Sparse Operator Compression of Elliptic Operators with Multiscale Coefficients

Thomas Y. Hou

Applied and Comput. Math, Caltech

February 1, 2017

Joint work with Pengchuan Zhang

Big Data Meets Computation, IPAM

Motivations

Dimension reduction appears nearly everywhere in science and engineering.

- Solving elliptic equations with multiscale coefficients: multiscale finite element basis for the elliptic operator.
- Principal component analysis (PCA): principle modes of the covariance operator.
- Quantum chemistry: eigen states of the Hamiltonian.

For computational efficiency and/or good interpretability, localized basis functions are preferred.

- Localized multiscale finite element basis: Babuska-Caloz-Osbron-94, Hou-Wu-1997, Hughes-Feijóo-Mazzei-98, E-Engquist-03, Owhadi-Zhang-07, Målqvist-Peterseim-14, Owhadi-15, etc.
- Sparse principle modes obtained by Sparse PCA or sparse dictionary learning: Zou-Hastie-Tibshirani-04, Witten-Tibshirani-Hastie-09, etc.
- Compressed Wannier modes: Ozoliņš-Lai-Caflisch-Osher-13, E-Li-Lu-10, Lai-Lu-Osher-15, etc.

Operator compression

Consider an elliptic operator in the divergence form

$$\mathcal{L}u = \sum_{0\le|\sigma|,|\gamma|\le k} (-1)^{|\sigma|} D^\sigma (a_{\sigma\gamma}(x) D^\gamma u), \tag{1}$$

where the coefficients $a_{\sigma\gamma} \in L^\infty(D)$, D is a bounded domain in $\mathbb{R}^d$, $\sigma = (\sigma_1, \ldots, \sigma_d)$ is a d-dimensional multi-index.

- $\mathcal{L}$ is self-adjoint and positive definite in a Hilbert space $H^k_{\mathcal{B}}(D)$. $H^k_{\mathcal{B}}(D) \subset H^k(D)$ incorporates the boundary condition for the elliptic operator.
- For any $f \in L^2(D)$, $\mathcal{L}u = f$ has a unique weak solution in $H^k_{\mathcal{B}}(D)$, denoted as $u := \mathcal{L}^{-1} f$.
- Given n basis functions $\Psi = [\psi_1, \ldots, \psi_n] \subset H^k_{\mathcal{B}}(D)$, we define the operator compression error:

$$E_{oc}(\Psi; \mathcal{L}^{-1}) := \min_{K_n \in \mathbb{R}^{n\times n},\ K_n \succeq 0} \|\mathcal{L}^{-1} - \Psi K_n \Psi^T\|_2, \tag{2}$$

  which is the optimal approximation error of $\mathcal{L}^{-1}$ among all positive semidefinite operators with range space spanned by Ψ.

Main results of sparse operator compression

Definition

Given n basis functions $\Psi = [\psi_1, \ldots, \psi_n] \subset H^k_{\mathcal{B}}(D)$, we define the operator compression error:

$$E_{oc}(\Psi; \mathcal{L}^{-1}) := \min_{K_n \in \mathbb{R}^{n\times n},\ K_n \succeq 0} \|\mathcal{L}^{-1} - \Psi K_n \Psi^T\|_2,$$

which is the optimal approximation error of $\mathcal{L}^{-1}$ among all positive semidefinite operators with range space spanned by Ψ.

For any $n \in \mathbb{N}$, we construct n localized basis functions $\{\psi_i^{loc}\}_{i=1}^n$ such that

- $$\left|\mathrm{supp}(\psi_i^{loc})\right| \leq \frac{C_l \log(n)}{n}, \quad \forall 1 \leq i \leq n. \tag{3}$$
- $$E_{oc}(\Psi^{loc}; \mathcal{L}^{-1}) \leq C_e \lambda_n(\mathcal{L}^{-1}), \tag{4}$$
- The constants C_l and C_e are independent of n and multiscale features in $a_{\sigma\gamma}$.

Potential Applications I. Solving elliptic equations.

$\mathcal{L}$ is an elliptic operator of order $2k$ ($k \geq 1$) with rough multiscale coefficients in $L^\infty(D)$, and the load $f \in L^2(D)$.

$$\mathcal{L}u = f, \qquad u \in H_0^k(D). \tag{5}$$

- $k = 1$: heat equation, subsurface flow; $k = 2$: beam equation, plate equation, etc...
- We construct nearly optimally localized basis functions $\{\psi_i^{loc}\}_{i=1}^n \subset H_0^k(D)$. For a given mesh h, we have

$$\left|\text{supp}(\psi_i^{loc})\right| \leq C_l h \log(1/h) \quad 1 \leq i \leq n.$$

- The multiscale finite element solution $u_{ms} := \Psi^{loc} L_n^{-1} \left(\Psi^{loc}\right)^T f$ satisfies

$$\|u - u_{ms}\|_H \leq C_e h^k \|f\|_2 \quad \forall f \in L^2(D),$$

  where $\|\cdot\|_H$ is the energy norm, C_e is indep. of small scale of $a_{\sigma\gamma}$.
- Sparsity/locality: computational efficiency.

Potential Applications II. Sparse PCA.

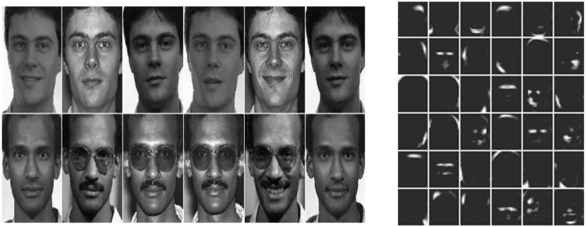


Figure: Left: samples of human faces. Right: sparse principal modes. [1]

The Matérn class covariance in spatial statistics

$$K_\nu(x,y) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right) \tag{6}$$

- $\nu = 1/2$: $K_{1/2}(x,y) = \sigma^2 \exp(-|x-y|/\rho)$
- $\nu \to \infty$: $\lim_{\nu\to\infty} K_\nu(x,y) = \sigma^2 \exp\left(-\frac{|x-y|^2}{2\rho^2}\right)$.

[1] Wang-Jia-Hu-Turk, IJPRAI, 2005

Potential Applications II. Sparse PCA continued.

The Matérn class covariance

$$K_\nu(x,y) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right)^\nu K_\nu \left(\sqrt{2\nu}\frac{|x-y|}{\rho}\right),$$

- It is the solution operator of high-order elliptic operators

$$\mathcal{L} = C_{\nu,\lambda}\sigma^2 \left(\frac{2\nu}{\lambda^2} - \Delta\right)^{\nu+d/2}.$$

- We construct nearly optimally localized basis functions $\{\psi_i^{loc}\}_{i=1}^n$:

$$\left|\mathsf{supp}(\psi_i^{loc})\right| \leq \frac{C_l \log(n)}{n} \quad 1 \leq i \leq n.$$

- We can approximate K_ν by rank-n operator with optimal accuracy:

$$\left\|K_\nu - \Psi^{loc} K_n \left(\Psi^{loc}\right)^T\right\|_2 \leq C_e \lambda_n(K_\nu).$$

- Sparsity/locality: better interpretability and computational efficiency.

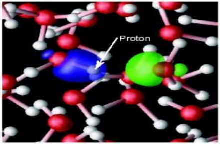
Potential Applications III. Quantum chemistry.

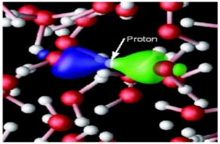
Maximally-localized generalized Wannier functions for composite energy bands

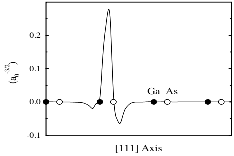
Nicola Marzari and David Vanderbilt
Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08855-0849, USA
(July 10, 1997)

We discuss a method for determining the optimally-localized set of generalized Wannier functions associated with a set of Bloch bands in a crystalline solid. By "generalized Wannier functions" we mean a set of localized orthonormal orbitals spanning the same space as the specified set of Bloch bands. Although we minimize a functional that represents the total spread $\sum_n \langle r^2\rangle_n - \langle \mathbf{r}\rangle_n^2$ of the Wannier functions in real space, our method proceeds directly from the Bloch functions as represented on a mesh of k-points, and carries out the minimization in a space of unitary matrices $U_{mn}^{(\mathbf{k})}$ describing the rotation among the Bloch bands at each k-point. The method is thus suitable for use in connection with conventional electronic-structure codes. The procedure also returns the total electric polarization as well as the location of each Wannier center. Sample results for Si, GaAs, molecular C_2H_4 and LiCl will be presented. [2]

$$\mathrm{i}\hbar\partial_t u(t,x) = \underbrace{\left(-\frac{\hbar^2}{2}\Delta_x + V(x)\right)}_{\text{Hamiltonian : } \mathcal{L}} u \quad \Rightarrow \quad \begin{cases} \mathcal{L}e_m = \lambda_m e_m \\ u(x,t) = \sum \alpha_m(t) e_m(x) \end{cases}$$







Sparsity/locality: better interpretability and computational efficiency.

[2] Marzari-Vanderbilt, PRB (56), 97

Potential Applications III. Quantum chemistry continued.

- We construct nearly optimally localized basis functions $\{\psi_i^{loc}\}_{i=1}^n$ that optimally approximates the eigenspace in the sense of

$$E_{oc}(\Psi^{loc};\mathcal{L}^{-1}) := \min_{\substack{K_n\in\mathbb{R}^{n\times n}\\ K_n\succeq 0}} \|\mathcal{L}^{-1} - \Psi^{loc}K_n\left(\Psi^{loc}\right)^T\|_2 \le C_e\lambda_n(\mathcal{L}^{-1}),$$

- Another natural choice to define the compression error:

$$\widetilde{E}_{oc}(\Psi) = \|\mathcal{P}_{V_n} - \mathcal{P}_\Psi\|_2 = \left\|\sum_{i=1}^n e_i e_i^T - \mathcal{P}_\Psi\right\|_2,$$

where V_n is the first n-dimensional eigenspace span$\{e_1,\ldots,e_n\}$ and $\mathcal{P}_V$ is the orthogonal projection from $L^2(D)$ to its subspace V.

-

$$E_{oc}(\Psi;\mathcal{L}^{-1}) = \min_{K_n\in\mathbb{R}^{n\times n},\ K_n\succeq 0} \left\|\sum_{i=1}^\infty \frac{1}{\lambda_i}e_i e_i^T - \Psi\mathcal{K}_n\Psi^T\right\|_2.$$

We believe that $E_{oc}(\Psi;\mathcal{L}^{-1})$ is a better criterion for operator compression because it takes into consideration the decay of the eigenvalues of the solution operator $\mathcal{L}^{-1}$.

Our construction and theoretical results

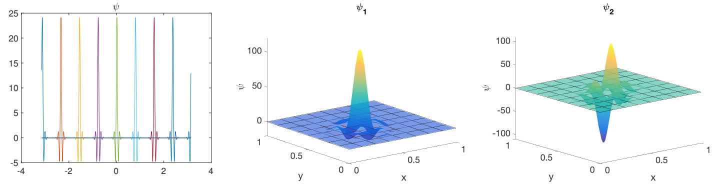


Figure: Left: 8 localized basis functions for $-\Delta$ with periodic BC. Middle and right: 2 localized basis functions for Δ^2 with homogeneous Dirichlet BC.

Our construction of $\{\psi_i^{loc}\}_{i=1}^n$

1. Choose $h > 0$. Partition the physical domain D using a regular partition $\{\tau_i\}_{i=1}^m$ with mesh size h.
2. Choose $r > 0$, say $r = 2h\log(1/h)$. For each patch τ_i, S_r is the union of the subdomains $\tau_{i'}$ intersecting $B(x_i, r)$ (for some $x_i \in \tau_i$).
3. $\mathcal{P}_{k-1}(\tau_i)$ is the space of all d-variate polynomials of degree at most $k-1$ on the patch τ_i. $Q = \binom{k+d-1}{d}$ is its dimension. $\{\varphi_{i,q}\}_{q=1}^Q$ is a set of orthogonal basis functions for $\mathcal{P}_{k-1}(\tau_i)$.

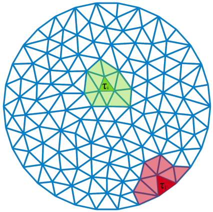


Figure: A regular partition, local patch τ_i and its associated S_r.

$$
\begin{aligned}
\psi_{i,q}^{loc} = \arg\min_{\psi \in H_{\mathcal{B}}^k} \quad & \|\psi\|_H^2 \\
\text{s.t.} \quad & \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \quad \forall 1 \le j \le m,\ 1 \le q' \le Q, \\
& \psi(x) \equiv 0, \quad x \in D \backslash S_r,
\end{aligned}
$$

where $H_{\mathcal{B}}^k$ is the solution space (with some prescribed BC), $\|\cdot\|_H$ is the energy norm associated with $\mathcal{L}$ and the BC.

Our construction $\Psi^{loc} := \{\psi_{i,q}^{loc}\}_{i=1,q=1}^{m,Q}$

Theorem (Hou-Zhang-2016)

Suppose $H_{\mathcal{B}}^k = H_0^k(D)$ *and* $\mathcal{L}u = (-1)^k \sum_{|\sigma|=|\gamma|=k} D^\sigma(a_{\sigma\gamma} D^\gamma u)$. *Assume that* $\mathcal{L}$ *is self-adjoint, positive definite and strongly elliptic, and that there exists* $\theta_{min}, \theta_{max} > 0$ *such that*

$$\theta_{min} \|\xi\|^{2k} \leq \sum_{|\sigma|=|\gamma|=k} a_{\sigma\gamma} \xi^\sigma \xi^\gamma \leq \theta_{max} \|\xi\|^{2k}, \quad \forall \xi \in \mathbb{R}^d.$$

Then for $r \geq C_r h \log(1/h)$, *we have*

1. $$\|\mathcal{L}^{-1} f - \Psi^{loc} L_n^{-1} (\Psi^{loc})^T f\|_H \leq \frac{C_e h^k}{\sqrt{\theta_{min}}} \|f\|_2 \quad \forall f \in L^2(D), \tag{7}$$

   where L_n *is the stiffness matrix under basis functions* Ψ^{loc}.

2. $$E_{oc}(\Psi^{loc}; \mathcal{L}^{-1}) \leq \frac{C_e^2 h^{2k}}{\theta_{min}}. \tag{8}$$

Here, the constant C_r *only depends on the contrast* $\frac{\theta_{max}}{\theta_{min}}$, *and* C_e *is independent of the coefficients.*

Several remarks

- Theorem (Hou-Zhang-2016) also applies to $\mathcal{L}$ with low order terms, i.e. $\mathcal{L}u = (-1)^k \sum_{|\sigma|,|\gamma| \le k} D^\sigma(a_{\sigma\gamma} D^\gamma u)$.
- Theorem (Hou-Zhang-2016) also applies to other homogeneous boundary conditions, like periodic BC, Robin BC and mixed BC.
- For $H^k_{\mathcal{B}} = H^1_0(D)$, i.e. second order elliptic operators with zero Dirichlet BC, Theorem (Hou-Zhang-2016) have been proved in Owhadi-2015. A similar result for $H^k_{\mathcal{B}} = H^1_0(D)$ was also provided in Målqvist-Peterseim-2014. In this case, Our proof improves the estimates of the constants C_r and C_e.
- For other BCs, operators with lower order terms, and high-order elliptic operators, new techniques and concepts have been developed. Among them, the most important three new techniques are
  - a projection-type polynomial approximation property in $H^k(D)$,
  - the notion of the strong ellipticity [3],
  - an inverse energy estimate for functions in $\Psi := \mathrm{span}\{\psi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$.

[3] Equivalent to uniform ellipticity when $d = 1, 2$ or $k = 1$. Slightly stronger than uniform ellipticity in other cases; counter examples exist but difficult to construct.

Roadmap of the proof: Error estimate

Theorem (An error estimate based on projection-type approximation)

Suppose there is a n-dimensional subspace $\Phi \subset L^2(D)$ with basis $\{\varphi_i\}_{i=1}^n$ such that

$$\|u - \mathcal{P}_{\Phi}^{(L^2)} u\|_{L^2} \leq k_n \|u\|_H \qquad \forall u \in H^k(D). \tag{9}$$

Let Ψ be the n-dimensional subspace in $H^k(D)$ (also in $H^k_{\mathcal{B}}(D)$) spanned by $\{\mathcal{L}^{-1}\varphi_i\}_{i=1}^n$. Then

1. *For any $f \in L^2(D)$ and $u = \mathcal{L}^{-1} f$, we have*

$$\|u - \mathcal{P}_{\Psi}^{(H^k_{\mathcal{B}})} u\|_H \leq k_n \|f\|_{L_2}\,. \tag{10}$$

2. *We have*

$$E_{oc}(\Psi; \mathcal{L}^{-1}) \leq k_n^2\,. \tag{11}$$

- $k = 1$: Φ piecewise constant functions. By the Poincare inequality, it is easy to obtain $\|u - \mathcal{P}_{\Phi}^{(L^2)} u\|_{L^2} \leq \frac{C_p h}{\sqrt{\theta_{min}}} \|u\|_H$.
- $k \geq 2$: Φ piecewise polynomials with degree no more than $k-1$. By a projection-type polynomial approximation property in $H^k(D)$, see Thm 3.1 in Hou-Zhang-PartII, we have $\|u - \mathcal{P}_{\Phi}^{(L^2)} u\|_{L^2} \leq \frac{C_p h^k}{\sqrt{\theta_{min}}} \|u\|_H$.

Roadmap of the proof: Error estimate, discussions

Take $H^k_{\mathcal{B}} = H^1_0(D)$ as an example, where Φ is the space of piecewise constant functions.

- Based on a projection-type approximation property, we obtain the error estimates of the GFEM in the energy norm, i.e.

  $$\|u - \mathcal{P}_{\Phi}^{(L^2)} u\|_{L^2} \leq C_{proj} h \|u\|_H \Rightarrow \|u - \mathcal{P}_{\Psi}^{(H_0^1)} u\|_H \leq C_{proj} h \|f\|_{L_2}.$$

  C_{proj} does not depends on the small scales in the coefficients.
- Tranditional interpolation-type estimation requires higher regularity of the solution u: assume $u \in H^2(D)$

  $$|u - \mathcal{I}_h u|_{1,2,D} \leq C h |u|_{2,2,D} \Rightarrow \|u - \mathcal{I}_h u\|_H \leq C_{interp} h \|f\|_{L_2}.$$

  C_{interp} depends on the small scales in the coefficients.
- Basis functions for $\mathcal{I}_h u$: optimally localized linear nodal basis
  Basis functions for $\mathcal{P}_{\Psi}^{(H_0^1)} u$: global basis functions $\{\mathcal{L}^{-1}\varphi_i\}_{i=1}^n$

Roadmap of the proof: Basis with exponential decay

Global energy minimizing basis functions

$$\boxed{\begin{aligned} \psi_{i,q} = & \operatorname*{arg\,min}_{\psi \in H^k_{\mathcal{B}}} \quad \|\psi\|_H^2 \\ & \text{s.t.} \quad \int_D \psi_{i,q} \varphi_{j,q'} = \delta_{iq,jq'}, \forall 1 \le q' \le Q, 1 \le j \le m. \end{aligned}} \tag{12}$$

Theorem (Energy minimizing basis functions with exponential decay)

- $\{\psi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$ *and* $\{\mathcal{L}^{-1}\varphi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$ *span the same space* Ψ*.*
- $\psi_{i,q}$ *decays exponentially fast away from its associated patch* τ_i*.*

Roadmap of the proof: Basis with exponential decay

Global energy minimizing basis functions

$$\boxed{\begin{aligned} \psi_{i,q} = \arg\min_{\psi \in H^k_{\mathcal{B}}} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_D \psi_{i,q}\varphi_{j,q'} = \delta_{iq,jq'}, \forall 1 \le q' \le Q, 1 \le j \le m. \end{aligned}} \tag{12}$$

Theorem (Energy minimizing basis functions with exponential decay)

- $\{\psi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$ *and* $\{\mathcal{L}^{-1}\varphi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$ *span the same space* Ψ*.*
- $\psi_{i,q}$ *decays exponentially fast away from its associated patch* τ_i*.*

Intuition: We apply a linear transform to $\{\mathcal{L}^{-1}\varphi_{i,q} : 1 \le i \le m, 1 \le q \le Q\}$ such that the new basis function $\psi_{i,q}$ has zero moments up to the $(k-1)$-th order on any patch other than τ_i.

Roadmap of the proof: Basis with exponential decay

$$\boxed{\begin{aligned}\psi_{i,q} = &\mathop{\arg\min}_{\psi \in H^k_{\mathcal{B}}} \quad \|\psi\|_H^2 \\ &\text{s.t.} \quad \int_D \psi_{i,q}\varphi_{j,q'} = \delta_{iq,jq'}, \forall 1 \le q' \le Q, 1 \le j \le m.\end{aligned}}$$

1D second order elliptic operator with Robin BC:

$$\mathcal{L}u = -\frac{1}{2}u'' + \frac{1}{2}u,$$
$$u(0) - u'(0) = 0,\ u(1) + u'(1) = 0,$$

1D Matérn covariance with $\nu = 1/2$:

$$K_{1/2}(x,y) = \exp(-|x-y|),$$
$$\mathcal{L}^{-1}f = \int_0^1 K_{1/2}(x,y)f(y)\mathrm{d}y.$$

Roadmap of the proof: Basis with exponential decay

$$\boxed{\begin{aligned} \psi_{i,q} &= \arg\min_{\psi \in H^k_{\mathcal{B}}} \quad \|\psi\|_H^2 \\ &\text{s.t.} \quad \int_D \psi_{i,q}\varphi_{j,q'} = \delta_{iq,jq'}, \forall 1 \le q' \le Q, 1 \le j \le m. \end{aligned}}$$

1D second order elliptic operator with Robin BC:

$$\mathcal{L}u = -\frac{1}{2}u'' + \frac{1}{2}u,$$
$$u(0) - u'(0) = 0, \ u(1) + u'(1) = 0,$$

1D Matérn covariance with $\nu = 1/2$:

$$K_{1/2}(x,y) = \exp(-|x-y|),$$
$$\mathcal{L}^{-1}f = \int_0^1 K_{1/2}(x,y)f(y)\mathrm{d}y.$$

$$H^k_{\mathcal{B}} = H^1([0,1]), \quad \|u\|_H^2 = \frac{1}{2}\left(u(0)^2 + u(1)^2 + \int_0^1 (u')^2 + \int_0^1 u^2\right).$$

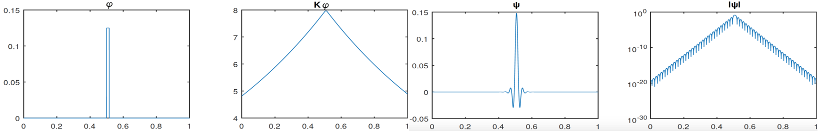


Figure: The basis function associated with patch $[1/2-h, 1/2]$, $h = 1/64$.

Roadmap of the proof: Localized basis

- Localized energy minimizing basis functions

$$\psi_{i,q}^{loc} = \underset{\psi \in H_{\mathcal{B}}^k}{\arg\min} \quad \|\psi\|_H^2$$

$$\text{s.t.} \quad \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \quad \forall 1 \le j \le m,\, 1 \le q' \le Q,$$

Roadmap of the proof: Localized basis

- Localized energy minimizing basis functions

$$\boxed{\begin{aligned} \psi_{i,q}^{loc} = \arg\min_{\psi \in H_{\mathcal{B}}^k} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \quad \forall 1 \le j \le m,\, 1 \le q' \le Q, \\ & \psi(x) \equiv 0, \quad x \in D \backslash S_r. \end{aligned}}$$

Roadmap of the proof: Localized basis

- Localized energy minimizing basis functions

$$\boxed{\begin{aligned} \psi_{i,q}^{loc} = \arg\min_{\psi \in H_{\mathcal{B}}^k} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \quad \forall 1 \le j \le m,\, 1 \le q' \le Q, \\ & \psi(x) \equiv 0, \quad x \in D \backslash S_r. \end{aligned}}$$

- Because $\psi_{i,q}$ decays exponentially fast away from patch τ_i, $r = \mathcal{O}(h \log(1/h))$ is sufficient to preserve the good error estimate of Ψ:

$$\|u - \mathcal{P}_{\Psi}^{(H_{\mathcal{B}}^k)} u\|_H \le \frac{C_p h^k}{\sqrt{\theta_{min}}} \|f\|_{L_2} \Rightarrow \|u - \mathcal{P}_{\Psi^{loc}}^{(H_{\mathcal{B}}^k)} u\|_H \le \frac{2 C_p h^k}{\sqrt{\theta_{min}}} \|f\|_{L_2}.$$

Roadmap of the proof: Localized basis

- Localized energy minimizing basis functions

$$\boxed{\begin{aligned} \psi_{i,q}^{loc} = \arg\min_{\psi \in H_{\mathcal{B}}^k} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \quad \forall 1 \le j \le m,\, 1 \le q' \le Q, \\ & \psi(x) \equiv 0, \quad x \in D \backslash S_r. \end{aligned}}$$

- Because $\psi_{i,q}$ decays exponentially fast away from patch τ_i, $r = \mathcal{O}(h \log(1/h))$ is sufficient to preserve the good error estimate of Ψ:

$$\|u - \mathcal{P}_{\Psi}^{(H_{\mathcal{B}}^k)} u\|_H \le \frac{C_p h^k}{\sqrt{\theta_{min}}} \|f\|_{L_2} \Rightarrow \|u - \mathcal{P}_{\Psi^{loc}}^{(H_{\mathcal{B}}^k)} u\|_H \le \frac{2C_p h^k}{\sqrt{\theta_{min}}} \|f\|_{L_2}.$$

- With the Aubin-Nistche duality argument, we have proved

$$E_{oc}(\Psi^{loc}; \mathcal{L}^{-1}) \le \frac{4C_p^2 h^{2k}}{\theta_{min}}.$$

Roadmap of the proof: Localized basis

$$\mathcal{L}u = -\frac{1}{2}u''(x) + \frac{1}{2}u, \quad u(0) - u'(0) = 0,\ u(1) + u'(1) = 0.$$

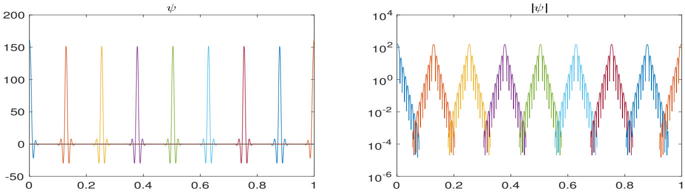


Figure: A few basis functions for the case $m = 2^7$ and $r = 2.4h\log_2(1/h)$.

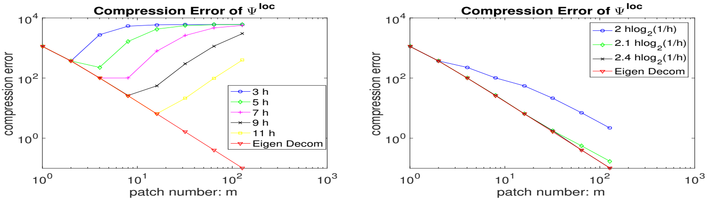


Figure: $E(\Psi^{loc}; \mathcal{L}^{-1})$ with localized basis functions Ψ^{loc}.

Compare with the l^1-minimization approach

Sparse Operator Compression

$$\boxed{\begin{aligned} \min_{\psi\in H^k_{\mathcal{B}}} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_{S_r} \psi\varphi_{j,q'} = \delta_{iq,jq'},\ \forall j, q' \\ & \psi(x) \equiv 0, \quad x \in D\backslash S_r, \end{aligned}}$$

Sparsity/locality from moment condition and exponential decay

l^1 minimization

$$\boxed{\begin{aligned} \min_{\Psi\subset H^k_{\mathcal{B}}} \quad & \sum_{i=1}^n \|\psi_i\|_H^2 + \mu \sum_{i=1}^n \|\psi_i\|_1, \\ \text{s.t.} \quad & \int_D \psi_i\psi_j = \delta_{i,j}\ \forall 1 \le i, j \le n. \end{aligned}}$$

Sparsity/locality from the l^1 penalty

See e.g. [Ozoliņš-Lai-Caflisch-Osher, PNAS, 2013].

Sparse OC vs l^1 minimization: math formulation

Sparse Operator Compression

$$\begin{aligned} \min_{\psi \in H^k_{\mathcal{B}}} \quad & \|\psi\|_H^2 \\ \text{s.t.} \quad & \int_{S_r} \psi \varphi_{j,q'} = \delta_{iq,jq'}, \ \forall j, q' \\ & \psi(x) \equiv 0, \quad x \in D \backslash S_r, \end{aligned}$$

l^1 minimization

$$\begin{aligned} \min_{\Psi \subset H^k_{\mathcal{B}}} \quad & \sum_{i=1}^n \|\psi_i\|_H^2 + \mu \sum_{i=1}^n \|\psi_i\|_1, \\ \text{s.t.} \quad & \int_D \psi_i \psi_j = \delta_{i,j} \ \forall 1 \le i, j \le n. \end{aligned}$$

- Linear constraints, convex quadratic optimization v.s. orthogonality constraints, non-convex optimization
- Decoupled, parallel implementation v.s. coupled, not easy for parallel computing
- The computational complexity to obtain all n localized basis functions $\{\psi_i^{loc}\}_{i=1}^n$ is only of order $N \log(N)$, where N is the degree of freedom in the discretization of $\mathcal{L}$.
- The SOC algorithm [4] solves the l^1 minimization in an iterative manner, where the computational cost of *each iteration* is comparable with the total cost of the Sparse OC.

[4] Lai-Osher, SIAM-JSC, 2014

Sparse OC vs l^1 minimization: 1D free electron

Free electron with periodic boundary condition:

$$\mathcal{L} = -\frac{1}{2}\Delta, \quad D = [0, 50].$$ [5]

- Discretization $\mathcal{L} \in R^{1024\times 1024}$.
- Number of compressed/localized modes $n = 128$.
- Sparse OC takes 0.035 sec to obtain all 128 localized modes, without parallel computing.
- After 390 iterations, the l^1 approach achieves 1e-7 relative energy decrease, and the iteration is stopped. The total time is 4.426 secs. Each iteration takes 0.013 sec.

[5] Lai-Osher, SIAM-JSC, 2014

Sparse OC vs l^1 minimization: 1D free electron

Localized/compressed modes

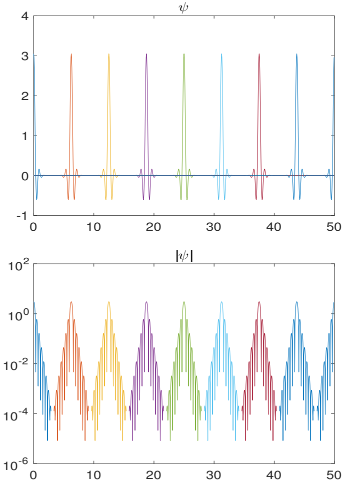


Figure: A few basis functions for the case $m = 2^7$ and $r = h\log_2(1/h)$.

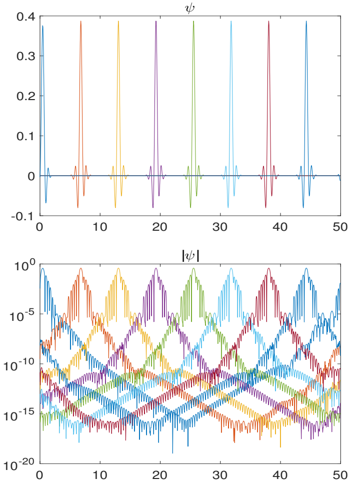


Figure: A few compressed modes, $m = 2^7, \mu = 0.84$

Sparse OC vs l^1 minimization: 1D free electron

Approximate eigenvalues

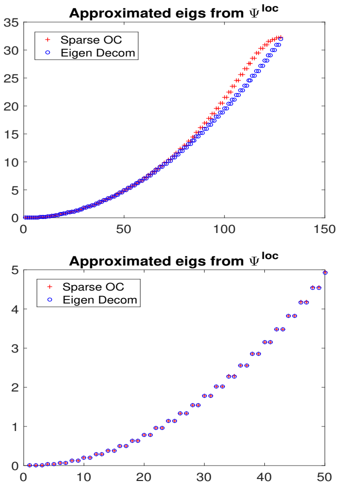


Figure: The eigenvalues of $Q^T H Q$ and H; Q is an orthonormal basis of Ψ^{loc}.

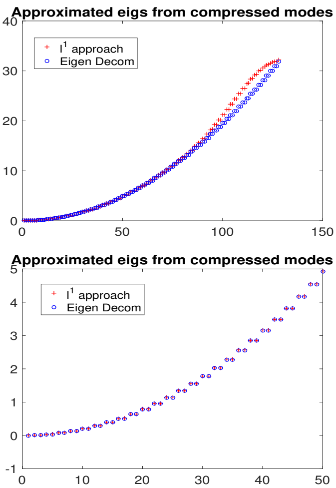


Figure: The eigenvalues of $\Psi^T H \Psi$, $m = 2^7, \mu = 10$

Sparse OC vs l^1 minimization: 1D free electron

Operator compression error

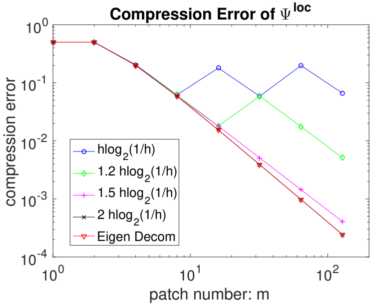


Figure: The operator compression error $E(\Psi;(\mathcal{L}+1)^{-1})$ for the Hamiltonian with localized basis functions Ψ^{loc}.

Sparse OC vs l^1 minimization: 1D free electron

Operator compression error

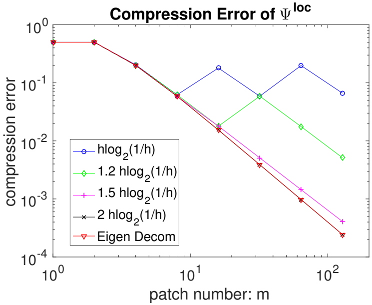


Figure: The operator compression error $E(\Psi; (\mathcal{L}+1)^{-1})$ for the Hamiltonian with localized basis functions Ψ^{loc}.

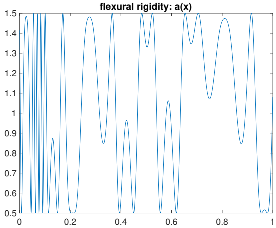
Other related work

- E-Li-Lu, PNAS, 2010 : localization using weight function (algebraical decay)
- Lai-Lu-Osher, CMS, 2015 : convex relaxation of the l^1 approach
- Hou-Li-Zhang, SIAM-MMS, 2016: ISMD for low rank covariance matrices

Fourth order elliptic operators

The 1D biharmonic equation

$$\frac{d^2}{dx^2}\left(a(x)\frac{d^2u}{dx^2}\right) = f(x), \quad 0 < x < 1,$$
$$u(0) = u'(0) = 0, \quad u(1) = u'(1) = 0, \tag{13}$$



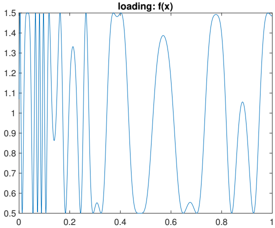
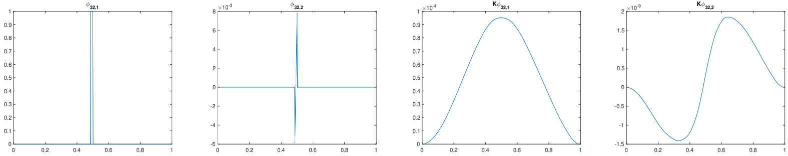
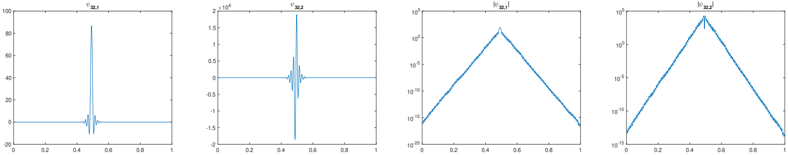


Figure: Highly oscillatory flexural rigidity $a(x)$ and load $f(x)$: no scale separation.

The 1D biharmonic equation: basis functions



(a) Left: $\varphi_{32,1}, \varphi_{32,2}$ for piecewise linear Φ; Right: $\mathcal{L}^{-1}\varphi_{32,1}, \mathcal{L}^{-1}\varphi_{32,2}$



(b) $\psi_{32,1}, \psi_{32,2}$: normal scale and log-scale

Figure: 1D biharmonic operator: piecewise linear Φ. There are two basis functions associated with each patch. The multiscale effect is visible in the basis functions, but the decay rate only depends on the contrast $\frac{a_{max}}{a_{min}}$.

The 1D biharmonic equation: finite element solutions

- Φ_0 space of piecewise constant functions $\Rightarrow \Psi_0 \Rightarrow \Psi_0^{loc} \Rightarrow u_{0,h}$
- Φ_0 space of piecewise linear functions $\Rightarrow \Psi_1 \Rightarrow \Psi_1^{loc} \Rightarrow u_{1,h}$

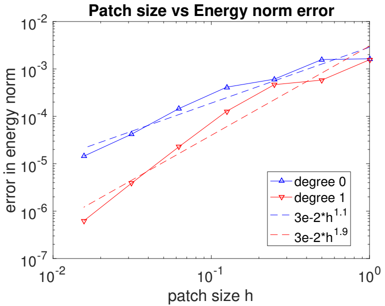


Figure: Error of the finite element solutions: $\|u_{h,0} - u\|_H$ and $\|u_{h,1} - u\|_H$.

The 1D biharmonic equation: finite element solutions

- Φ_0 space of piecewise constant functions $\Rightarrow \Psi_0 \Rightarrow \Psi_0^{loc} \Rightarrow u_{0,h}$
- Φ_0 space of piecewise linear functions $\Rightarrow \Psi_1 \Rightarrow \Psi_1^{loc} \Rightarrow u_{1,h}$

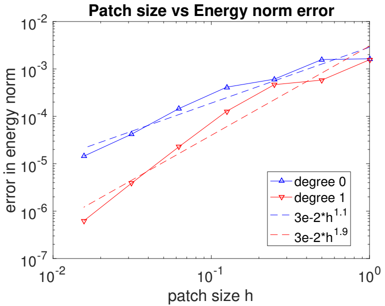


Figure: Error of the finite element solutions: $\|u_{h,0} - u\|_H$ and $\|u_{h,1} - u\|_H$.

To obtain the optimal convergence rate h^k, it is necessary to take Φ as the space of piecewise polynomial space of degree no more than $k-1$.

The 2D biharmonic operator

$$\mathcal{L} = \Delta^2, \quad H^2_{\mathcal{B}} = H^2_0([0,1]^2) \tag{14}$$

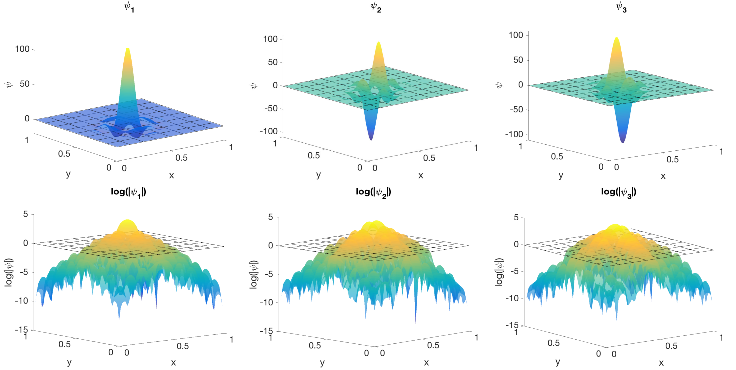


Figure: The three basis functions associated with patch $[1/2 - h_x, 1/2] \times [1/2 - h_y, 1/2]$. They clearly show exponential decay.

Ongoing work and conclusions

Discrete setting: graph Laplacians

$$\mathcal{L}u = -\frac{\mathrm{d}}{\mathrm{d}x}\left(a(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right),$$
$$u(0) = u(1).$$

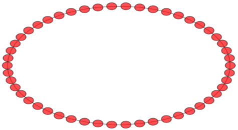


Figure: A 1D circular graph.

$$\mathcal{L}u = -\nabla \cdot (a(x)\nabla u),$$
$$u|_{\partial D} = 0.$$

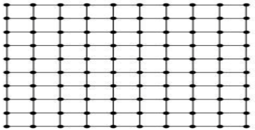


Figure: A 2D lattice graph.

$$\mathcal{L}u = f$$

$\mathcal{L}$: a graph Laplacian

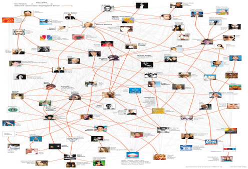


Figure: A social network graph.

- Social networks and transportation networks; genetic data and web pages; spectral clustering of images; electrical resistor circuits; elliptic partial differential equations discretized by finite elements; etc
- Fundamental problems: fast algorithms for $\mathcal{L}u = f$ and eigen decomposition of $\mathcal{L}$.

Discrete setting: graph Laplacians

$$\mathcal{L}u = f$$

- Spielman-Teng (STOC-04, SICOMP-13, SIMAX-14): Nearly-Linear Time Algorithms for Graph Partitioning and Solving Linear Systems
  - Maximal spanning tree, support-graph preconditioners, graph sparsification, etc.
  - Theoretical results, impractical algorithms.
  - Gödel Prize 2008, 2015.
- Livne-Brandt-2012: Lean Algebraic Multigrid. Practical nearly-linear time algorithm, no theoretical guarantee.
- Sparse operator compression for graph Laplacians? The key is an efficient algorithm to find a partition $\{\tau_i\}_{i=1}^m$ of the graph vertices such that

$$\|u - \mathcal{P}_{\Phi}^{(L^2)} u\|_{L^2} \le C_p \sqrt{\lambda_n(\mathcal{L}^{-1})} \|u\|_H,$$

  which is the Poincare inequality on graphs.
- Implementing the sparse operator compression in a multigrid manner leads to a nearly-linear time algorithm.

Conclusions

- We have developed a general strategy to compress self-adjoint second-order and high-order elliptic operators by localized energy-minimizing basis functions.
- For a self-adjoint, bounded and strongly elliptic operator of order $2k$ ($k \geq 1$), we have proved that with support size $h\log(1/h)$, our localized basis functions can obtain the optimal operator compression rate $O(h^{2k})$.
- We have applied our new operator compression strategy in different applications: solving elliptic equations with multiscale coefficients, Sparse PCA for the Matérn class covariance, and compressing Hamiltonians in quantum chemistry.
- Ongoing work on compressing elliptic operators with high contrast coefficients, new multi-grid algorithms for elliptic operators, and fast algorithms for graph partitioning and solving graph Laplacians.

References

1. T. Y. Hou and P. Zhang, “Sparse operator compression of elliptic operators Part I: Second order elliptic operators”, preprint, submitted to RMS, 2016.
2. T. Y. Hou and P. Zhang, “Sparse operator compression of elliptic operators Part II: Higher order elliptic operators”, preprint, submitted to RMS, 2016.