Recent Progess in Linear Scaling Quantum Monte Carlo Algorithms

* Introduction: Scaling in Classical and Quantum Monte Carlo
* Linear Scaling Algorithm for Quantum Simulations
* Progress in Solving the Conditioning Problem
* Conclusions/Future Directions

Z. Bai (UCD, Computer Science and Applied Math)

I. Yamazaki (UCD, Computer Science)

R.T. Scalettar (UCD, Physics)

NSF ITR 0313390

ITR Significance

* Example of interplay between computer science/applied mathematics and an important materials theory problem
* Addressing linear algebra problem at center of Quantum Monte Carlo for lattice fermion models

Materials Science Objective

Applications to Correlated Electron Models:

* Hubbard Hamiltonian
* Periodic Anderson Hamiltonian
* Impurity solvers in Dynamical Mean Field Theory

Applications to Physics of Materials with Strongly Correlated Bands:

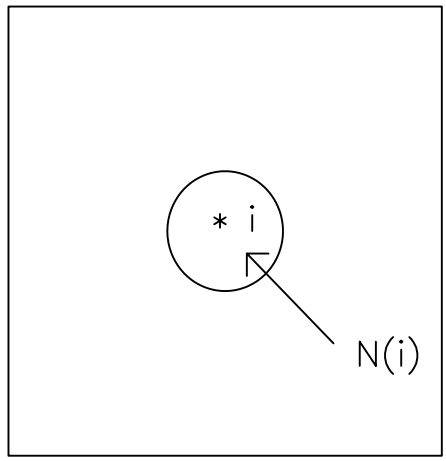
* Magnetism
* Metal-Insulator Transitions

Classical Monte Carlo

Energy of set of degrees of freedom x_i is local:

$$E = \sum_i^N x_i \sum_{j \in \mathcal{N}(i)} \kappa_{ij}\, x_j$$

x_i couples only to x_j only in some neighborhood $\mathcal{N}(i)$.



Suggest change $x_i \to x_i'$. Need to evaluate

$$e^{-\beta E'}/e^{-\beta E} = e^{-\beta \Delta E} \qquad\qquad \beta = 1/T$$

Since $\mathcal{N}(i)$. is independent of system size, so is time to update x_i.

CONCLUSION:

Time to do update all degrees of freedom is linear in the lattice size.

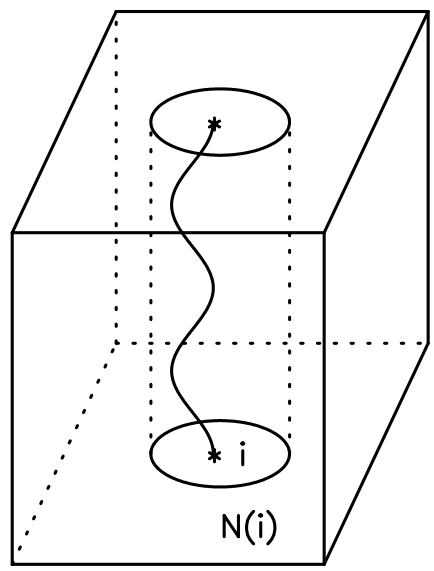
Quantum Monte Carlo

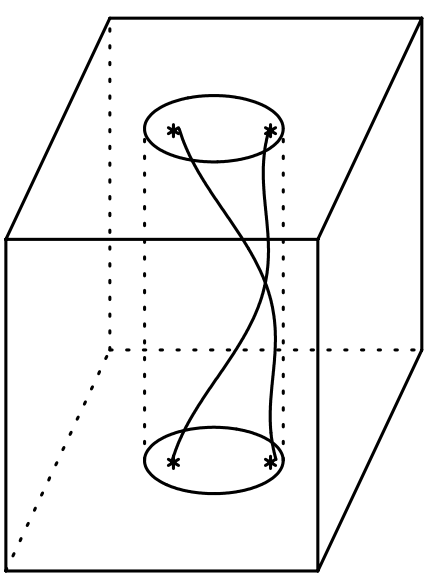
Classical Boltzmann weight, an exponential of a number, E, becomes the exponential of an operator, H:

$$e^{-\beta E} \rightarrow e^{-\beta \hat{H}}$$

Express this as a path integral.
Extra "imaginary time" dimension of extent $\beta = 1/T$.

* If $\hat{H}$ is local, time is still linear in spatial lattice size.
* Cost is only extra dimension of lattice, a factor $\propto 1/T$.





This works for quantum spins and bosons (unfrustrated lattices). But, "sign problem" for fermions. If electron world lines exchange, the contribution to partition function is negative.

Solution: Eliminate fermion operators with a "Hubbard-Stratonovich" transformation. This introduces classical variables $x(i,\tau)$ living on space-time lattice. Replace

$$e^{-\beta E} \rightarrow \det \mathcal{M}(x)$$

$\mathcal{M}$ is a matrix of dimension NL where N is the spatial lattice size and L is proportional to the inverse temperature $\beta = 1/T$.

To do the simulation one needs to compute the ratio of the determinant of $\mathcal{M}$ before and after a monte carlo move $x \rightarrow x'$.

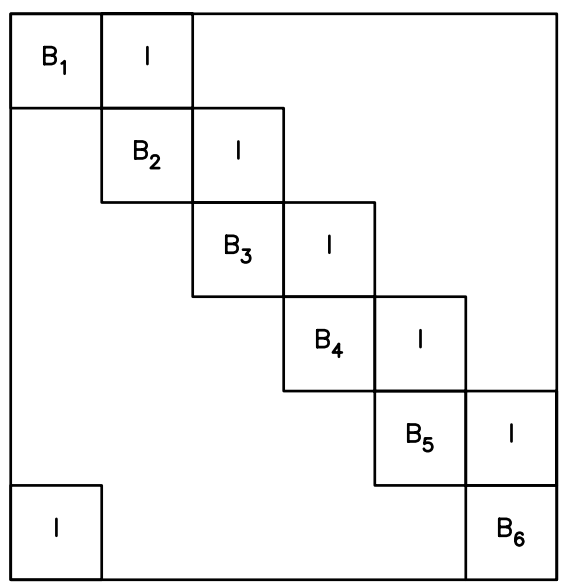
$$\det\mathcal{M}(x')/\det\mathcal{M}(x)$$

* CPU time $\propto (NL)^3$ to update one $x(i,\tau)$.
* Time to sweep entire lattice $\propto (NL)^4$.
* Take advantage of special structure of $\mathcal{M}$ and locality of $\hat{H}$ to reduce scaling to $N^3\,L$.

B_l = sparse dimension N matrices.

I = dimension N identity matrix.

$\mathcal{M}=$



Present Simulations

N= 10x10=100 (2-d lattice)

L =100 (to reach low temperature)

Sweeps = 10^4 (monte carlo averaging)

Time $\propto$	10^2	100^3	100	10^4	/	10^9	$\approx 10^5$ seconds
	prefactor	N^3	L	sweeps	/	clock speed	one day

Goal

3-d lattices (or larger 2-d lattices)

N=10x10x10 = 10^3

Time $\approx 10^8$ seconds

Linear Scaling Algorithm

Multidimensional Gaussian Integral

$$\int d\Phi \, e^{-\Phi^T \mathcal{A} \Phi} \propto (\det \mathcal{A})^{-\frac{1}{2}}$$

A : a real symmetric matrix

Φ : a vector

Recasting,

$$\int d\Phi \, e^{-\Phi^T (\mathcal{M}\mathcal{M}^T)^{-1} \Phi} \propto \det \mathcal{M}$$

Quantum Monte Carlo

$$\int dx \int d\Phi e^{-\Phi^T (\mathcal{M}\mathcal{M}^T)^{-1} \Phi} \propto \det \mathcal{M}$$

Integrate over both

* $x(i,\tau)$: Classical Hubbard-Stratonovich variables introduced to eliminate fermion interactions.
* $\Phi(i,\tau)$: Variables which replace determinant by Gaussian integral.
* Entries in $\mathcal{M}$ depend on $x(i,\tau)$

Linear Scaling

Updating Φ in order NL is trivial:

$$\Phi = \mathcal{M} R$$

R : a vector of independent Gaussian random numbers

Recall $\mathcal{M}$ is sparse.

Updating $x \to x + \delta x$ requires computation of

$$\delta \left(\Phi^T \left(\mathcal{M}\mathcal{M}^T\right)^{-1} \Phi \right) = \Phi^T \left(\mathcal{M}\mathcal{M}^T\right)^{-1} \delta \left(\mathcal{M}\mathcal{M}^T\right)^{-1} \left(\mathcal{M}\mathcal{M}^T\right)^{-1} \Phi$$

Need Iterative Solver for $(\mathcal{M}\mathcal{M}^T)^{-1}\Phi$

$\mathcal{M}$ and $\mathcal{M}^T$ are sparse: multiplication on a vector is NL.

But they are poorly conditioned!

* Conjugate Gradient iteration number grows rapidly with N and L

Conditioning

Use (analytically known) matrix inverse in certain physical limits:

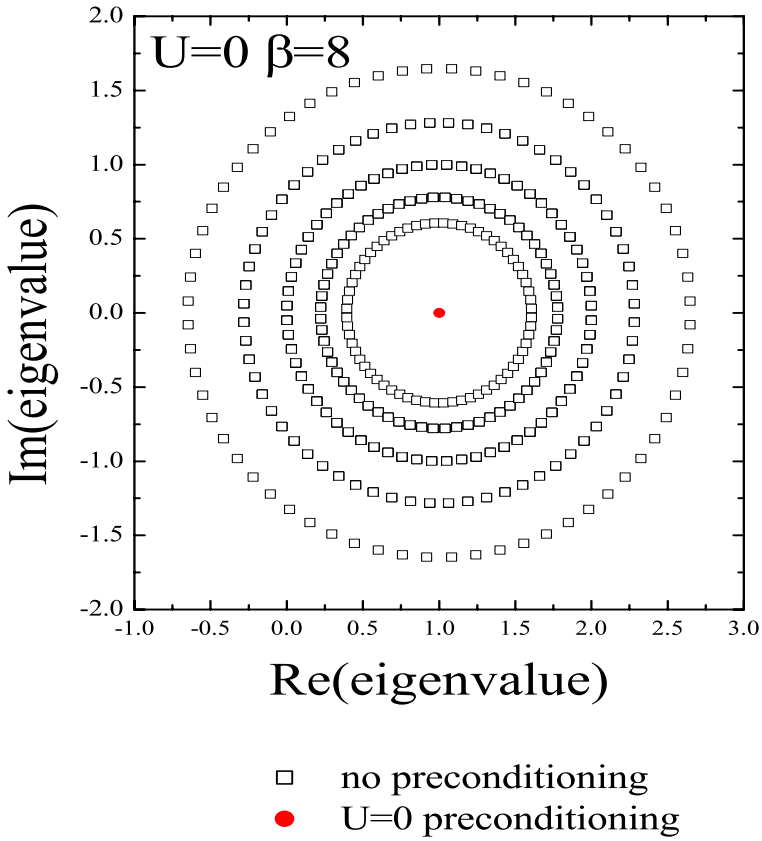
$U = 0$ electron-electron interactions are zero

$t = 0$ electron kinetic energy is zero

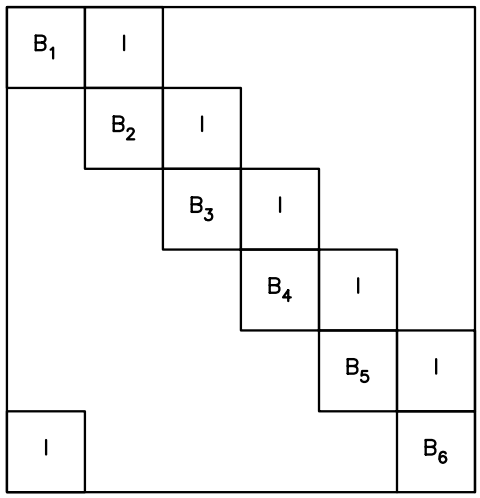
Hubbard Hamiltonian

$$H = -t \sum_{\langle ij \rangle \sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) + U \sum_i n_{i\uparrow} n_{i\downarrow}$$

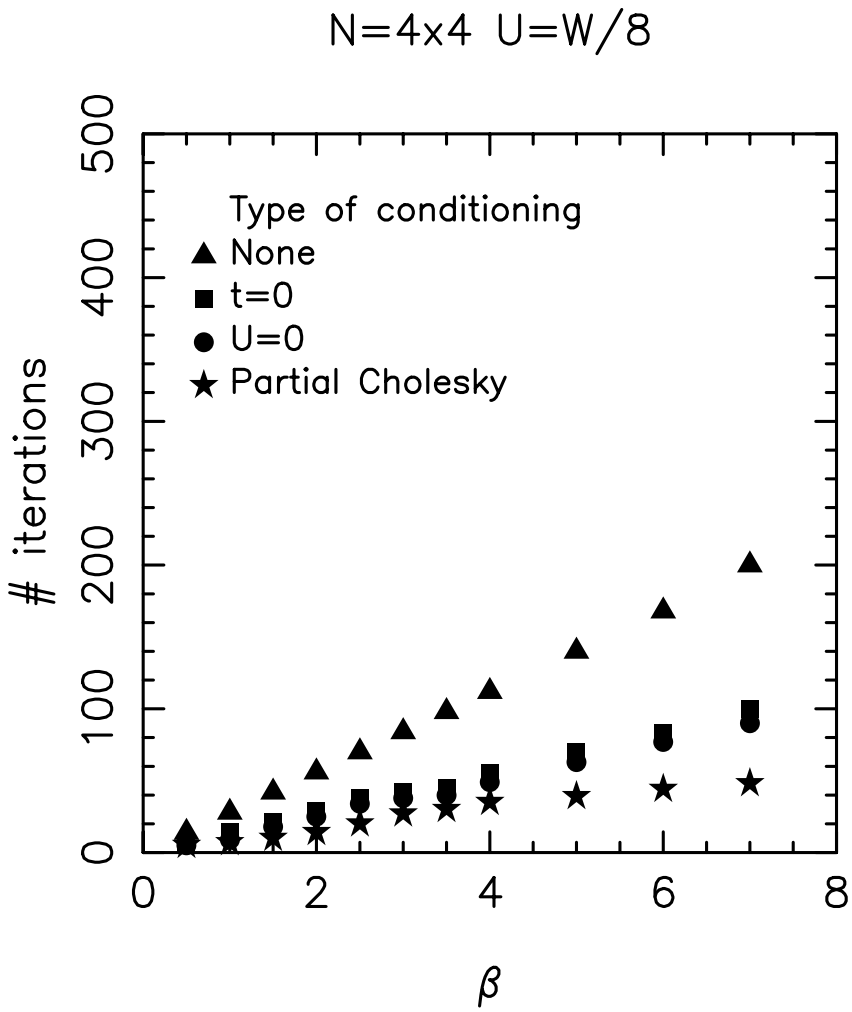
Eigenspectrum on Noninteracting Electrons



Special pattern of eigenvalues comes from structure of matrix.



Effect of Conditioning



This should have been an easy case:

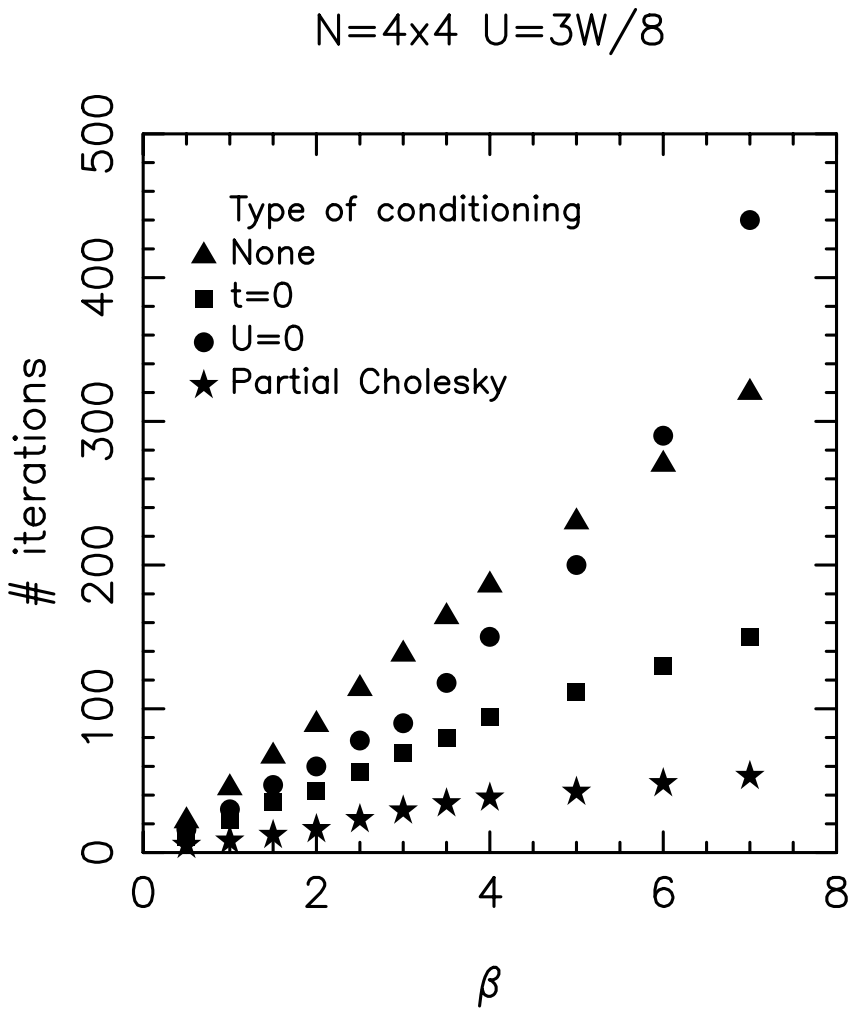
Small lattice (16 sites)

Weak coupling

But iteration number grows linearly with system size NL.

Incomplete Cholesky is okay?

Effect of Conditioning

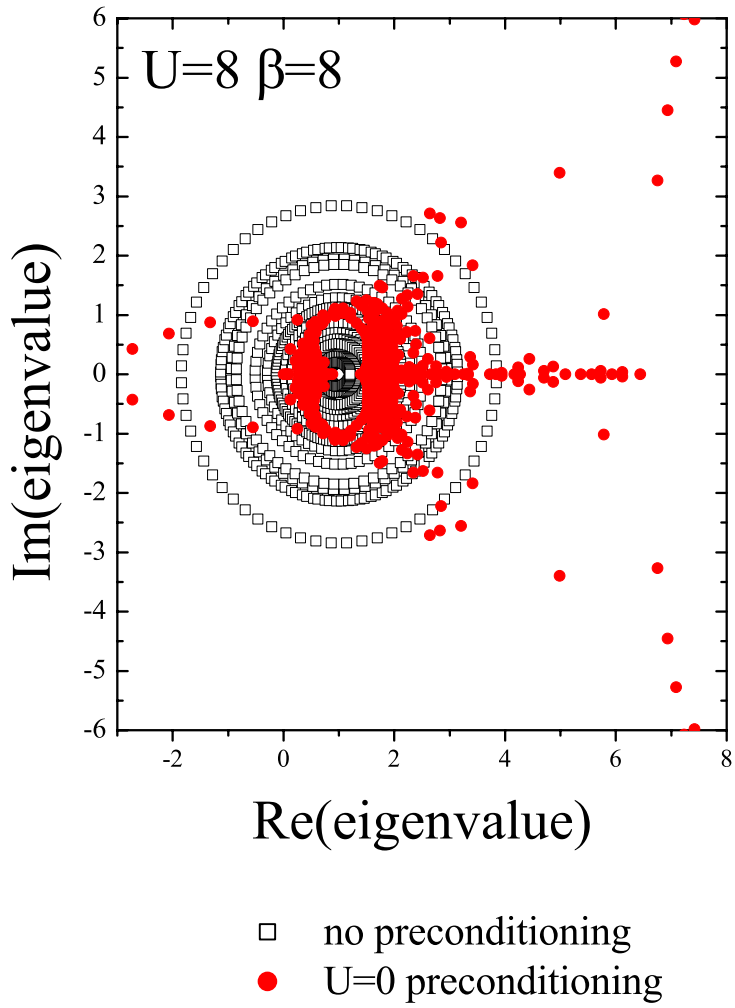


Somewhat more realistic case:

Still Small lattice (16 sites)

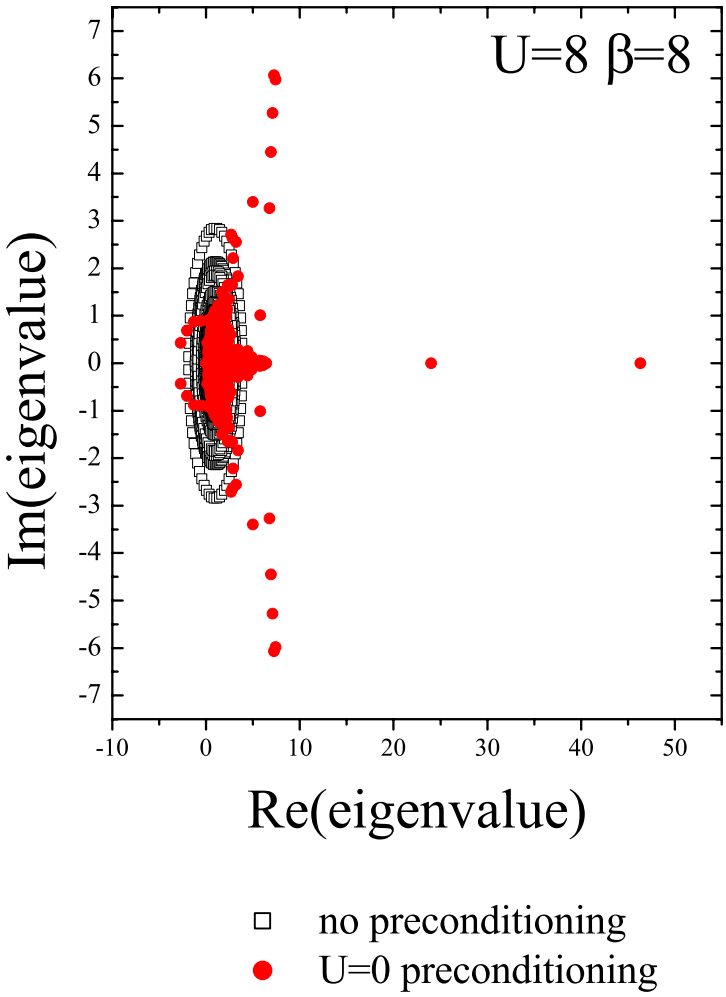
Moderate coupling (strong enough to see magnetic ordering)

$U = 0$ Conditioning



Eigenvalue spectrum worsened

$U = 0$ Conditioning



Entire spectrum

Cholesky Decomposition

$$\mathcal{M}\mathcal{M}^T = \underset{\substack{\text{lower}\\\text{triangular}}}{L} \quad \underset{\text{diagonal}}{D} \quad \underset{\substack{\text{upper}\\\text{triangular}}}{U}$$

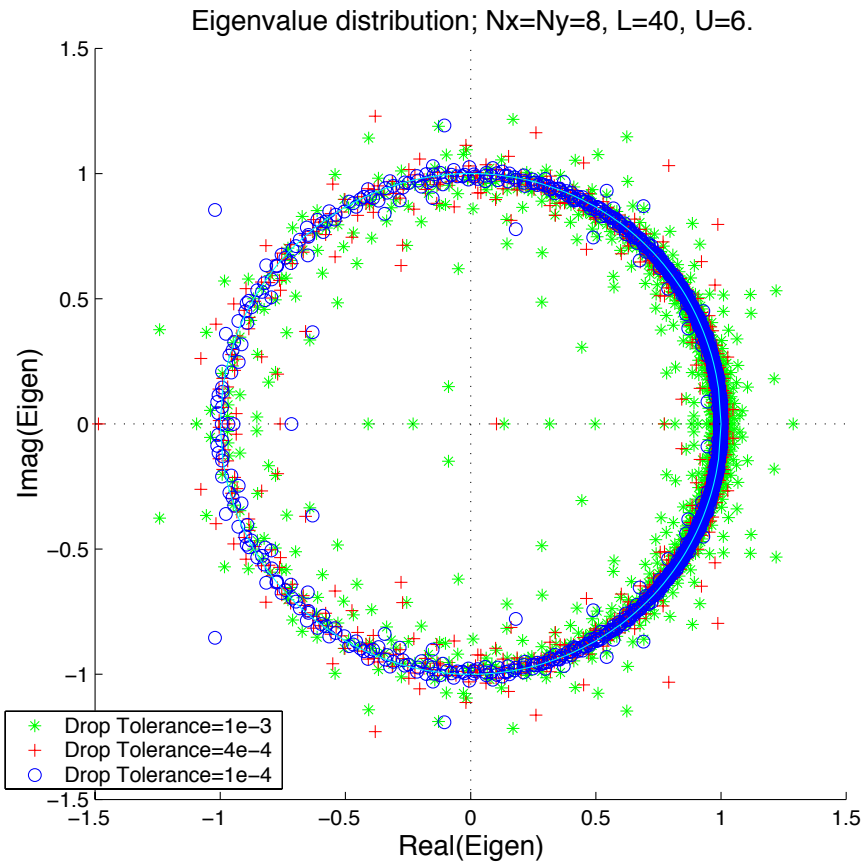
$U^{-1}D^{-1}L^{-1}$ is exact inverse of $\mathcal{M}\mathcal{M}^T$.

* conjugate gradient converges in a single iteration
* but costs $(NL)^3$ to compute
* L and U do not preserve sparsity pattern of $\mathcal{M}\mathcal{M}^T$.

Incomplete Cholesky Decomposition

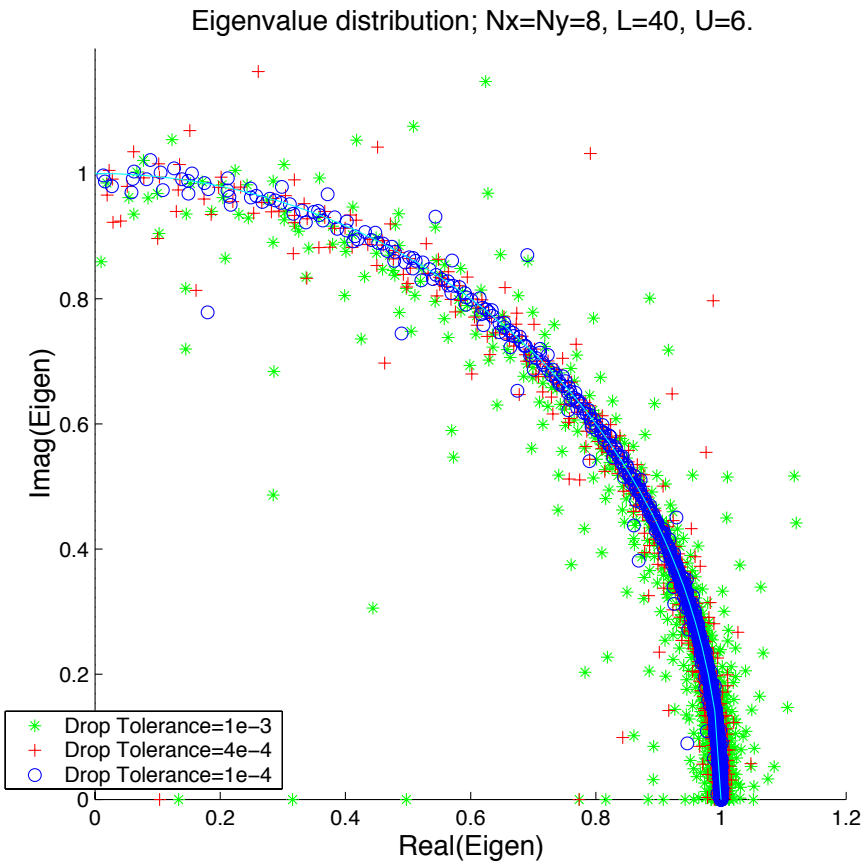
* Insist L and U have sparsity pattern of $\mathcal{M}\mathcal{M}^T$.
* $(NL)^1$
* Good, but not good enough.
* Same sort of blow-up of iteration number as with $t = 0$ and $U = 0$ conditioning at physically interesting lattice sizes, temperature, interaction strengths.

Improving Incomplete Cholesky



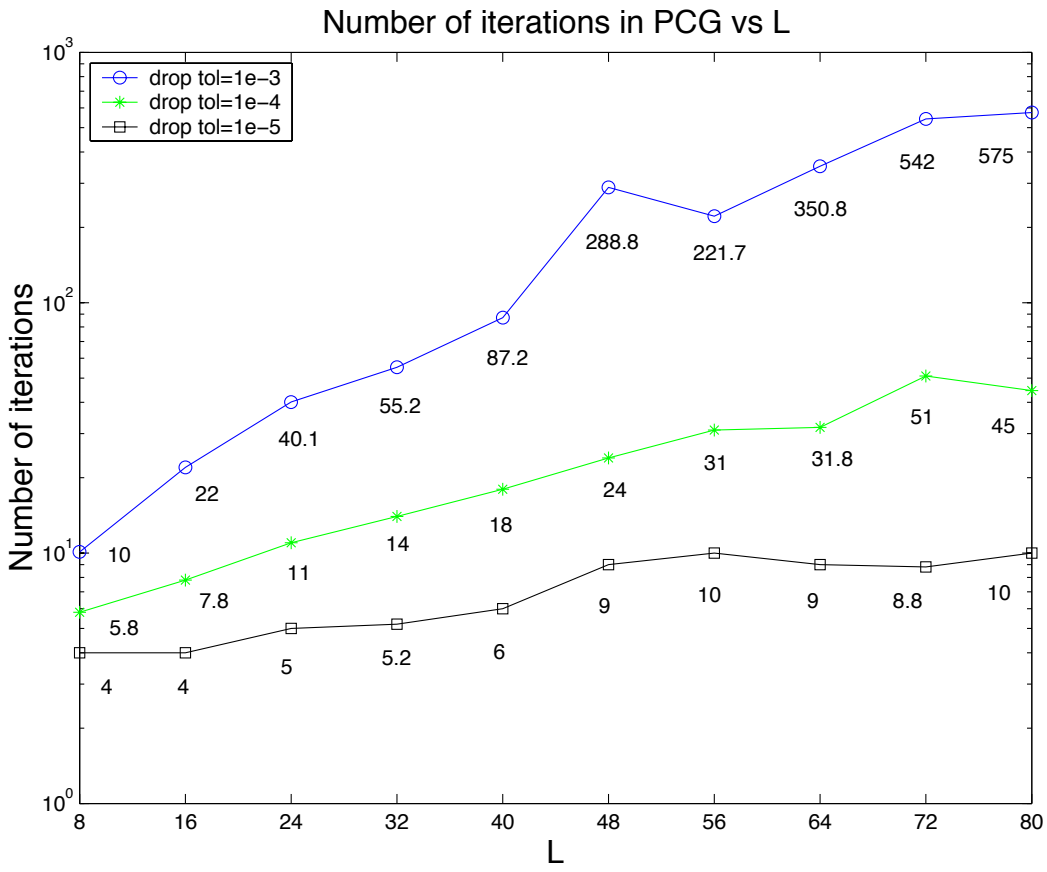
* Bai and Yamazaki introduce "drop tolerance", a control parameter to allow a somewhat larger number of fill-ins of triangular matrices than in original $\mathcal{M}\mathcal{M}^T$.
* Original Incomplete Cholesky: drop tolerance $\approx 10^{-2} - 10^{-3}$.
* Observe eigenvalues of conditioned matrix collapse better and better on the unit circle $|\lambda| = 1$ in the complex plane.

Improving Incomplete Cholesky



* Detail of first quadrant of previous distribution.

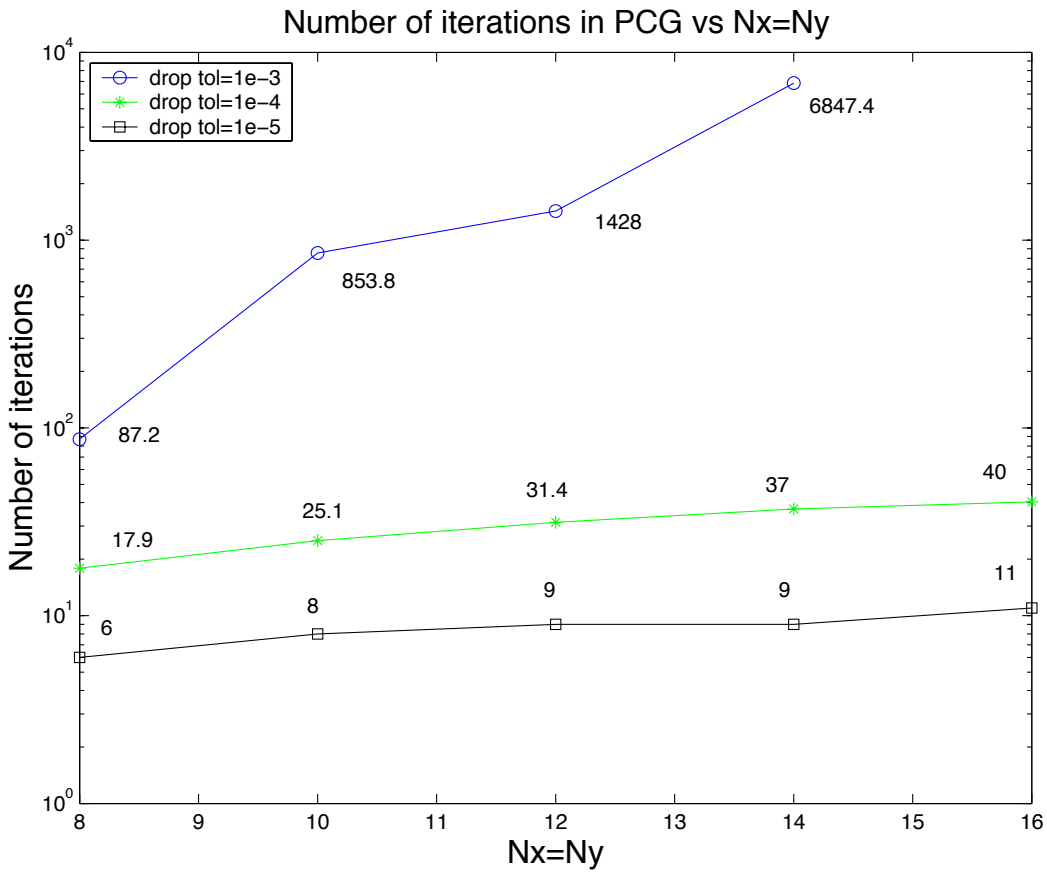
Scaling with Inverse Temperature



* Number of iterations is only weakly growing with L (roughly $L^{4/3}$)

* Entire algorithm will scale linearly with inverse temperature.

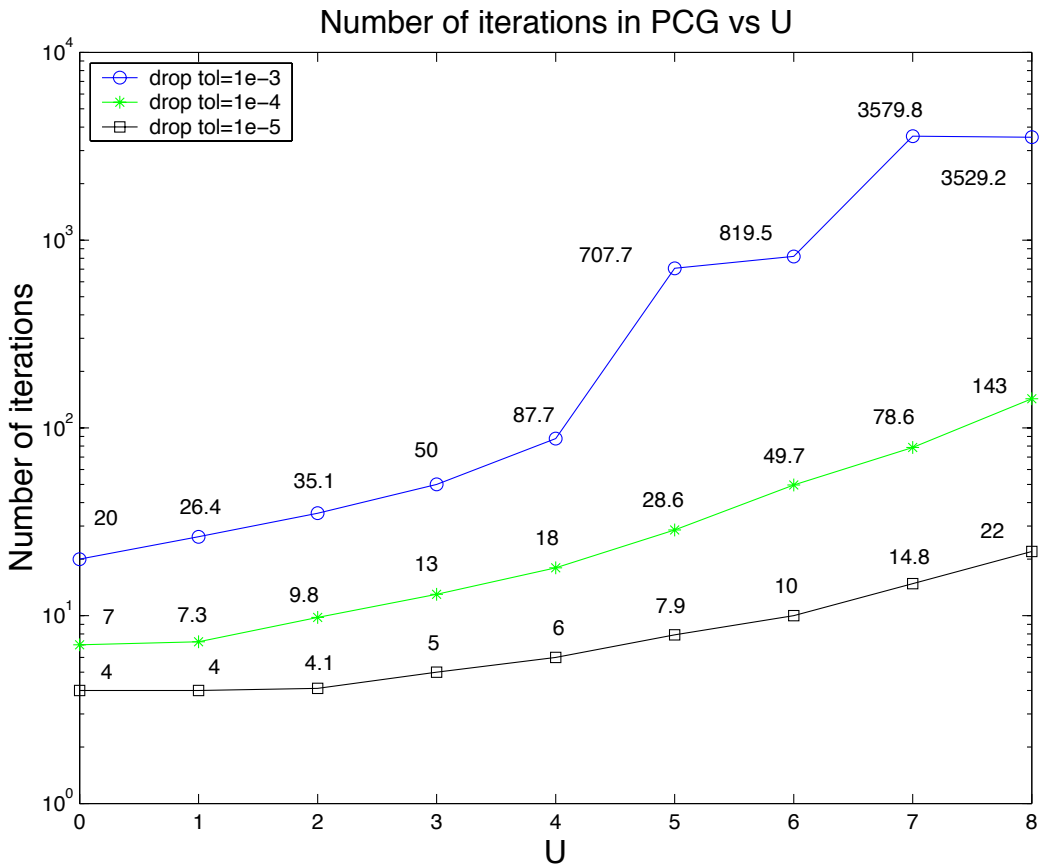
Scaling with Spatial Size



* Number of iterations is only weakly growing with N

* Entire algorithm will scale linearly with spatial size

Scaling with Interaction Strength



* Number of iterations is grows with U (as expected)

* Simulations about four times as expensive at $U = 8$ than at $U = 4$.

Conclusions

* Have constructed a robust conditioner for matrices arising in determinant Quantum Monte Carlo simulations.
* ITR crucial: Fostered computer science/physics interaction.

Future

* Implement in Full QMC code
* Investigate parallelization
* Look at quadratic form instead of full vector
* Address analogous issues in "dynamical mean field theory" simulations with many orbitals.