NLA techniques for efficient, reliable and asymptotically exact eigenvalue estimation

Luka Grubišić

Department of Mathematics, University of Zagreb

joint work with: Jeffrey S. Ovall

IWASEP VII, Dubrovnik

Model Problem

- **Problem:** $\Omega \subset \mathbf{R}^2$, bounded, polygonal, possibly non-convex. Find $(\lambda, \phi) \in \mathbb{R} \times \mathcal{H}$, $\psi \neq 0$, $\mathcal{H} = H_0^1(\Omega)$, such that

$$\int_\Omega \nabla\phi \cdot \nabla v = \lambda \int_\Omega \phi v \text{ for all } v \in \mathcal{H}$$

- **Objective:** Adaptively compute a cluster of m eigenvalues $\lambda_i \in [\tau_{\min}, \tau_{\max}]$ and a corresponding invariant subspace $E = \text{span}\{v_1, \cdots v_m\}$. We assume that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and that the dimension of S_h matches the collective multiplicity $[\tau_{\min}, \tau_{\max}]$.

- **Recent Work:** see MPI-MSI preprint for comments on similarities/differences

  Larson '00; Heuveline and Rannacher '01; Neymeyr '02

  Durán, Padra and Rodriguez '03; Mao, Shen and Zhou '06

Discretized Model Problem

- **Triangulation:** Conforming, shape-regular mesh $\mathcal{T}_h$ (mesh parameter h)

  All vertices $z \in \bar{\mathcal{V}}$, interior vertices $z \in \mathcal{V}$, interior edges $e \in \mathcal{E}$

- **Linear Finite Element Space:** $L(\mathcal{T}_h) = \mathrm{span}\{\ell_z | z \in \mathcal{V}\}$

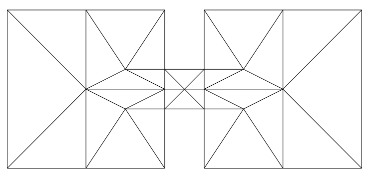
  ℓ_z continuous, linear on $\tau \in \mathcal{T}_h$ and $\ell_z(z') = \delta_{zz'}$ for $z' \in \bar{\mathcal{V}}$

- **Discrete Problem:** Given $\tau \approx \lambda_q$ and $m \approx$ mult(λ_q), solve

  $$\int_\Omega \nabla\psi_i^h \cdot \nabla v = \mu_i^h \int_\Omega \psi_i^h v \text{ for all } v \in L(\mathcal{T}_h),\ i = 1, \ldots, m$$

$K\mathbf{x} = \mu M\mathbf{x}$, solved near τ

(MATLAB eigs)



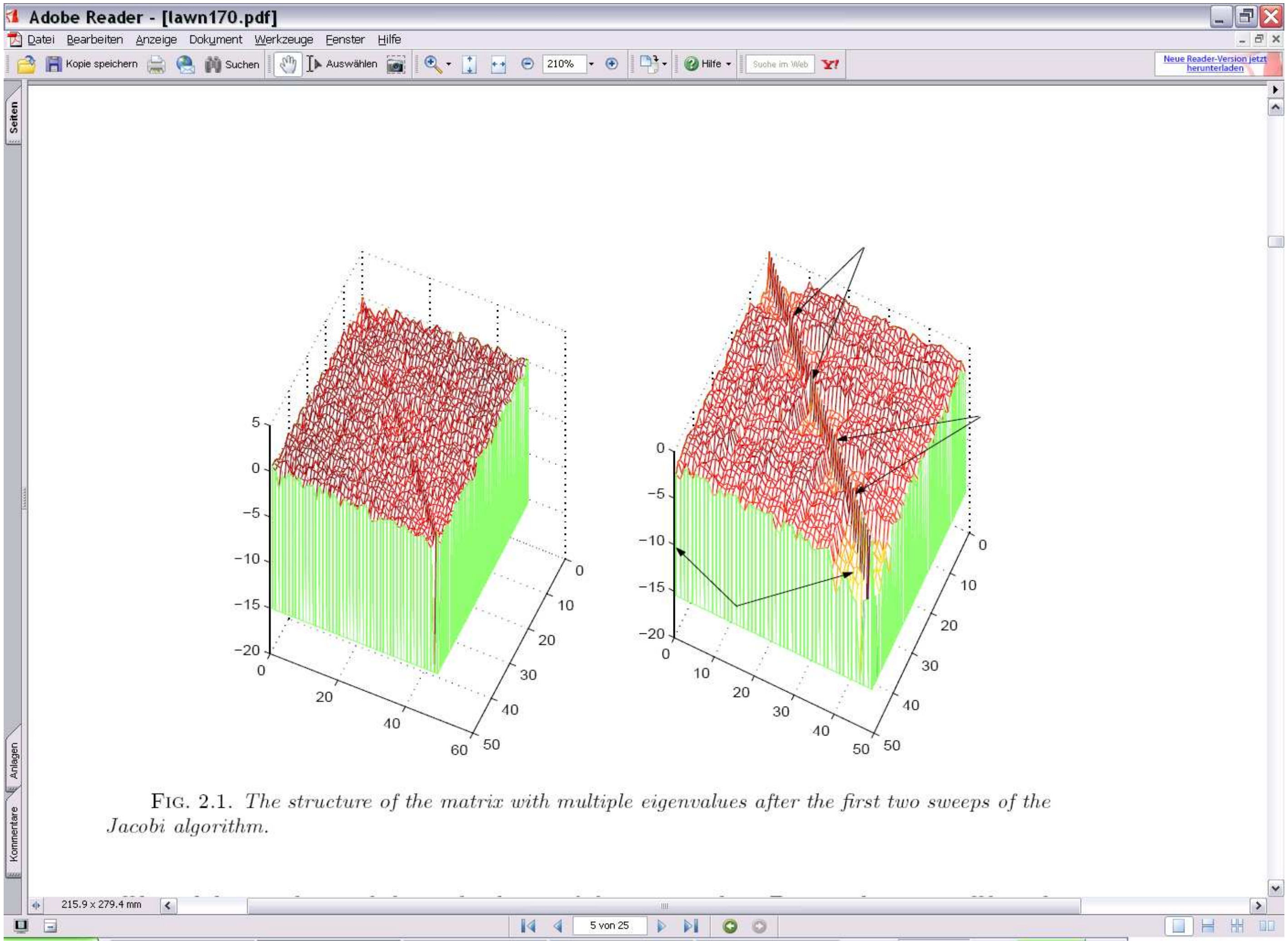
Conceptual challenges/NLA intuition

1. How to **see a residual** which is a functional—not a vector—on $H^1_0(\Omega)$
2. Answer: SCALE IT!
3. Convergence of scaled iterates in diagonalization processes, i.e.

$$\textsc{Scaled} := |\sqrt{\operatorname{diag}(H)} \backslash H / \sqrt{\operatorname{diag}(H)} - I|$$

   has special structure.
4. Scaling means solving very simple auxiliary systems!

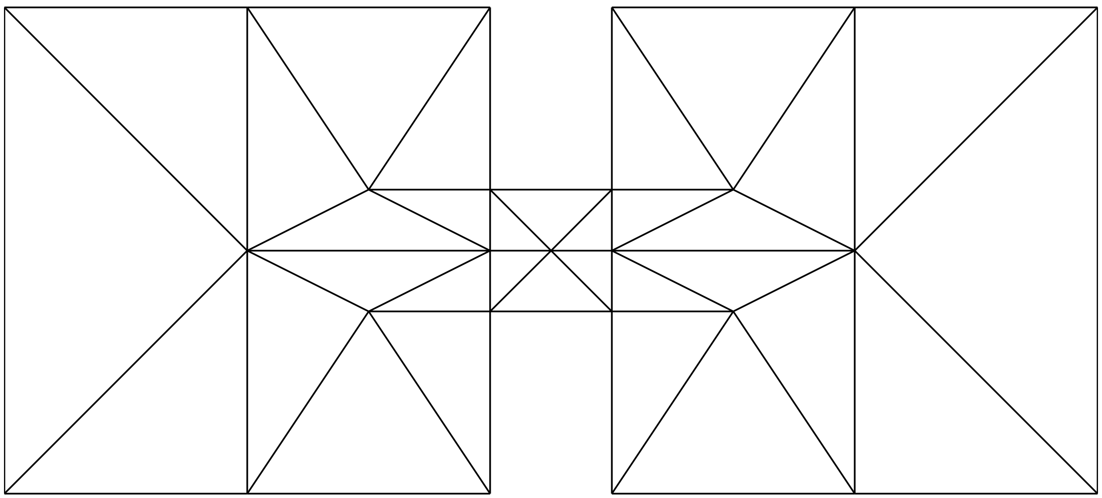
The block structure of quasi-diagonal matrices, plot SCALED



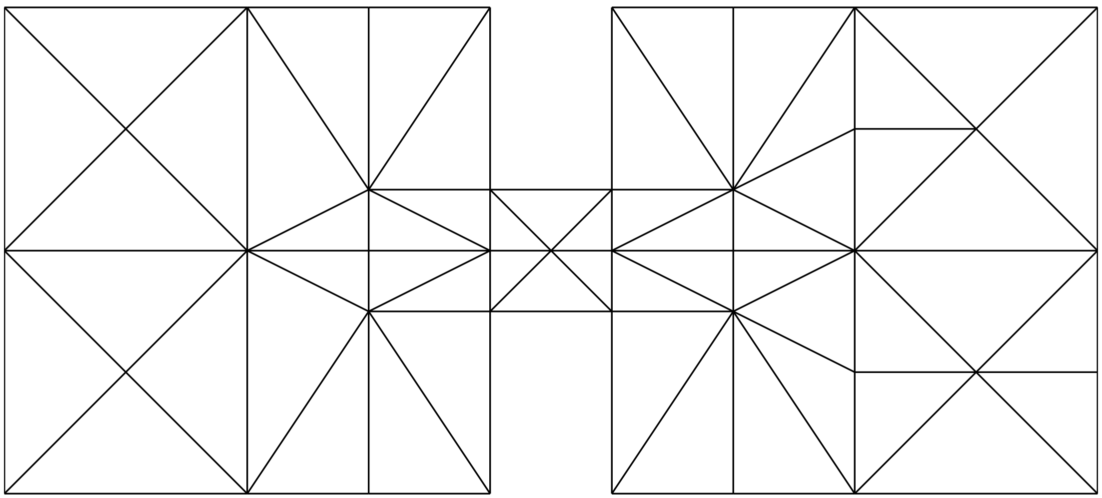
a

[a] Drmač and Veselić, New fast and accurate jacobi svd algorithm: II., SIMAX, 2007

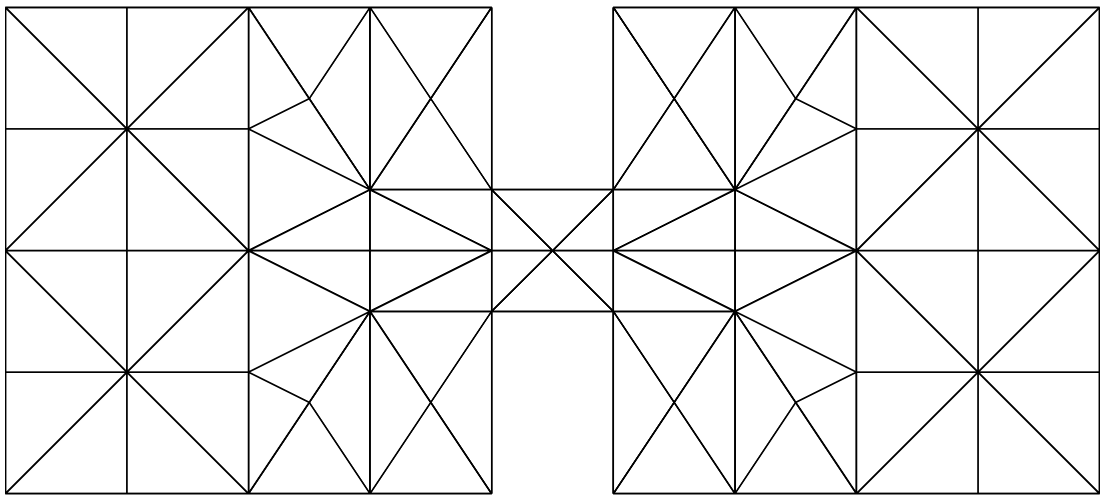
Refined using longest-edge bisection (Rivara '97)



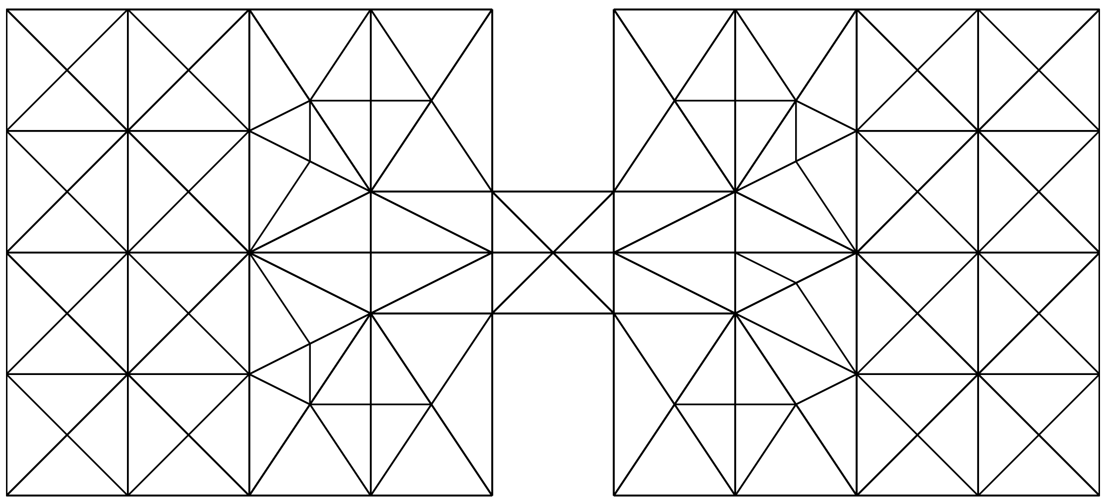
Refined using longest-edge bisection (Rivara '97)



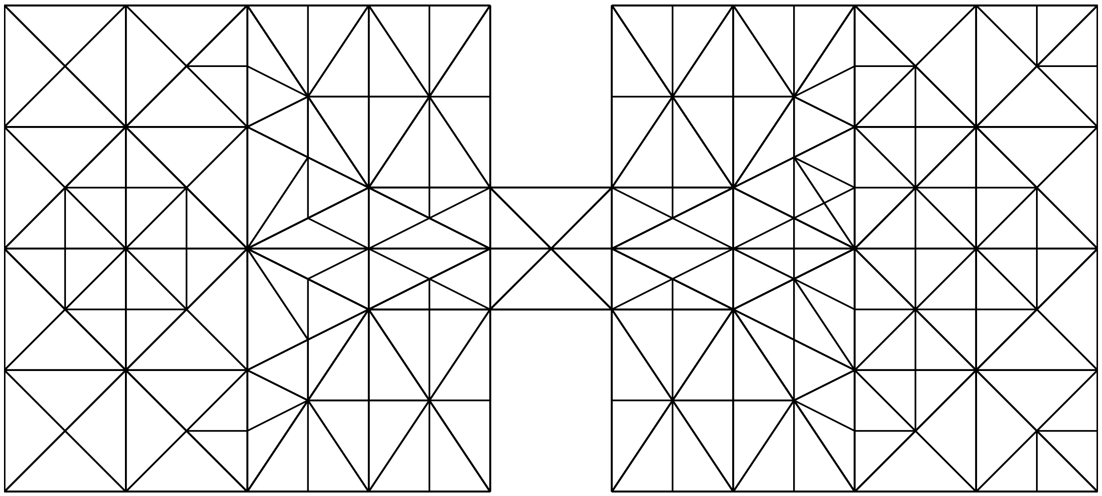
Refined using longest-edge bisection (Rivara ’97)



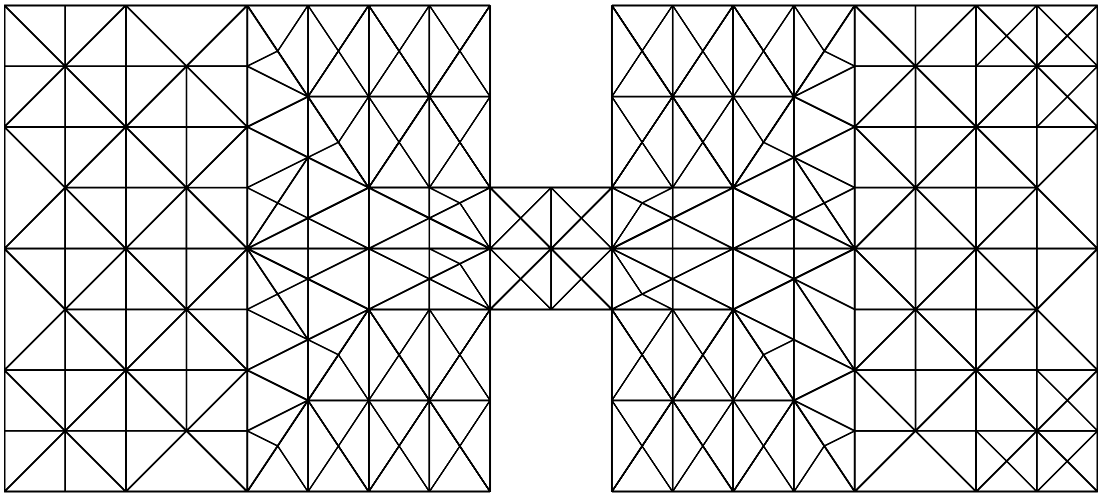
Refined using longest-edge bisection (Rivara ’97)



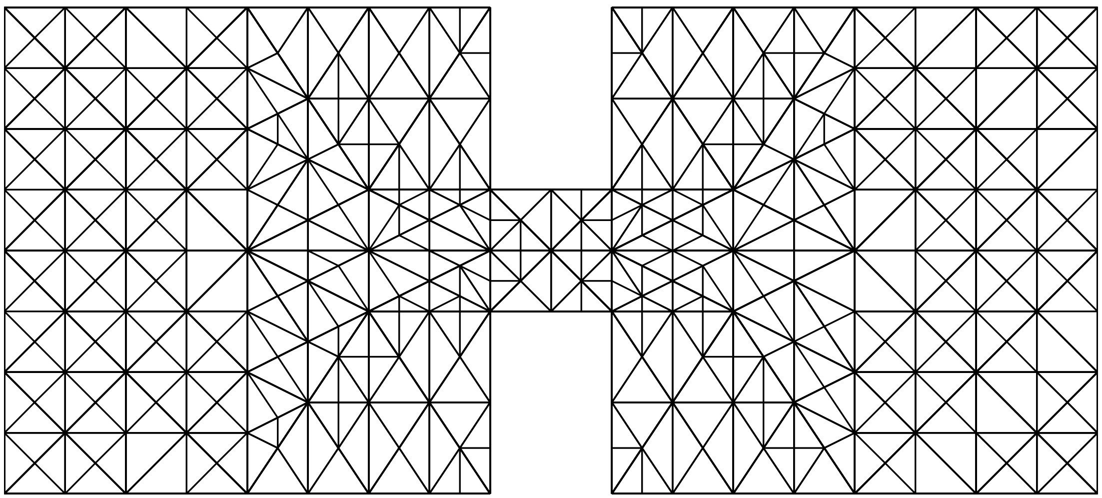
Refined using longest-edge bisection (Rivara '97)



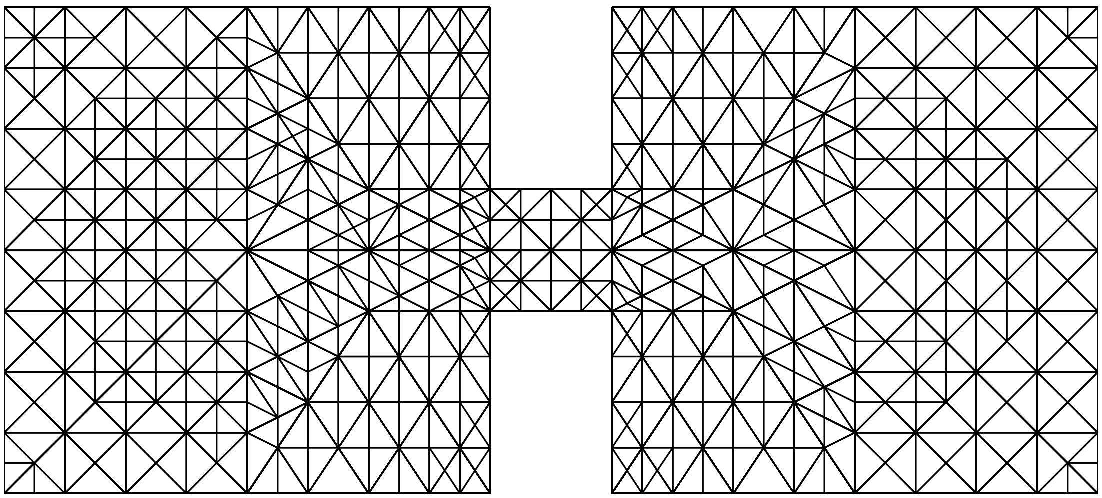
Refined using longest-edge bisection (Rivara '97)



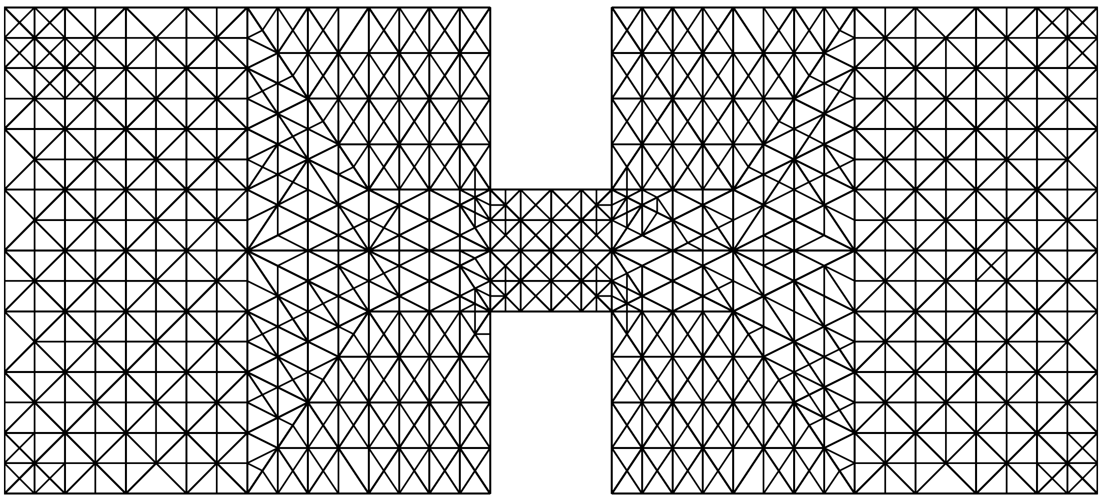
Refined using longest-edge bisection (Rivara '97)



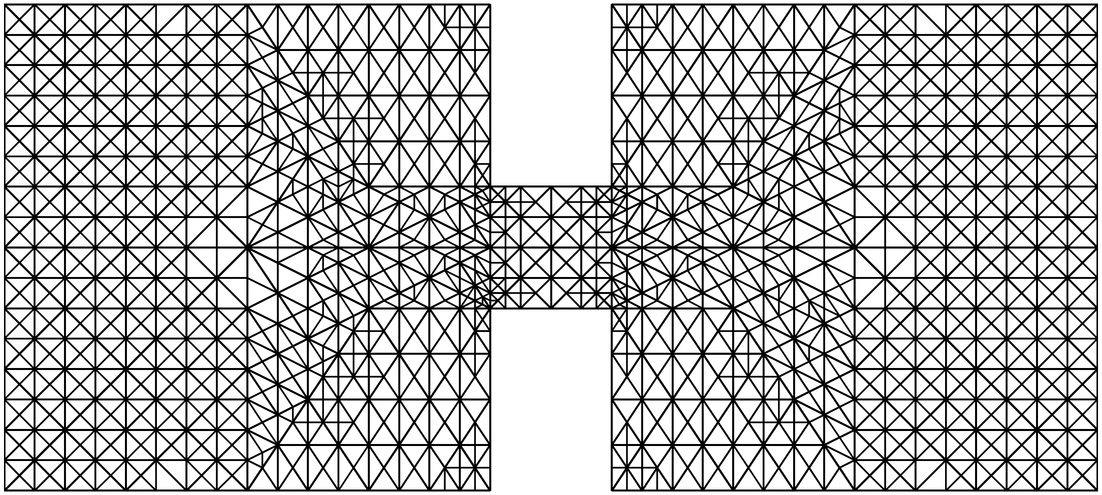
Refined using longest-edge bisection (Rivara '97)



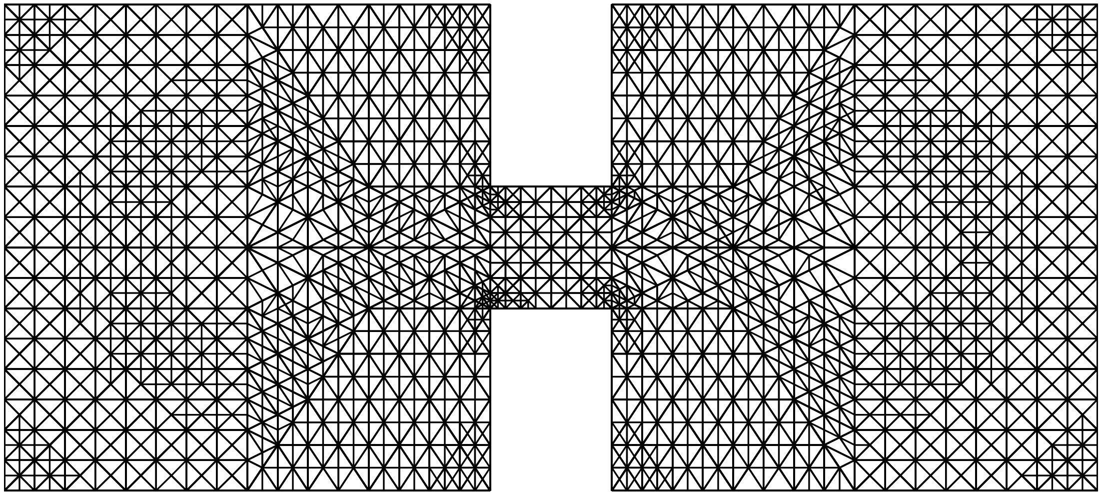
Refined using longest-edge bisection (Rivara '97)



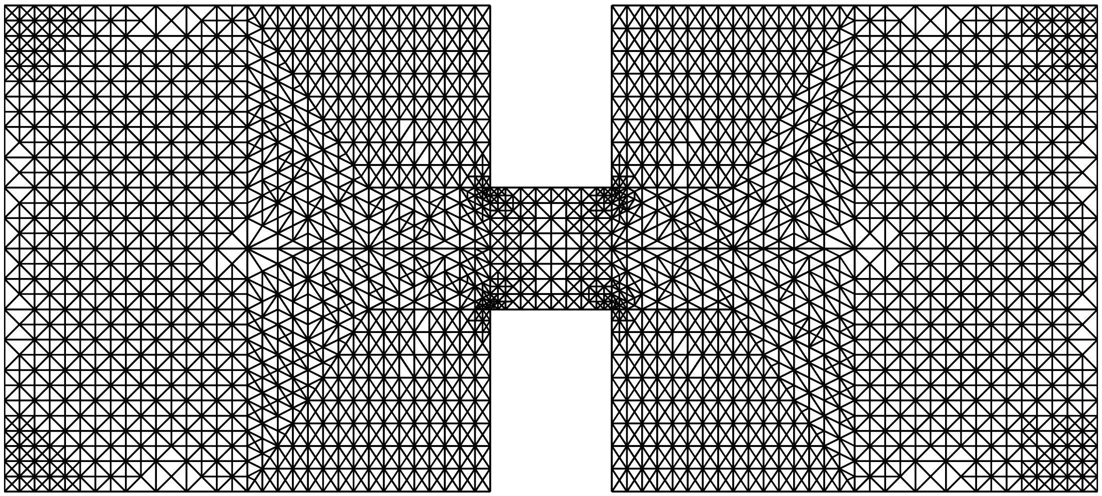
Refined using longest-edge bisection (Rivara '97)



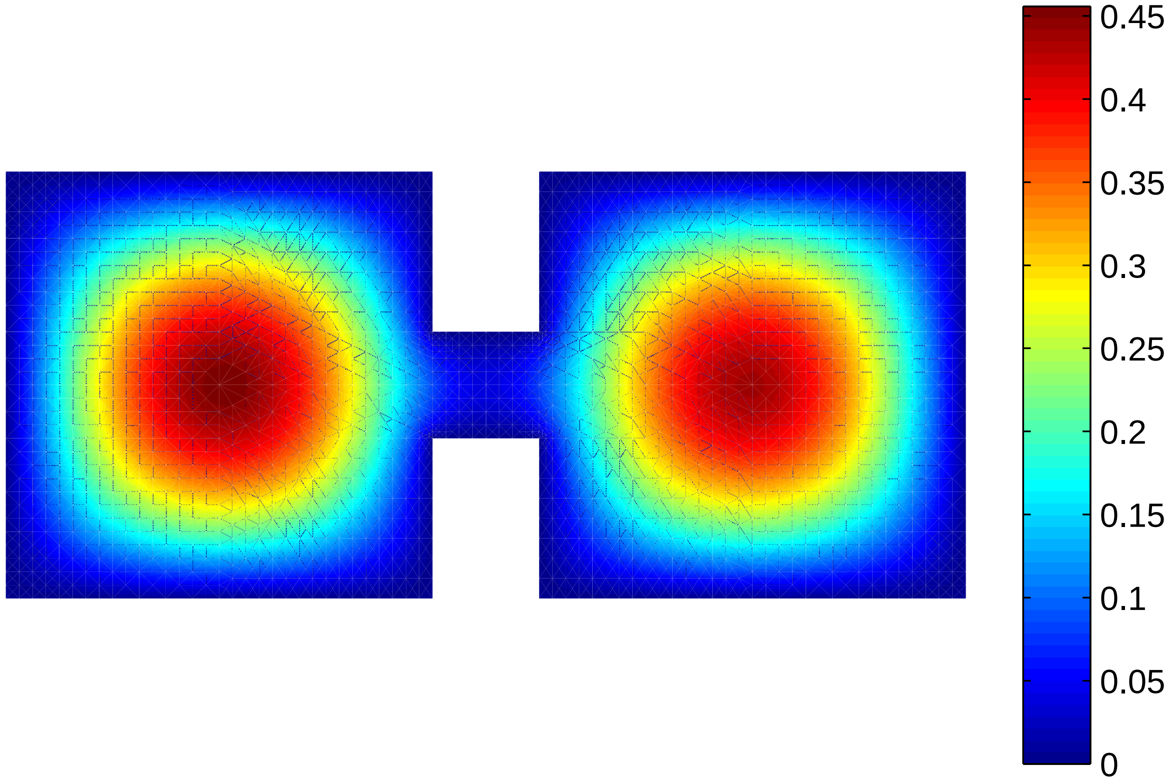
Refined using longest-edge bisection (Rivara '97)



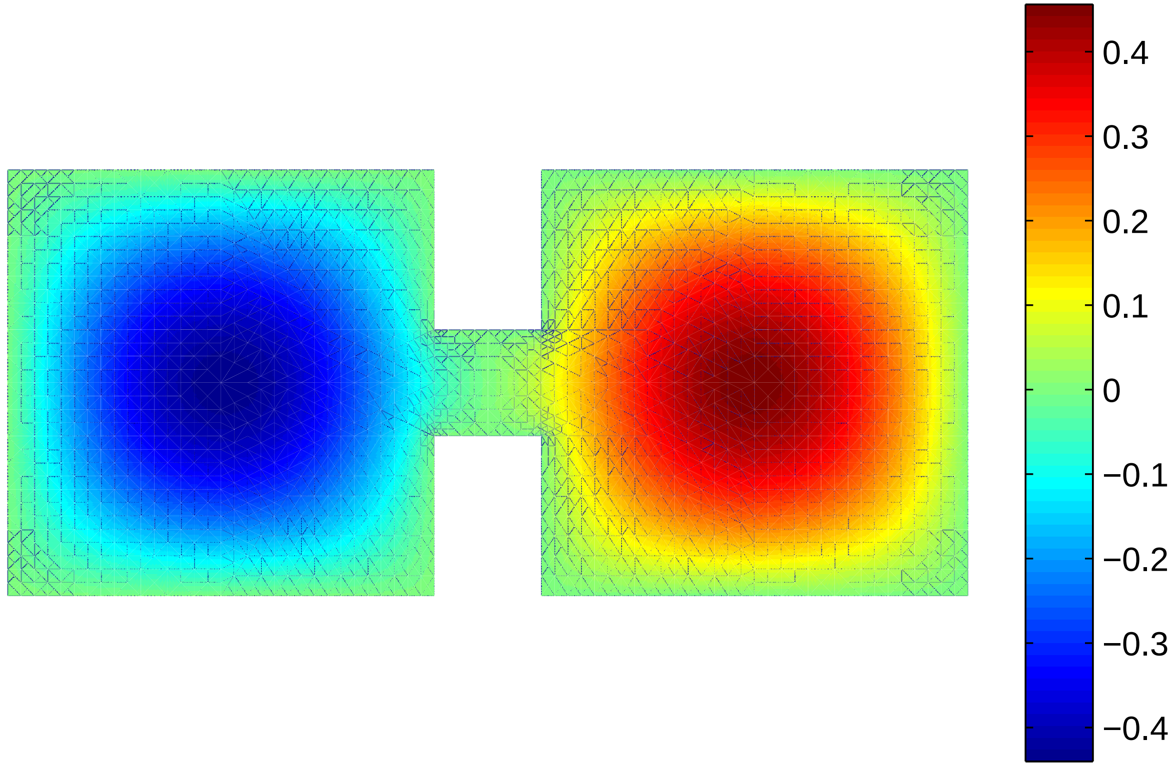
Refined using longest-edge bisection (Rivara ’97)



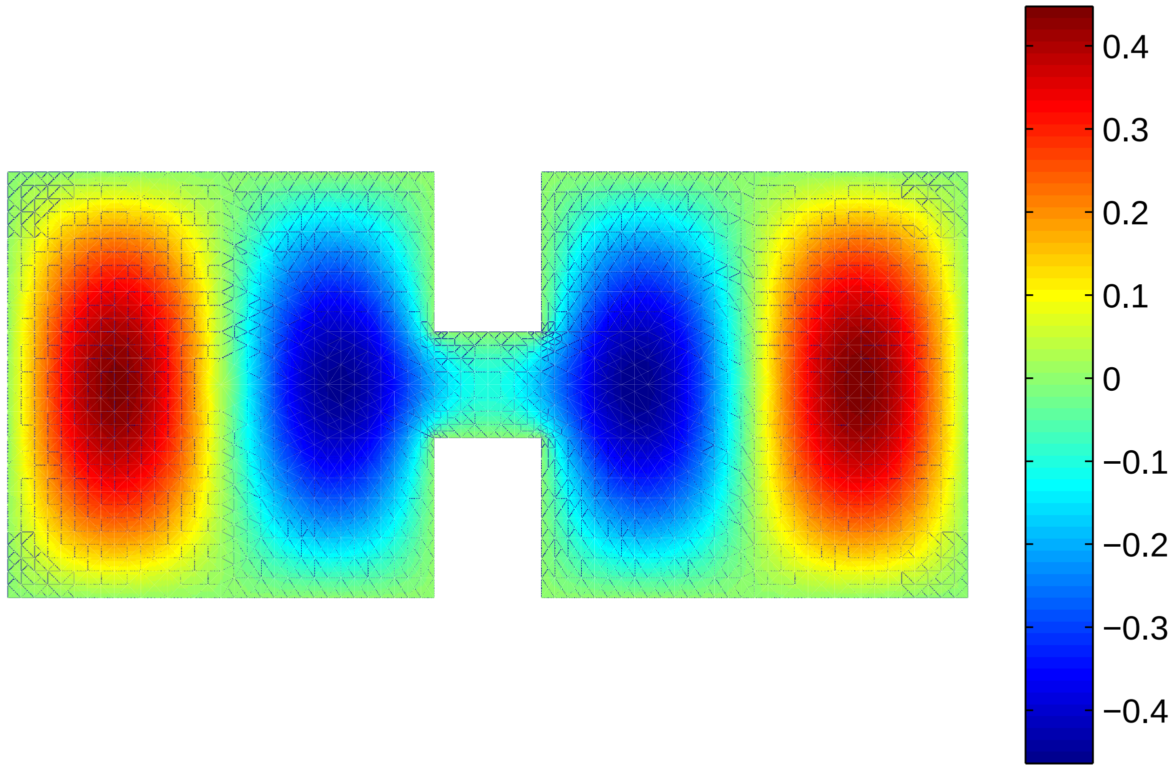
Refined using longest-edge bisection (Rivara ’97)



Refined using longest-edge bisection (Rivara '97)



Refined using longest-edge bisection (Rivara '97)



Convergence in numbers

Effectivity: $EFF_1 = \left(\sum_{i=1}^{6} \frac{|\mu_i^h - \lambda_i|}{\mu_i^h} \right) / \left(\sum_{i=1}^{6} \sigma_i^2(\text{RES\_SCALED}) \right)$

N	μ_6	μ_5	μ_4	μ_3	μ_2	μ_1	EFF_1
85	5.6138	5.6138	5.5491	5.5184	2.0846	2.0809	1.1725
243	5.2183	5.2128	5.0333	5.0083	1.9967	1.9929	1.1130
650	5.0778	5.0765	4.9095	4.8847	1.9759	1.9718	1.1065
1693	5.0279	5.0278	4.8630	4.8345	1.9673	1.9626	1.1058
4274	5.0093	5.0093	4.8428	4.8143	1.9630	1.9583	1.0944

What does the job?

- Write the problem as a 2×2 block operator matrix and use the standard **Wilkinson's Schur Complement Trick** to obtain

$$\text{ERROR} = \text{RES\_SCALED}^{\text{DUAL}} \cdot \text{GAP} \cdot \text{RES\_SCALED}$$

which can be written as

$$\text{ERROR} = \text{RES\_SCALED}^{\text{DUAL}} \cdot \text{RES\_SCALED} + \text{RES\_SCALED}_{\mathbf{2}}^{\text{DUAL}} \cdot \text{GAP} \cdot \text{RES\_SCALED}_{\mathbf{2}}.$$

- Use some Resolvent tricks on the quadratic estimates from Parlett's book or Drmač–Hari relative version, cf. Mathias on RQ estimates.

Assessment of Approximations — Approximation Defects

- **Full Solution:** Given $\psi \in \mathsf{R}(P_h) = \text{span}\{\psi_k^h\}_{k=1}^m$, solve for $u(\psi) \in \mathcal{H}$,
$$\int_\Omega \nabla u(\psi) \cdot \nabla v = \int_\Omega \psi v \text{ for all } v \in \mathcal{H}$$
- **PW-Linear Solution:** Given $\psi \in \mathsf{R}(P_h)$, solve for $u_h(\psi) \in L(\mathcal{T}_h)$,
$$\int_\Omega \nabla u_h(\psi) \cdot \nabla v = \int_\Omega \psi v \text{ for all } v \in L(\mathcal{T}_h)$$
- **Approximation Defects:**
$$\eta_i^2(P_h) = \max_{\substack{\mathcal{S} \subset \mathsf{R}(P_h) \\ \dim \mathcal{S} = m-i+1}} \min_{\substack{\psi \in \mathcal{S} \\ \|\psi\|=1}} \frac{\|\nabla u(\psi) - \nabla u_h(\psi)\|^2}{\|\nabla u(\psi)\|^2}$$
- **Error and Gradient Matrices:** $E, G \in \mathbb{R}^{\mathbf{m} \times \mathbf{m}}$ then solve for η
$$E\mathbf{x} = \eta^2 G\mathbf{x}$$

Our theorems

- **Asymptotic Exactness:** Under standard convergence, non-degeneracy assumptions (cf. Hackbusch '79, Durán/Padra/Rodriguez '03)

$$\lim_{h\to 0} \frac{\sum_{i=1}^m \frac{|\lambda_q - \mu_i^h|}{\mu_i^h}}{\sum_{i=1}^m \eta_i^2(P_h)} = 1 \quad , \quad \lim_{h\to 0} \frac{\sum_{i=1}^m \frac{\|\nabla\phi_i - \nabla\psi_i^h\|^2}{\|\nabla\phi_i\|^2}}{\sum_{i=1}^m \eta_i^2(P_h)} = 1$$

- **Equivalence:** Without such assumptions, or in pre-asymptotic range

  Equivalence of $\sum_{i=1}^m \eta_i^2(P_h)$ and $\sum_{i=1}^m \frac{|\lambda_q - \mu_i^h|}{\mu_i^h}$ or $\sum_{i=1}^m \frac{\|\nabla\phi_i - \nabla\psi_i^h\|^2}{\|\nabla\phi_i\|^2}$

  Sharp upper and lower bounds – Grubišić '05, Grubišić/Ovall '07

- Results also hold for similar measures of relative error

Motivation for estimating approximation defects, $\eta_i(P_h)$

Estimating the Approximation Defects

Approximation Defects

- Trustworthy estimates of relative eigenvalue/vector errors
- Relative discretization errors for standard "source problems"
- Sources from computed subspace $\mathsf{R}(P_h)$
- "Computed" via small generalized eigenvalue problem $E\mathbf{x} = \eta^2 G\mathbf{x}$

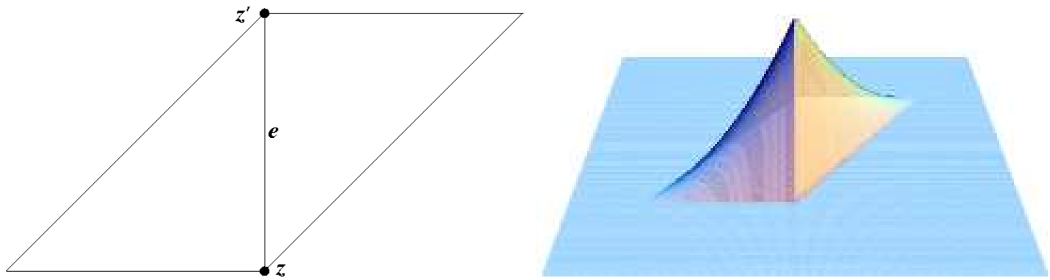
Estimating Approximation Defects

- Solve a nearby generalized eigenvalue problem $\tilde{E}\mathbf{x} = \tilde{\eta}^2 \tilde{G}\mathbf{x}$
- $\tilde{E}$ and $\tilde{G}$ constructed from **computable** estimates of $\nabla(u(\psi) - u_h(\psi))$ and $\nabla u(\psi)$, for $\psi = \psi_j^h,\ j = 1, \ldots, m$
- Some viable estimation methods: gradient recovery, hierarchical basis, equilibrated residual

Hierarchical Basis Error Estimation

Edge Bump Space: $B(\mathcal{T}_h) = \mathrm{span}\{b_e = 4\ell_z \ell_{z'} | e \in \mathcal{E}\}$

Quadratic Space: $Q(\mathcal{T}_h) = L(\mathcal{T}_h) \oplus B(\mathcal{T}_h) \rightsquigarrow$ DISCR $= \begin{bmatrix} LL & LB \\ BL & BB \end{bmatrix}$



Capture the scaled residual in the bump space $B(\mathcal{T}_h)$. Bump space is:

- Sufficiently large to resolve the residual.
- The bump-bump matrix is almost diagonal (spectral equivalence to a diagonal with explicit constants).

Our Estimates of Approximation Defects

Approximation Defects and Estimates: Only $H^1(\Omega)$ assumed!

$$\eta_i^2(P_h) = \max_{\substack{\mathcal{S}\subset \mathsf{R}(P_h) \\ \dim \mathcal{S}=m-i+1}} \min_{\substack{\psi\in\mathcal{S} \\ \|\psi\|=1}} \frac{\|\nabla(u(\psi)-u_h(\psi))\|^2}{\|\nabla u(\psi)\|^2}$$

$$\tilde{\eta}_i^2(P_h) = \max_{\substack{\mathcal{S}\subset \mathsf{R}(P_h) \\ \dim \mathcal{S}=m-i+1}} \min_{\substack{\psi\in\mathcal{S} \\ \|\psi\|=1}} \frac{\|\nabla \varepsilon_h(\psi)\|^2}{\|\nabla u_h(\psi)\|^2 + \|\nabla \varepsilon_h(\psi)\|^2}$$

The η estimator:

$$1 \le \frac{\eta_i(P_h)}{\tilde{\eta}_i(P_h)} \le C_1(\mathcal{T}_h) + \max_{\substack{\psi\in \mathsf{R}(P_h) \\ \|\psi\|=1}} \frac{\mathsf{osc}(\psi,\mathcal{T}_h)}{\|\nabla\varepsilon_h(\psi)\|}$$

$$\le C_1(\mathcal{T}_h) + \max_{\substack{\psi\in \mathsf{R}(P_h) \\ \|\psi\|=1}} \frac{C_2(\mathcal{T}_h)h^2\sqrt{\mu_m^h}}{\|\nabla\varepsilon_h(\psi)\|}.$$

- $C_1(\mathcal{T}_h)$, $C_2(\mathcal{T}_h)$ scale-invariant, can be **(over)-estimated in terms of mesh angles**

Experiment 1 — Simple Problem

Problem: Estimate smallest six eigenvalues on unit square

$10\pi^2\ ,\ 10\pi^2\ ,\ 8\pi^2\ ,\ 5\pi^2\ ,\ 5\pi^2\ ,\ 2\pi^2 \qquad (k^2+n^2)\pi^2 \leftrightarrow 2\sin(k\pi x)\sin(n\pi y)$

Effectivity: $EFF_1 = \left(\sum_{i=1}^{6} \frac{|\mu_i^h - \lambda_i|}{\mu_i^h}\right) / \left(\sum_{i=1}^{6} \tilde{\eta}_i^2(P_h)\right)$

N	μ_6	μ_5	μ_4	μ_3	μ_2	μ_1	EFF_1
45	113.1352	112.4708	91.0246	53.3258	53.3074	20.4309	1.1270
111	106.5602	106.1753	82.5317	50.9809	50.9774	19.9609	1.0852
295	101.6302	101.3446	80.7847	50.1042	50.0528	19.8777	1.0742
731	99.7023	99.6333	79.7179	49.6299	49.6294	19.7851	1.0634
1765	99.2275	99.1885	79.2009	49.4685	49.4685	19.7592	1.0619
4248	98.8957	98.8811	79.1084	49.4033	49.4031	19.7494	1.0546

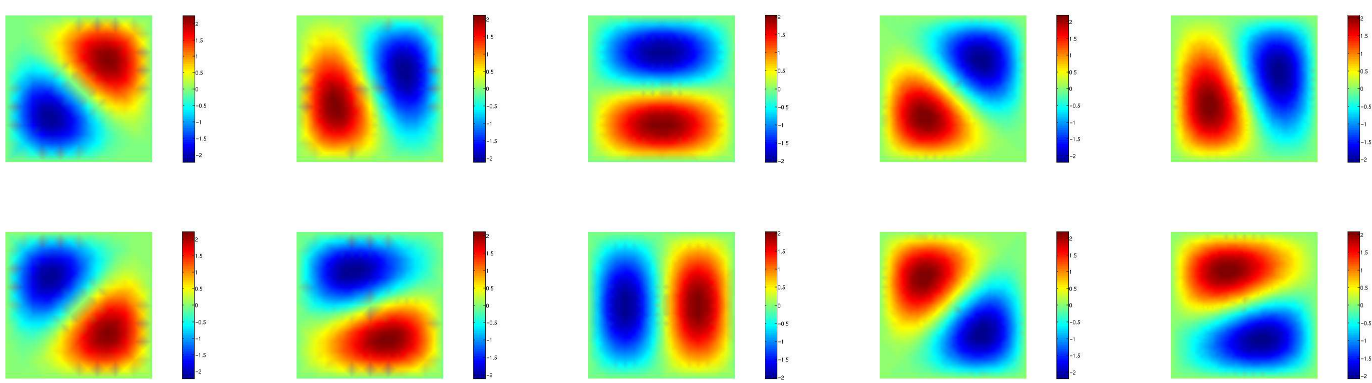
Experiment 2 — Simple Problem

Problem: Estimate subspace for degenerate eigenpair on unit square

$$5\pi^2 \leftrightarrow \text{span}\{2\sin(2\pi x)\sin(\pi y),\ 2\sin(\pi x)\sin(2\pi y)\} = \text{span}\{v_1\ ,\ v_2\}$$

Effectivity: $EFF_2 = \left(\sum_{i=1}^{2} \frac{\min_{\psi\in \mathrm{R}(P_h)} \|\nabla(\psi - v_i)\|^2}{5\pi^2}\right) \Big/ \left(\sum_{i=1}^{2} \tilde{\eta}_i^2(P_h)\right)$

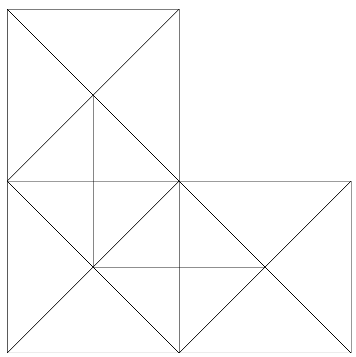
N	5	9	15	29	45	73	129
EFF_2	1.8844	1.7255	1.4649	1.1744	1.1408	1.1119	1.0940
N	215	353	589	891	1385	2193	3369
EFF_2	1.0721	1.0754	1.0637	1.0551	1.0633	1.0608	1.0534

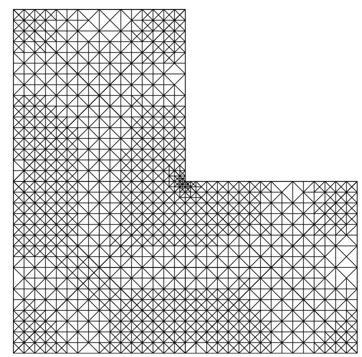


Experiment 3 — L shape

Problem: Estimate smallest six eigenvalues on L shape domain

$41.474510\ ,\ 31.912636\ ,\ 29.521481\ ,\ 19.739209\ ,\ 15.197252\ ,\ 9.6397238$[a]

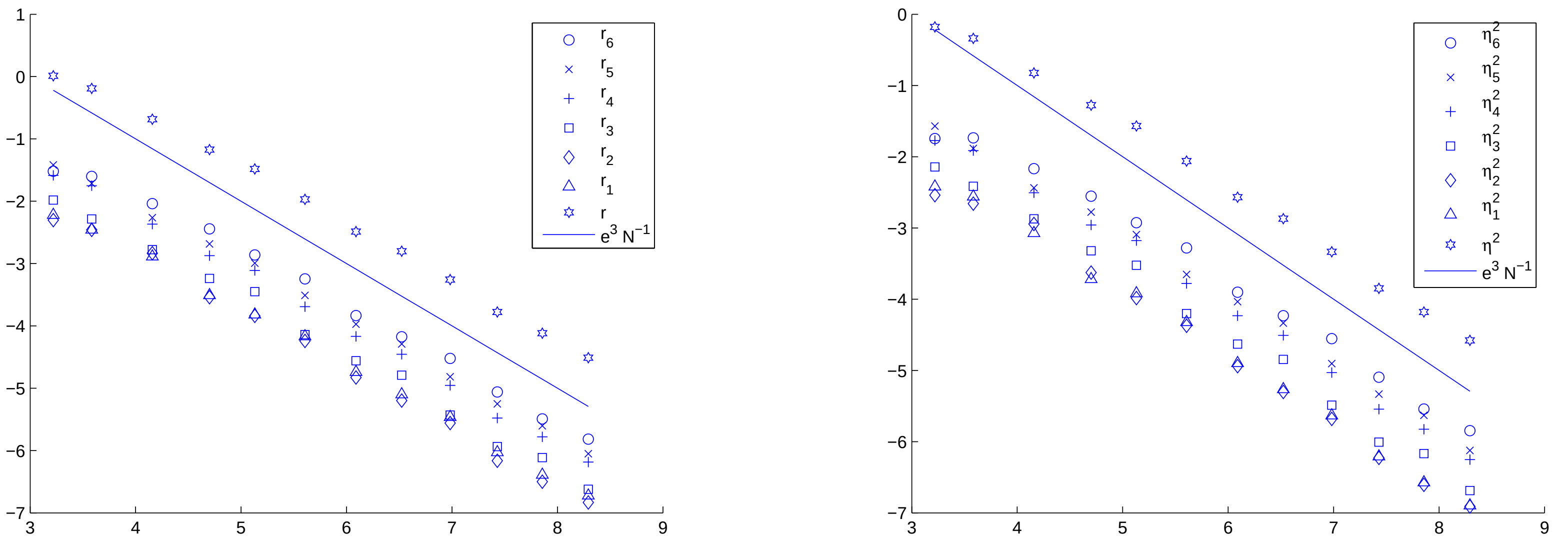




N	eff$_6$	eff$_5$	eff$_4$	eff$_3$	eff$_2$	eff$_1$	EFF
36	1.1413	1.1829	1.1780	1.1390	1.2196	1.1092	1.1617
110	1.1166	1.0989	1.0870	1.0843	1.0897	1.2348	1.1107
680	1.0581	1.0460	1.0534	1.0568	1.1105	1.1788	1.0700
3998	1.0280	1.0747	1.0693	1.0669	1.0816	1.1889	1.0718

[a] Trefethen and Betcke, 2006

Convergence histories for η-s and rel_error



Outlook

- **Accomplished:**

  Reliable estimates of relative eigenvalue/vector approximation error

  Based on hierarchical basis error estimation

  Other error estimation techniques possible in our framework

- **Gradient Recovery Version:**

  $\nabla u_h \mapsto \mathcal{R}\nabla u_h$ (Zienkiewicz/Zhu, Bank/Xu, Zhang/Naga, etc.)

  $$\tilde{\eta}_i(P_h) = \max_{\substack{\mathcal{S}\subset \mathsf{R}(P_h) \\ \dim \mathcal{S} = m-i+1}} \min_{\substack{\psi\in\mathcal{S} \\ \|\psi\|=1}} \frac{\|\mathcal{R}\nabla u_h(\psi) - \nabla u_h(\psi)\|}{\|\mathcal{R}\nabla u_h(\psi)\|}$$

  Compare effectivity and computational/memory costs with original

- **Extensions of Error Estimation Argument:** (future work)

  More general differential operators and boundary conditions

  Problems in $\mathbb{R}^n$, $n > 2$

References

1. R. E. Bank and J. Xu. Asymptotically exact a posteriori error estimators. I. Grids with superconvergence. *SIAM J. Numer. Anal.*, 41 (6): 2294–2312 (electronic), 2003.

2. R. E. Bank and J. Xu. Asymptotically exact a posteriori error estimators. II. General unstructured grids. *SIAM J. Numer. Anal.*, 41 (6): 2313–2332 (electronic), 2003.

3. C. Carstensen. Quasi-Interpolation and A Posteriori Error Analysis in Finite Element Methods. *M2AN Math. Model. Numer. Anal.* 33(6): 1187–1202, 1999.

4. S. Chua and R. L. Wheeden. Estimates of Best Constants for Weighted Poincaré Inequalities on Convex Domains. *Proc. London Math. Soc.* 93(3):197–226, 2006.

References

5. W. Dörfler and R. H. Nochetto. Small data oscillation implies the saturation assumption. *Numer. Math.*, 91(1):1–12, 2002.

6. R. G. Durán, C. Padra, and R. Rodriguez. A posteriori error estimates for the finite element approximation of eigenvalue problems. *Math. Models Methods Appl. Sci.*, 13(8):1219–1229, 2003.

7. L. Grubišić. A posteriori estimates for eigenvalue/vector approximations. *PAMM Proc. Appl. Mathi. Mech.*, 6(1):59 – 62, 2006.

8. L. Grubišić. On Temple–Kato like inequalities and applications. *submitted* 2005–Preprint available from `http://arxiv.org/abs/math/0511408`.

References

9. W. Hackbusch. On the computation of approximate eigenvalues and eigenfunctions of elliptic operators by means of a multi-grid method. *SIAM J. Numer. Anal.*, 16(2), 201–215, 1979.
10. M. G. Larson. A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems. *SIAM J. Numer. Anal.*, 38(2):608–625 (electronic), 2000.
11. J. F. Maitre and F. Musy. The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems. *Multigrid methods (Cologne, 1981)*. Lecture Notes in Math., 960, Springer, Berlin-New York, pp. 535–544, 1982.
12. D. Mao, L. Shen, and A. Zhou. Adaptive finite element algorithms for eigenvalue problems based on local averaging type a posteriori error estimates. *Adv. Comput. Math.*, 25(1-3):135–160, 2006.

References

13. P. Morin, R. H. Nochetto, and K. G. Siebert. Local Problems on Stars: A Posteriori Error Estimators, Convergence, and Performance. *Math. Comp.*, 72 (243): 1067–1097 (electronic), 2002.

14. K. Neymeyr. A posteriori error estimation for elliptic eigenproblems. *Numer. Linear Algebra Appl.*, 9(4):263–279, 2002.

15. J. S. Ovall. Function, Gradient and Hessian Recovery Using Quadratic Edge-Bump Functions. To appear in *SIAM J. Numer. Anal*.

16. R. Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4):695–713, 1999.

References

13. P. Morin, R. H. Nochetto, and K. G. Siebert. Local Problems on Stars: A Posteriori Error Estimators, Convergence, and Performance. *Math. Comp.*, 72 (243): 1067–1097 (electronic), 2002.

14. K. Neymeyr. A posteriori error estimation for elliptic eigenproblems. *Numer. Linear Algebra Appl.*, 9(4):263–279, 2002.

15. J. S. Ovall. Function, Gradient and Hessian Recovery Using Quadratic Edge-Bump Functions. To appear in *SIAM J. Numer. Anal.*

16. R. Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4):695–713, 1999.

Thanks for your attention!

Our Adaptive Refinement Strategy

Local Indicators: For $\tau \in \mathcal{T}$, $\tilde{\eta}_\tau^2(P_h) = \sum_{i=1}^{m} \tilde{\eta}_{i,\tau}^2(P_h)$

$$\tilde{\eta}_{i,\tau}^2(P_h) = \frac{\|\nabla \varepsilon_h(\bar{\psi}_i)\|_\tau^2}{\|\nabla \varepsilon_h(\bar{\psi}_i)\|^2 + \|\nabla u_h(\bar{\psi}_i)\|^2}$$ where $\bar{\psi}_i \in \mathrm{R}(P_h)$ satisfies

$$\tilde{\eta}_i^2(P_h) = \frac{\|\nabla \varepsilon_h(\bar{\psi}_i)\|^2}{\|\nabla \varepsilon_h(\bar{\psi}_i)\|^2 + \|\nabla u_h(\bar{\psi}_i)\|^2}$$ $\bar{\psi}_i$ obtained from $\tilde{E}\mathbf{x}_i = \tilde{\eta}_i^2 \tilde{G}\mathbf{x}_i$

Independent of the computed basis for $\mathrm{R}(P_h)$

Marking and Refining:

- Marked triangles whose local indicator was larger than the median
- Refined using longest-edge bisection (Rivara '97)