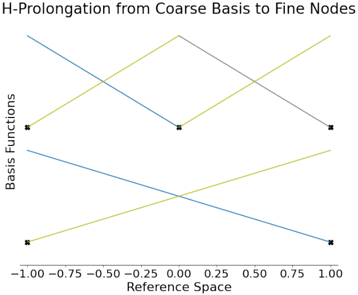
Validation: P-Multigrid

p_{fine} to p_{coarse}	$k = 2$		$k = 3$		$k = 4$	
	LFA	libCEED	LFA	libCEED	LFA	libCEED
$p = 2$ to $p = 1$	0.253	0.222	0.076	0.058	0.041	0.033
$p = 4$ to $p = 2$	0.277	0.251	0.111	0.097	0.062	0.050
$p = 4$ to $p = 1$	0.601	0.587	0.416	0.398	0.295	0.276
$p = 8$ to $p = 4$	0.398	0.391	0.197	0.195	0.121	0.110
$p = 8$ to $p = 2$	0.748	0.743	0.611	0.603	0.506	0.469
$p = 8$ to $p = 1$	0.920	0.914	0.871	0.861	0.827	0.814

LFA and experimental two-grid convergence factors with Chebyshev smoothing for 3D Laplacian

H-Multigrid Transfer Operators

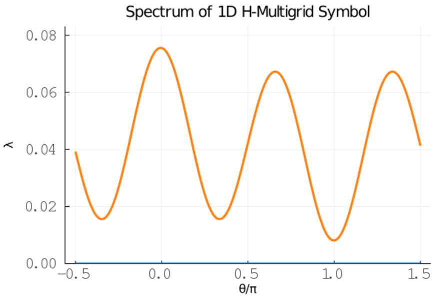
H-multigrid prolongation can be represented as an interpolation from the coarse grid to fine grid macro-elements

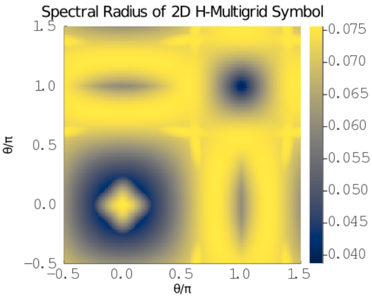


$$\begin{aligned} P_{\text{ctof}} &= P_f^T G_f^T P_e G_c P_c \\ P_e &= I D_{\text{scale}} B_{\text{ctof}} \end{aligned} \tag{23}$$

D scales for node multiplicity

Example: H-Multigrid





h-multigrid with third order Chebyshev on linear elements

low frequencies - $\theta \in T^{\text{low}} = [-\pi/2, \pi/2)^d$
high frequencies - $\theta \in T^{\text{high}} = [-\pi/2, 3\pi/2)^d \setminus T^{\text{low}}$

Validation: P-Multigrid

p, d	$\nu = (0, 1)$		$\nu = (1, 1)$		$\nu = (2, 2)$	
	ρ	ω	ρ	ω	ρ	ω
$p = 2, d = 1$	0.821	1.000	0.821	1.000	1.279	1.000
$p = 2, d = 1$	0.526	0.838	0.495	0.838	0.302	0.838
$p = 2, d = 1$	0.291	0.709	0.249	0.709	0.064	0.709
$p = 3, d = 1$	0.491	0.650	0.337	0.650	0.131	0.650
$p = 4, d = 1$	0.608	0.640	0.559	0.640	0.331	0.640
$p = 2, d = 2$	0.452	1.000	0.288	1.000	0.091	1.000

Two-grid convergence factor and Jacobi smoothing parameter for high-order h-multigrid

Results agree with previous work [3]

Big Picture

- High-order matrix-free representations of PDEs are better suited to modern hardware than sparse matrices
- High-order matrix-free representations require preconditioned iterative solvers
- Local Fourier Analysis (LFA) provides sharp convergence estimates for these preconditioners
- We develop LFA of p-multigrid and Balancing Domain Decomposition by Constraints (BDDC) on high-order element subdomains
- Further, we investigate LFA of p-multigrid with a BDDC smoother

Big Picture

- High-order matrix-free representations of PDEs are better suited to modern hardware than sparse matrices
- High-order matrix-free representations require preconditioned iterative solvers
- Local Fourier Analysis (LFA) provides sharp convergence estimates for these preconditioners
- We develop LFA of p-multigrid and Balancing Domain Decomposition by Constraints (BDDC) on high-order element subdomains
- Further, we investigate LFA of p-multigrid with a BDDC smoother

Local Fourier Analysis of Domain Decomposition and Multigrid Methods for High-Order Matrix-Free Finite Elements

Jeremy L Thompson

University of Colorado Boulder

jeremy@jeremylt.org

July 8, 2021

Ahmad Abdelfattah, Valeria Barra, Natalie Beams, Jed Brown, Jean-Sylvain Camier, Veselin Dobrev, Yohann Dudouit, Leila Ghaffari, Tzanio Kolev, David Medina, Thilina Rathnayake, Jeremy L Thompson, and Stanimire Tomov.
libCEED User Manual, September 2020.

Jed Brown.
Efficient nonlinear solvers for nodal high-order finite elements in 3D.
Journal of Scientific Computing, 45(1-3):48–63, 2010.

Yunhui He and Scott MacLachlan.
Two-level fourier analysis of multigrid for higher-order finite-element discretizations of the laplacian.
Numerical Linear Algebra with Applications, 27(3):e2285, 2020.

Karl Rupp.
CPU-GPU-MIC comparision charts, 2020.

Local Fourier Analysis of Domain Decomposition and Multigrid Methods for High-Order Matrix-Free Finite Elements

Jeremy L Thompson

University of Colorado Boulder

jeremy@jeremylt.org

July 8, 2021