The Quasi-Sparse Eigenvector Method

Ilse Ipsen

North Carolina State University

Joint work with:

Dean Lee and Nathan Salwen, Physics, NCState

Goal

At low temperatures behaviour of quantum systems determined by properties of low energy eigenstates.

Determine low energy eigenstates for solids containing:

- many molecules
- molecules with many atoms
- atoms with many electrons
- nuclei with many protons and neutrons
- hadrons (mesons and baryons) with quarks and gluons

many degrees of freedom $\Longrightarrow$ infinite-dimensional problems

Applications

- Hadronic physics:
  quark structure of mesons and baryons
  creation of quark-gluon plasma $\Longrightarrow$ origin of universe
- Electron-hole model for superconductors
- Atomic physics: Be-like Ions $\Longrightarrow$ astrophysics
- Nuclear physics:
  accurate description of complex nuclear reactions
- Quantum electrodynamics

Example

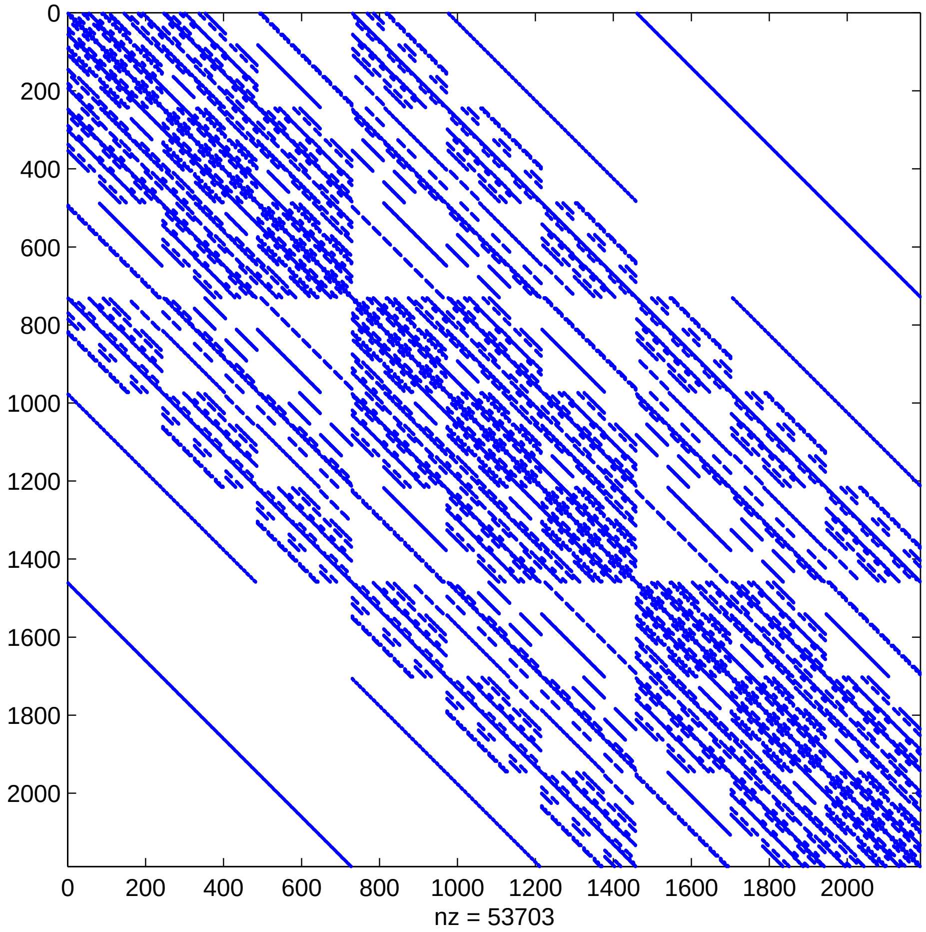
10×10 lattice of spins, 2 spins per site (up, down)

$$\underbrace{\left.\begin{matrix} \uparrow & \downarrow & \cdots & \uparrow \\ \downarrow & \uparrow & \cdots & \downarrow \\ \vdots & \vdots & & \vdots \\ \downarrow & \downarrow & \cdots & \downarrow \end{matrix}\right\}}_{10} \quad 10$$

Hilbert space of $2^{100} \approx 10^{30}$ dimensions

Computational Problem

Compute small eigenvalues & eigenvectors of infinite-dimensional Hamiltonian matrices



matrix order = 2187, n = 3, noccmax = 2, ell = 1, a2 = a4=1

Overview

- Computational methods
- Quasi-Sparse Eigenvector method (QSE)
- Structure of Hamiltonians
- Ideal QSE
- One iteration of (non-ideal) QSE
  - Absolute Bounds
  - Relative Bounds
- Accuracy Requirements
- Really Simple QSE

Computational Methods

- Monte Carlo
  - \+ little storage, fast, parallelizable, scales up
  - \- difficulty with sign & phase oscillations,
    no explicit information about eigenstate wave functions & excited states
- Diagonalization
  - \+ no sign problems
    info about spectrum & eigenstate wave functions
  - \- storage and time exponential in system size

Tamm-Dancoff truncation [Perry, Harindranath, Wilson 1950]

Similarity transformations [Wilson 1965, 1994]

Density matrix renormalization group [White 1992, 1993]

Stochastic diagonalization [Husslein, Fettes, Morgenstern 1997]

Diagonalization/Monte Carlo Method

[Lee, Salwen, Windoloski 2001], [Lee, Salwen, Lee 2001], [Salwen 2001]

1. Quasi-Sparse Eigenvector Method:
   Diagonalization restricted to a subspace
   Can find several eigenstates simultaneously
2. Monte Carlo:
   Stochastic error correction
   Samples contribution of missing basis vectors

Quasi-Sparse Eigenvector Method (QSE)

Finds algebraically smallest eigenvalue & eigenvector of H

Input: matrix H, basis for a subspace S

Do several times:

1. $H_S :=$ restriction of H to subspace S
   Compute smallest eigenvalue & eigenvector v of H_S
2. Discard basis vectors of S ‘corresponding to small’ components of v
3. Replace discarded basis vectors by new basis vectors
   New basis vectors connected to old ones through non-vanishing matrix elements of H
4. $S :=$ subspace of new and retained basis vectors

Efficiency of QSE

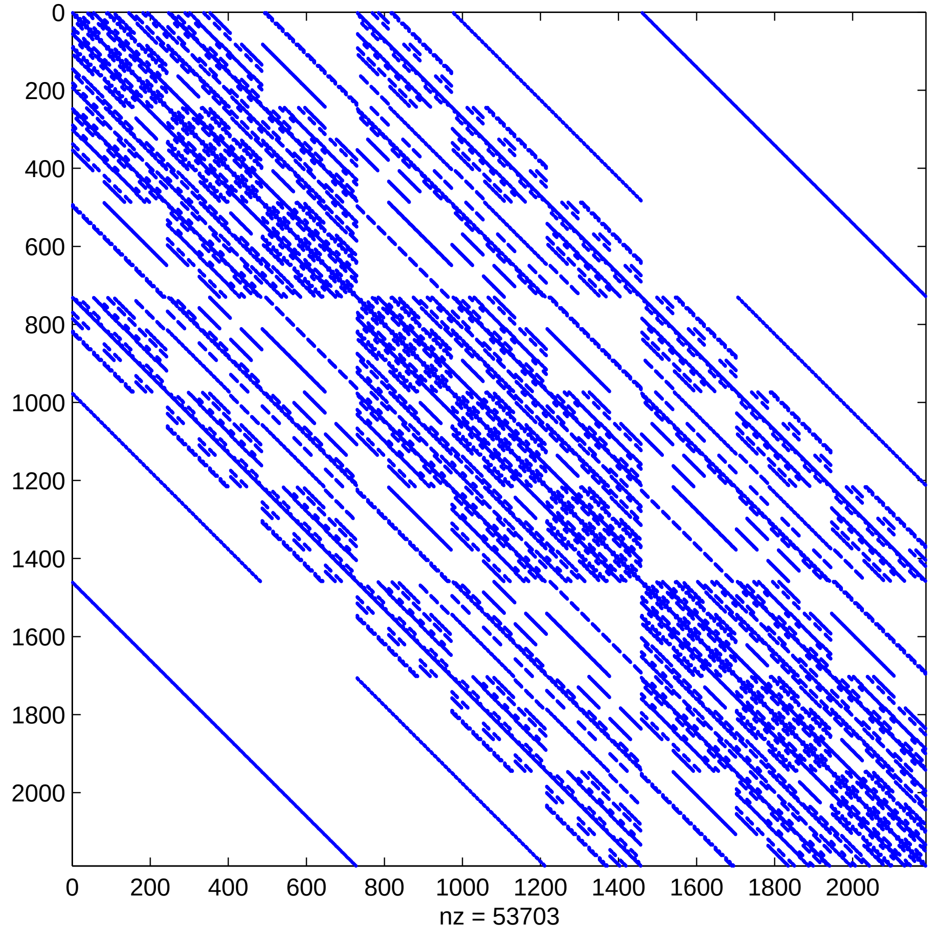
[Lee 2001], [Lee, Salwen, Lee 2001]

- If set of basis vectors sufficiently large then exact eigenvectors are stable fixed points of update process
- QSE efficient when eigenvector quasi-sparse, i.e. many eigenvector components small
- Quasi-sparsity related to eigenvalue separation

Why?

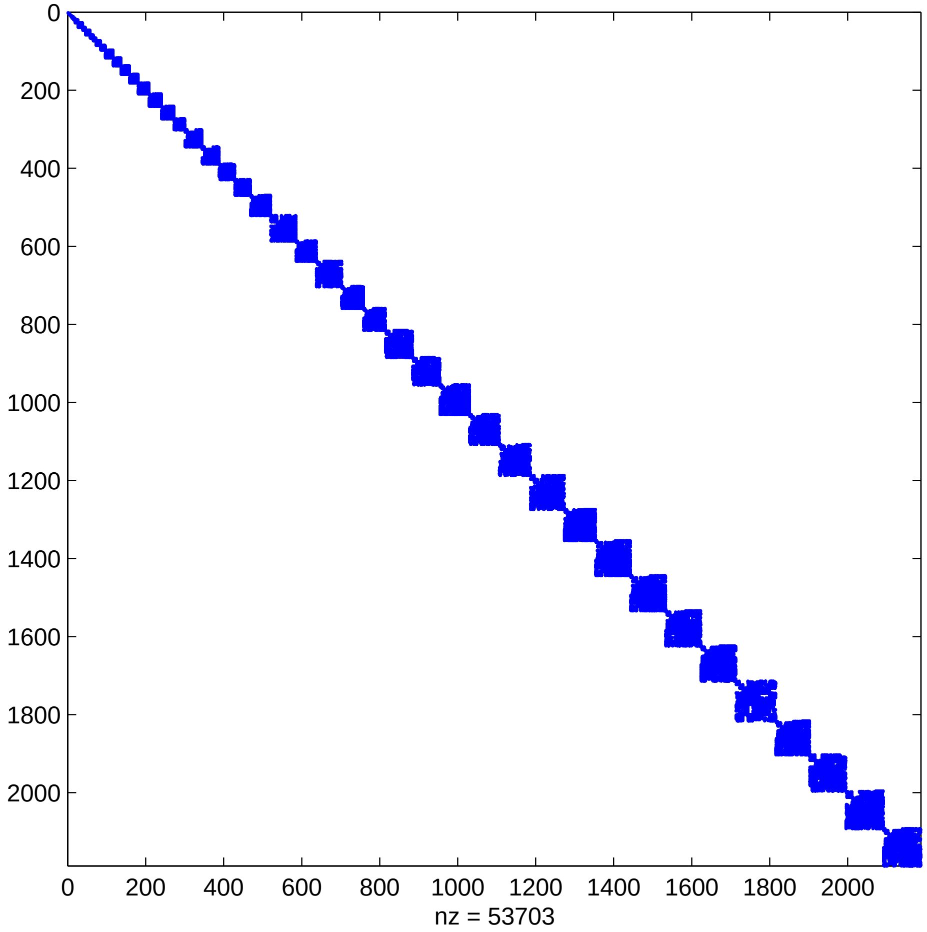
Look at structure of Hamiltonian

Structure of Hamiltonian

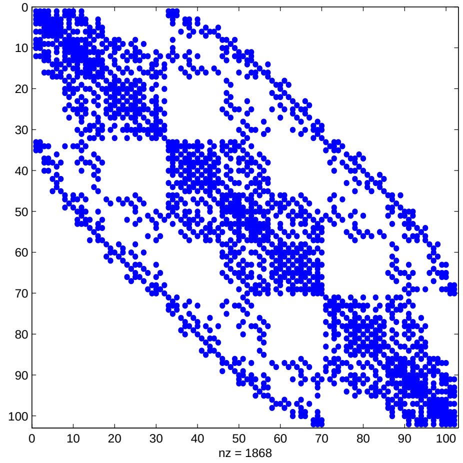


matrix order = 2187, n = 3, noccmax = 2, ell = 1, a2 = a4=1

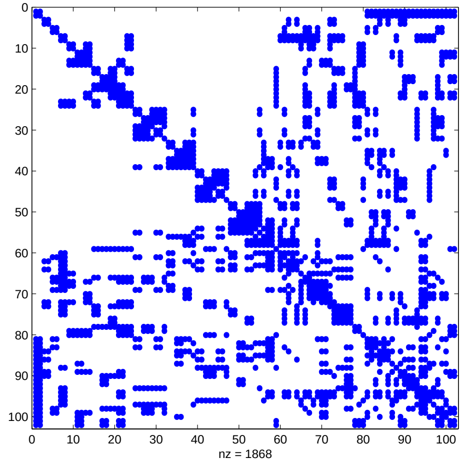
Symmetric Minimum Degree Reordering



Diagonal Block: 0 Momentum



Symmetric minimum degree reordering



matrix order = 102, n = 3, noccmax = 2, ell = 1, a2 = a4=1

Assumptions

- H is $n \times n$ Hermitian
- Compute smallest eigenvalue & eigenvector:
  $Hv = \lambda_{min} v$
- $\| \cdot \|$ is Euclidean norm
- $\|v\| = 1$
- Eigenvalues of H: $\lambda_n \leq \ldots \leq \lambda_1$

Ideal QSE

Use exact eigenvector components

$$Hv = \lambda v, \qquad H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

QSE: $\hat{v} \equiv \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \qquad \rho \equiv$ Rayleigh quotient of $\hat{v}$ and H

$$|\lambda - \rho| \leq 2\|H_{12}\| \, \|v_2\| + (|\rho| + \|H_{22}\|) \, \|v_2\|^2$$

$$\sin \angle(v, \hat{v}) = \|v_2\|$$

Ideal QSE accurate when neglected eigenvector components small (quasi-sparse)

When Are Eigenvectors Quasi-Sparse?

Eigenstates quasi-sparse if spacing between energy levels not too small compared to off-diagonal elements of Hamiltonian

$$Hv = \lambda v, \qquad H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

When is $\|v_2\|$ small?

$$\frac{\|H_{12}\|}{\max_i |\lambda_i(H_{22}) - \lambda|} \leq \|v_2\| \leq \frac{\|H_{12}\|}{\min_i |\lambda_i(H_{22}) - \lambda|}$$

One Iteration of (Non-Ideal) QSE

$$Hv = \lambda v, \qquad H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

Ritz Problem: $H_{11} w = \theta w, \qquad \hat{v} \equiv \begin{pmatrix} w \\ 0 \end{pmatrix}$

$$\min_i |\lambda_i - \theta| \leq \|H_{12}\|$$

$$\sin \angle(v, \hat{v}) \leq \frac{\|H_{12}\|}{\min_{\lambda_i \neq \lambda} |\lambda_i - \theta|}$$

Bounds small if $\|H_{12}\|$ small

Properties of Bounds

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

- Bounds independent of unknown block H_{22}
- H banded $\Rightarrow$ H_{12} 'finite' $\Rightarrow$ can estimate $\|H_{12}\|$
- Bounds small if off-diagonal block $\|H_{12}\|$ small

But: Offdiagonal Block Large

$$Hv = \lambda v, \qquad H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Example: matrix order = 520, n = 2, noccmax = 6, ell = 1, a2 = a4=1

$$\|H_{12}\| \approx 59, \qquad \|H_{11}\| \approx 385$$

but $\min_i |\lambda_i - \theta| \approx .02, \qquad \sin\angle(v, \hat{v}) \approx .04$

Ideal QSE residual:

$$\|(H - \lambda I)\begin{pmatrix} v_1 \\ 0 \end{pmatrix}\| \approx 1.7$$

$\|H_{12}\|$ too large, does not predict error

♣

Lehmann Bounds [Lehmann 1949], [Beattie, Ipsen 2002]

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}$$

Eigenvalues of H_{11}: $\theta_{\min} \leq \ldots \leq \theta_{\max}$

At least 1 eigenvalue of H in

$$\left\{ z : \ |z - \theta_{\min}| \leq \sigma_{\min} \begin{pmatrix} H_{11} - \theta_{\min} I \\ H_{12}^* \end{pmatrix} \right\}$$

If H is positive-definite then 1 eigenvalue of H in

$$\left\{ z : \ \left| \frac{z - \theta_{\min}}{\theta_{\min}} \right| \leq \sqrt{2} \max \left\{ \|H_{11}^{-1} H_{12}\|, \ 1 - \frac{\theta_{\min}}{\theta_{\max}} \right\} \right\}$$

Quality of Lehmann Bounds

absolute bound: Lehm = $\sigma_{\min} \begin{pmatrix} H_{11} - \theta_m I \\ H_{12}^* \end{pmatrix}$

(part of) relative bound: $\|H_{11}^{-1} H_{12}\|$

n	m	abs err	Lehm	rel err	$\|H_{11}^{-1} H_{12}\|$
102	7	.03	.9	.01	.3
28	5	.01	.8	.009	1.4
520	17	.01	.5	.008	.9
610	21	.01	.6	.004	.6

$\|H_{11}^{-1} H_{12}\|$ too large, does not predict relative error

Better Relative Bounds

$$Hv = \lambda v, \qquad H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Eigenvalue:

$$\min_i \left| \frac{\lambda - \theta_i}{\theta_i} \right| \leq \|H_{11}^{-1} H_{12}\| \, \frac{\|v_2\|}{\|v_1\|}, \qquad \theta_i \equiv \lambda_i(H_{11})$$

Eigenvector:

$$H_{11} w = \theta w, \qquad \text{relgap} \equiv \min_{\theta_i \neq \theta} \left| \frac{\lambda - \theta_i}{\theta_i} \right|$$

$$\sin \angle(v_1, w) \leq \frac{\|H_{11}^{-1} H_{12}\|}{\text{relgap}} \, \frac{\|v_2\|}{\|v_1\|}$$

Positive-Definite Matrices

Eigenvalues of H: $0 < \lambda_n < \lambda_{n-1} \leq \ldots$
Eigenvalues of H_{11}: $0 < \theta_m < \theta_{m-1} \leq \ldots$

Smallest eigenvalue: $Hv = \lambda_n v, \quad H_{11} w = \theta_m w$

$$\frac{\theta_m - \lambda_n}{\theta_m} \leq \|H_{11}^{-1} H_{12}\| \frac{\|v_2\|}{\|v_1\|}$$

$$\sin \angle(v_1, w) \leq \frac{\|H_{11}^{-1} H_{12}\|}{\text{relgap}} \frac{\|v_2\|}{\|v_1\|}$$

where

$$\text{relgap} \equiv \frac{\lambda_{n-1} - \lambda_n}{\lambda_{n-1}}$$

Positive-Definite Matrices

Eigenvalues of H: $0 < \lambda_n < \lambda_{n-1} \leq \dots$
Eigenvalues of H_{11}: $0 < \theta_m < \theta_{m-1} \leq \dots$

Next smallest eigenvalue: If also

$$\frac{\lambda_{n-1} - \lambda_n}{\lambda_n} > 1, \qquad \|H_{11}^{-1} H_{12}\| \, \frac{\|v_2\|}{\|v_1\|} \leq \frac{1}{2} \, \frac{\lambda_{n-1}}{\lambda_{m-1}}$$

then

$$\frac{\theta_{m-1} - \lambda_{n-1}}{\theta_{m-1}} \leq \|H_{11}^{-1} H_{12}\| \, \frac{\|v_2\|}{\|v_1\|}$$

θ_{m-1} has same error bound as θ_m

Quality of Eigenvalue Bound

$$\frac{\theta_m - \lambda_n}{\theta_m} \leq \|H_{11}^{-1} H_{12}\| \, \gamma \qquad \gamma \equiv \frac{\|v_2\|}{\|v_1\|}$$

n	m	rel err	bound	γ	$\|H_{11}^{-1} H_{12}\|$
102	7	.011	.013	.04	.3
28	5	.009	.02	.01	1.3
520	17	.008	.023	.026	.9
610	21	.004	.01	.02	.6

$\|v_2\|/\|v_1\|$ predicts relative eigenvalue error better than $\|H_{11}^{-1} H_{12}\|$

Quality of Eigenvector Bound

$$\sin \angle(v_1, w) \le \frac{\|H_{11}^{-1} H_{12}\|}{\text{relgap}} \gamma \qquad \gamma \equiv \frac{\|v_2\|}{\|v_1\|}$$

other $\equiv \|H_{11}^{-1} H_{12}\| / \text{relgap}$

n	m	$\sin$	bound	γ	other
102	7	.001	.01	.04	.4
28	5	.0142	.0196	.01	1.4
520	17	.002	.027	.026	1.04
610	21	.0005	.01	.02	.8

$\|v_2\| / \|v_1\|$ predicts relative eigenvector error better than $\|H_{11}^{-1} H_{12}\| / \text{relgap}$

♣

How to Make the Bounds Small

$$Hv = \lambda_n v, \qquad H_{11}w = \theta_m w, \qquad \lambda_n > 0$$

$$\frac{\theta_m - \lambda_n}{\theta_m} \leq \|H_{11}^{-1}H_{12}\| \, \frac{\|v_2\|}{\|v_1\|}$$

$$\sin\angle(v_1, w) \leq \frac{\|H_{11}^{-1}H_{12}\|}{\text{relgap}} \, \frac{\|v_2\|}{\|v_1\|}$$

Symmetrically permute H so that

- $\|H_{11}^{-1}H_{12}\|$ small
- $\|v_2\|/\|v_1\|$ small

Make $\|H_{11}^{-1}H_{12}\|$ Small

There exists permutation P such that

$$PHP^* = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \qquad \left|\left(H_{11}^{-1}H_{12}\right)_{ij}\right| \leq 1$$

$$\|H_{11}^{-1}H_{12}\| \leq \sqrt{m(n-m)}$$

Make $\|H_{11}^{-1}H_{12}\|$ Small

There exists permutation P such that

$$PHP^* = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \qquad \left| \left(H_{11}^{-1}H_{12}\right)_{ij} \right| \leq 1$$

$$\|H_{11}^{-1}H_{12}\| \leq \sqrt{m(n-m)}$$

But: [Eisenstat, Gu 1996]

$$\frac{\lambda_m}{m(n-m)} \leq \theta_m \leq \lambda_m \qquad \text{wrong eigenvalue}$$

Make $\|H_{11}^{-1}H_{12}\|$ Small

There exists permutation P such that

$$PHP^* = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \qquad \left|\left(H_{11}^{-1}H_{12}\right)_{ij}\right| \leq 1$$

$$\|H_{11}^{-1}H_{12}\| \leq \sqrt{m(n-m)}$$

But: [Eisenstat, Gu 1996]

$$\frac{\lambda_m}{m(n-m)} \leq \theta_m \leq \lambda_m \qquad \text{wrong eigenvalue}$$

But:

$$\frac{\theta_m - \lambda_m}{\lambda_m} \leq m(n-m)$$

Make $\|v_2\|/\|v_1\|$ Small

Order components of v in descending order

perm?	n	m	eval	bound	evec	bound
N	102	7	.011	.013	.001	.02
Y			.003	.01	.005	.01
N	520	17	.008	.02	.0018	.03
Y			.0009	.008	.0020	.009
N	610	21	.004	.01	.0006	.02
Y			.0002	.003	.002	.004

Bounds often decrease but errors may not

Observations

- Trying to make bounds small may not improve eigenvalue and eigenvector accuracy
- Selecting matrix pieces associated with large eigenvector components may not improve accuracy
- QSE: No theoretical justification yet for changing Ritz matrices

Question: How to choose H_{11} in each QSE iteration?

Simple Answer: Don't change H_{11}

Choose H_{11} = leading principal submatrix

Accuracy Requirements

$$\text{sep} \equiv \left| \frac{\lambda_{n-1} - \lambda_n}{\lambda_n} \right|, \qquad \text{ritz\_sep} \equiv \left| \frac{\theta_{m-1} - \theta_m}{\theta_m} \right|$$

Want: $\left| \frac{\text{sep} - \text{ritz\_sep}}{\text{ritz\_sep}} \right| < .05$

Have: $\lambda_n = \theta_m(1 - \epsilon_n), \quad \lambda_{n-1} = \theta_{m-1}(1 - \epsilon_{n-1})$

$$|\epsilon_i| \leq \epsilon, \qquad \epsilon \equiv \|H_{11}^{-1} H_{12}\| \frac{\|v_2\|}{\|v_1\|} \leq \frac{1}{2} \frac{\lambda_{n-1}}{\lambda_{m-1}}$$

Then:

$$\left| \frac{\text{sep} - \text{ritz\_sep}}{\text{ritz\_sep}} \right| \leq 2 \frac{\lambda_{n-1}}{\lambda_n} \epsilon + \mathcal{O}(\epsilon^2)$$

Really Simple QSE

Rayleigh Ritz: $m = 10$

$$\text{eval} \equiv \frac{\theta_m - \lambda_n}{\theta_m}, \quad \text{evec} \equiv \sin\angle(v_1, w), \quad \text{sep} \equiv \frac{\lambda_{n-1} - \lambda_n}{\lambda_n}$$

sep error $\equiv$ (sep $-$ ritz_sep)/ritz_sep

n	eval	evec	sep	sep error
102	.003	.0007	2.8	.0009
28	.0002	.0008	7.9	.006
520	.009	.004	4.8	.004
610	.005	.0008	2.6	.001

Summary

- Compute smallest eigenpair of Hermitian Hamiltonian
- Eigenvectors are quasi-sparse
- Quasi-sparsity $\sim$ large eigenvalue separation
- Analysis of quasi-sparse eigenvector method (QSE)
- Perturbation bounds for eigenvalues & eigenvectors
  - absolute (too pessimistic)
  - relative (mostly for positive-definite matrices)
- Simple QSE: 10×10 Ritz problem
- Ritzvalues & vectors computed to relative accuracy
- Ritz value separation computed to desired relative accuracy