Prospects for Quantum Monte Carlo Methods for calculating defects in materials

- Introduction to QMC methods
- Some results for defects
- Advantages of QMC for defect studies

Jaron Krogel and DMC Illinois

Jeongnim Kim Oak Ridge

Observation of Superflow in Solid Helium

E. Kim and M. H. W. Chan*

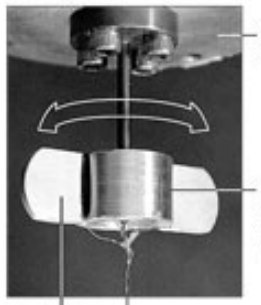
Science, September 2004

NewYorkTimes

Some Surprising Moves

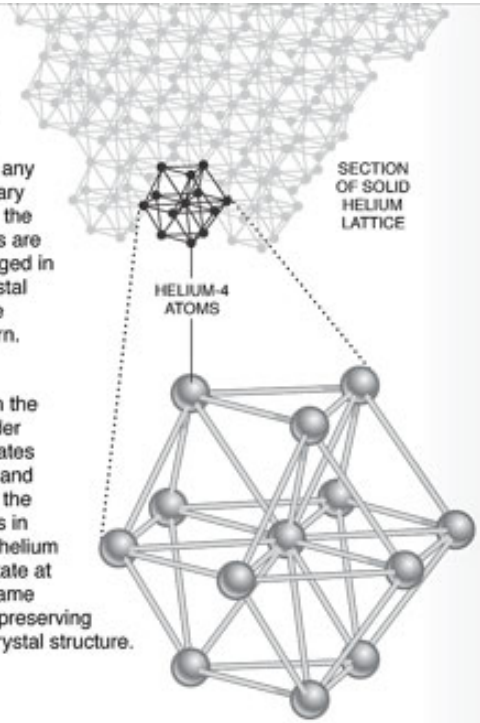
Helium may turn into a new state of matter called a supersolid — when rotated, it does not quite act solid.

1. Helium gas inside a cylinder (below, actual size) is chilled and squeezed until it turns into a solid. The cylinder then oscillates.



2. As in any ordinary solid, the atoms are arranged in a crystal lattice pattern.

When the cylinder oscillates back and forth, the atoms in solid helium all rotate at the same rate, preserving the crystal structure.

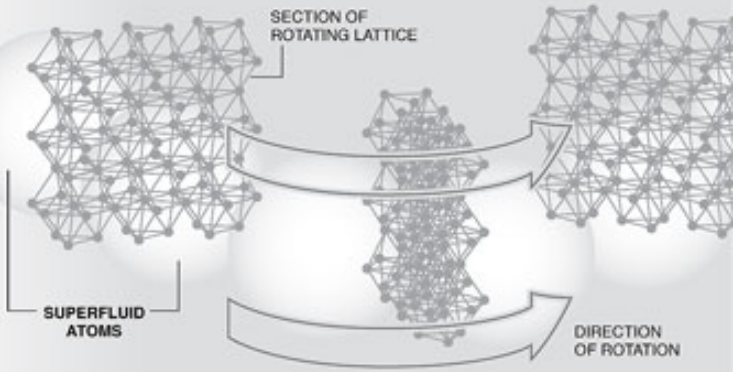


Sources: Dr. Moses H. W. Chan and Dr. David S. Weiss, Penn State University; Dr. Wayne Saslow, Texas A&M University

A NEW STATE OF MATTER, BOTH SOLID AND SUPERFLUID

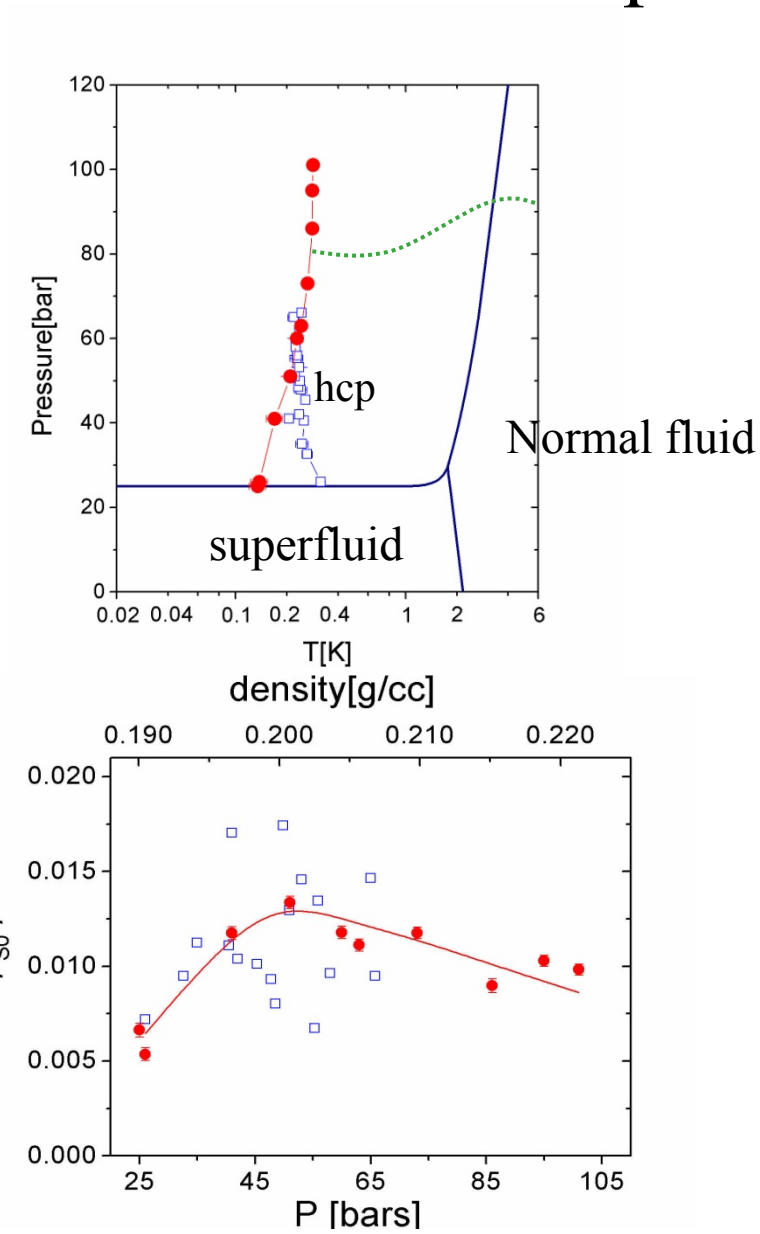
3. To turn solid into a supersolid, the temperature is lowered further, to almost absolute zero.

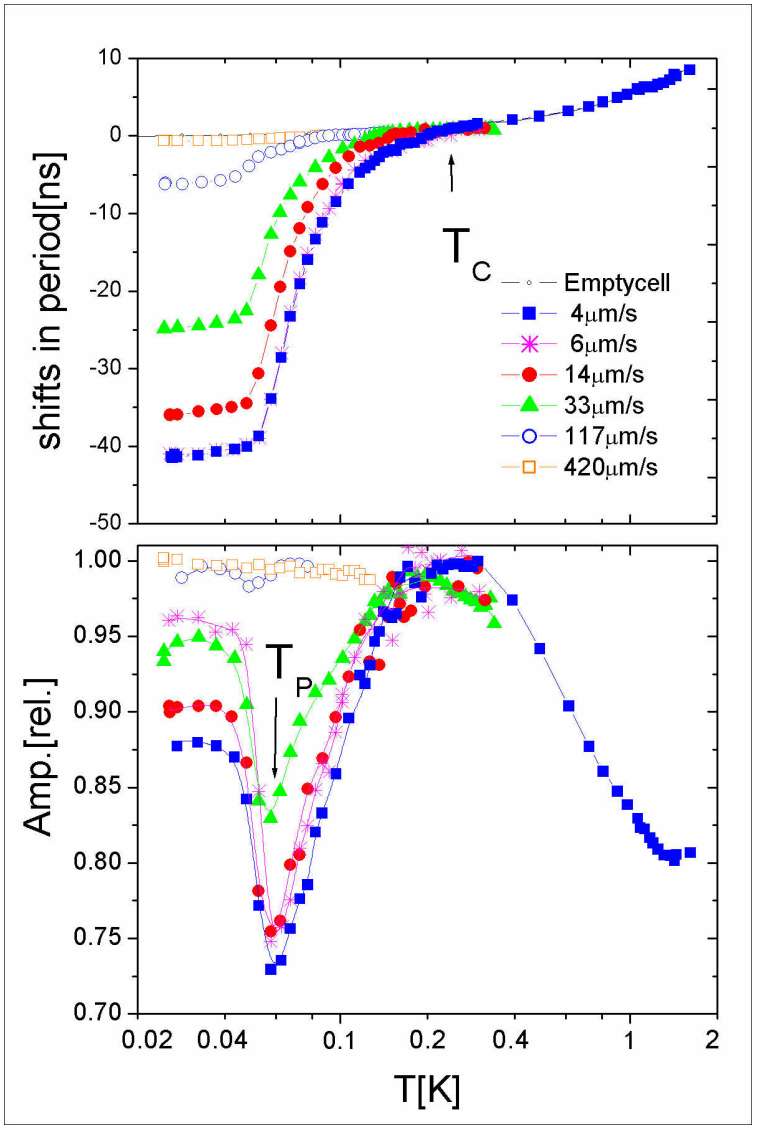
4. A small portion (1.5 percent) of the atoms are free to move away from their lattice sites, becoming superfluid. Their positions are blurred; they coalesce and suffuse the entire solid.



During rotation, as the lattice flows by, the blurry superfluid atoms stay in place. There is no friction between moving and stationary atoms.

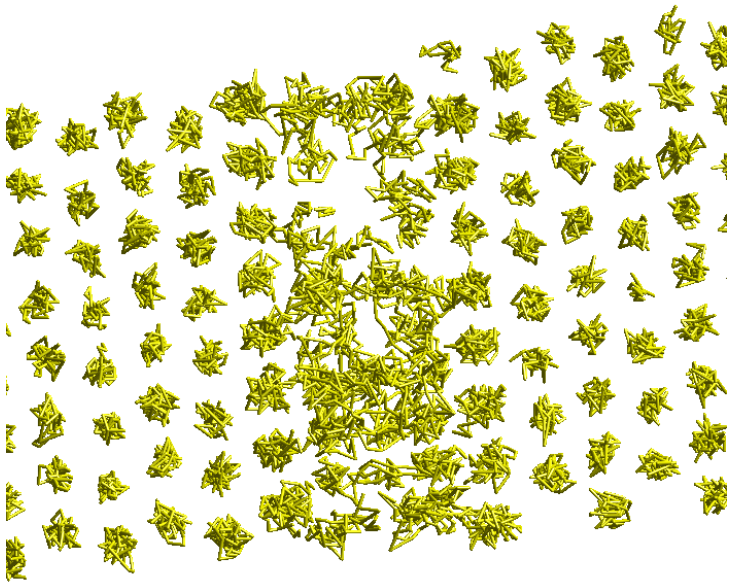
Kim-Chan experiments on bulk ^{4}He





Defect model of supersolidity

- Could the effect in bulk helium be due to surface superfluidity?
- With 2% of the sample showing NCRI, then the spacing between the layers would be 500nm
- Grain boundaries: (Burovski et al. PRL 2005) Are there enough grain boundaries?



- No consensus on mechanism

Quantum Monte Carlo

- Premise: we need to use simulation techniques to "solve" many-body quantum problems just as you need them classically.
- Both the wavefunction and expectation values are determined by the simulations. Correlation built in from the start.
- QMC gives most accurate method for general quantum many-body systems.
- QMC determined electronic energy is the standard for approximate LDA calculations. (but fermion sign problem!)
- Path Integral Methods provide a exact way to include effects of ionic zero point motion (include all anharmonic effects) and thermal effects
- A variety of stochastic QMC methods

  - **Variational Monte Carlo VMC (T=0)**
  - **Projector Monte Carlo (T=0)**
    - **Diffusion MC (DMC)**
    - **Reptation MC (RQMC)**
  - **Path Integral Monte Carlo (PIMC) (T>0)**
  - **Coupled Electron-Ion Monte Carlo (CEIMC)**

Variational Monte Carlo (VMC)

- Variational Principle. Given an appropriate trial function:
  - Continuous
  - Proper symmetry
  - Normalizable
  - Finite variance
- Quantum chemistry uses a product of single particle functions
- With MC we can use any "computable" function.
  - Sample R from $|\Psi|^2$ using MCMC.
  - Take average of local energy:
  - Optimize Ψ to get the best upper bound
- Error in energy is 2nd order
- Better wavefunction, lower variance!
- (non-classical) "zero variance" principle.

$$E_V = \frac{\int dR \langle \psi | H | \psi \rangle}{\int dR \langle \psi \psi \rangle} \geq E_0$$

$$\sigma^2 = \frac{\int dR \langle \psi | H^2 | \psi \rangle}{\int dR \langle \psi \psi \rangle} - E_V^2$$

$$E_L(R) = \Re\left[\psi^{-1}(R) H \psi(R)\right]$$

$$E_V = \langle E_L(R) \rangle_{\psi^2} \geq E_0$$

Fermions: antisymmetric trial function

- At mean field level the wavefunction is a Slater determinant. Orbitals for homogenous systems are a filled set of plane waves.
- We can compute this energy analytically (HF).
- To include correlation we multiply by a pseudopotential. We need MC to evaluate properties.
- New feature: how to compute the derivatives of a deteminant and sample the determinant. Use tricks from linear algebra.
- Can we do in O(N)?

$$\Psi_s(R) = Det\left\{e^{ik_i r_j}\eta_i(\sigma_j)\right\}$$

PBC: $k \cdot L = 2\pi n + \{\theta\}$

$$\Psi_{SJ}(R) = Det\{e^{ik_i r_j}\}e^{-\sum_{i<j}u(r_{ij})}$$

Slater-Jastrow trial function.

$$\det\left(\phi_k\left(r_j^T\right)\right) = \det\left(\phi_k\left(r_j\right)\right)\sum_k \phi_k\left(r_j^T\right)M^{-1}_{k,i}$$

$$\frac{1}{\det(M)}\frac{\partial \det(M)}{\partial a} = Tr\left\{M^{-1}\frac{\partial M}{\partial a}\right\}$$

Jastrow factor for correlated electrons

- Look at local energy either in r space or k-space:
- r-space: as 2 electrons get close gives cusp condition: $du/dr|_0=-1$
- k-space, charge-sloshing or plasmon modes.

$$2\rho u_k = \sqrt{\frac{V_k}{\lambda k^2}} \propto \frac{1}{k^2}$$

- Can combine 2 exact properties in the Gaskell form. Write E_V in terms structure factor making "random phase approximation." (RPA).

$$2\rho u_k = -\frac{1}{S_k} + \sqrt{\frac{1}{S_k^2} + \frac{V_k}{\lambda k^2}} \qquad S_k = \text{ideal structure factor}$$

$$\lim_{r\to\infty} u(r) = \begin{cases} r^{-1} & \text{3D} \\ r^{-1/2} & \text{2D} \\ \log(r) & \text{1D} \end{cases}$$

Long range properties

- Give rise to dielectric properties
- Van der Waals interaction

Generalized Feynman-Kacs formula

gives relation between trial function and exact wavefunction!

average "population" starting from a single point R_0 after a imaginary time "t":

$$P(R_0;t) = \int dR \frac{\psi(R)}{\psi(R_0)} \left\langle R \left| e^{-t(H-E_T)} \right| R_0 \right\rangle = \left\langle\!\left\langle e^{-\int_0^t dt E_L(t)} \right\rangle\!\right\rangle_{|\psi|}$$

expand the density matrix in terms of exact eigenstates

$$P(R_0;t) = \int dR \frac{\psi(R)}{\psi(R_0)} \sum_\alpha \phi_\alpha^*(R)\phi_\alpha(R_0) e^{-t(E_\alpha - E_T)}$$

$$\lim_{t\to\infty} P(R_0;t) = \frac{\phi_0(R_0)}{\psi(R_0)} \langle \psi\phi_0 \rangle \qquad E_L(R) = \psi^{-1}\hat{H}\psi(R)$$

$$\frac{\phi_0(R_0)}{\psi(R_0)} \sim e^{-\int_0^t dt \langle\langle E_L(t)\rangle\rangle_{|\psi|}}$$

Wavefunctions beyond Jastrow

- Use method of residuals construct a sequence of increasingly better trial wave functions. Justify from the Importance sampled DMC.
- Zeroth order is Hartree-Fock wavefunction
- First order is Slater-Jastrow pair wavefunction (RPA for electrons gives an analytic formula)
- Second order is 3-body backflow wavefunction
- Three-body form is like a squared force. It is a bosonic term that does not change the nodes.

$$\exp\{\sum_i [\sum_j \xi_{ij}(r_{ij})(\mathbf{r}_i - \mathbf{r}_j)]^2\}$$

smoothing

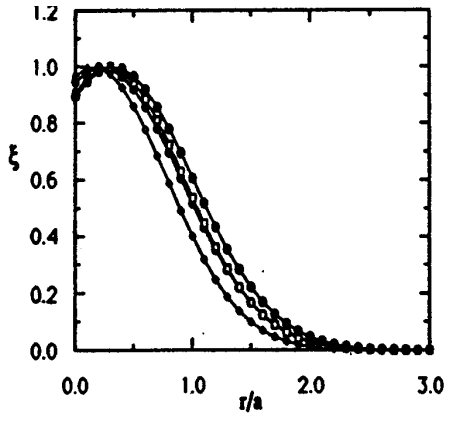
$$\phi_{n+1}(R) \approx \phi_n(R) e^{-\tau < \phi_n^{-1} H \phi_n >}$$

$$\phi_0 = e^{i\sum_j \mathbf{k}_j \bullet \mathbf{r}_j}$$

$$E_0 = V(R)$$

$$\phi_1 = \phi_0 e^{-U(R)}$$

$$E_1 = U(R) - [\nabla W(R)]^2 + i\sum_j \mathbf{k}_j \bullet (\mathbf{r}_j - \nabla_j Y(R))$$



Backflow wave function

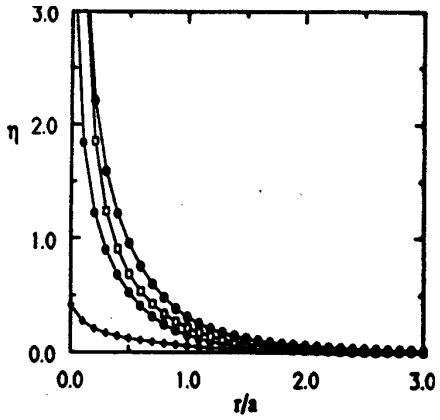
- Backflow means change the coordinates to quasi- coordinates.

$$Det\{e^{i\mathbf{k}_i\mathbf{r}_j}\} \Rightarrow Det\{e^{i\mathbf{k}_i\mathbf{x}_j}\}$$

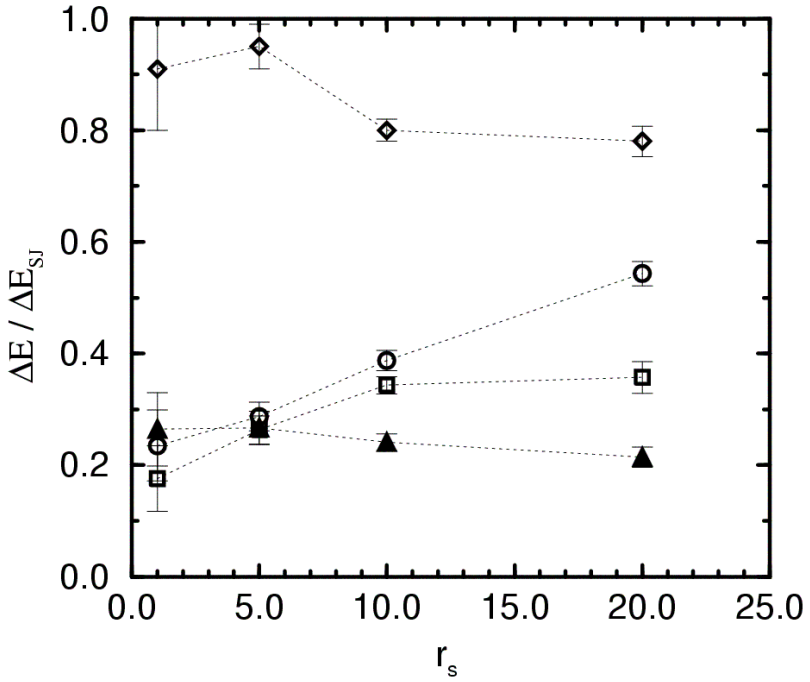
$$\mathbf{x}_i = \mathbf{r}_i + \sum_j \eta_{ij}(r_{ij})(\mathbf{r}_i - \mathbf{r}_j)$$

- Leads to a much improved energy and to improvement in nodal surfaces. Couples nodal surfaces together.

Kwon PRB 58, 6800 (1998).



3DEG



Projector Monte Carlo

(variants: Green's function MC, Diffusion MC, Reptation MC)

- Project single state using the Hamiltonian

$$\phi(t) = e^{-(\mathrm{H}-\mathrm{E_T})\mathrm{t}}\phi(0)$$

- We show that this is a diffusion + branching operator. Maybe we can interpret as a probability. But is this a probability?
- **Yes!** for bosons since ground state can be made real and non-negative.
- **But** all excited states must have sign changes. This is the "sign problem."
- For efficiency we do "importance sampling."
- Avoid sign problem with the fixed-node method.

Diffusion Monte Carlo

- How do we analyze this operator?

$$\psi(R,t) = e^{-(\mathrm{H}-\mathrm{E_T})\mathrm{t}}\psi(R,0)$$

- Expand into exact eigenstates of H.

$$H\phi_\alpha = E_\alpha\phi_\alpha$$

$$\psi(R,0) = \sum_\alpha \phi_\alpha(R)\left\langle \phi_\alpha \middle| \psi(0)\right\rangle$$

- Then the evolution is simple in this basis.

$$\psi(R,t) = \sum_\alpha \phi_\alpha(R) e^{-t(E_\alpha - E_T)} \left\langle \phi_\alpha \middle| \psi(0)\right\rangle$$

- Long time limit is lowest energy state that overlaps with the initial state, usually the ground state.

$$\lim_{t\to\infty} \psi(R,t) = \phi_0(R) e^{-t(E_0 - E_T)} \left\langle \phi_0 \middle| \psi(0)\right\rangle$$

$$E_0 \approx E_T \Rightarrow \textit{normalization fixed}$$

Importance Sampling

- Why should we sample the wavefunction? The physically correct pdf is $|\phi|^2$.
- Importance sample (multiply) by trial wave function.

$$f(R,t) \equiv \psi_T(R)\phi(R,t) \qquad \lim_{t\to\infty} f(R,t) \equiv \psi_T(R)\phi_0(R)$$

$$-\frac{\partial f(R,t)}{\partial t} = \psi_T(R) H\left[f(R,t)/\psi_T(R)\right]$$

Commute Ψ through H

$$-\frac{\partial f(R,t)}{\partial t} = -\lambda\nabla^2 f - \lambda\nabla\left(2f\nabla\ln\psi_T(R)\right) + \left(\psi_T{}^{-1} H \psi_T\right) f(R,t)$$

Evolution = diffusion + drift + branching

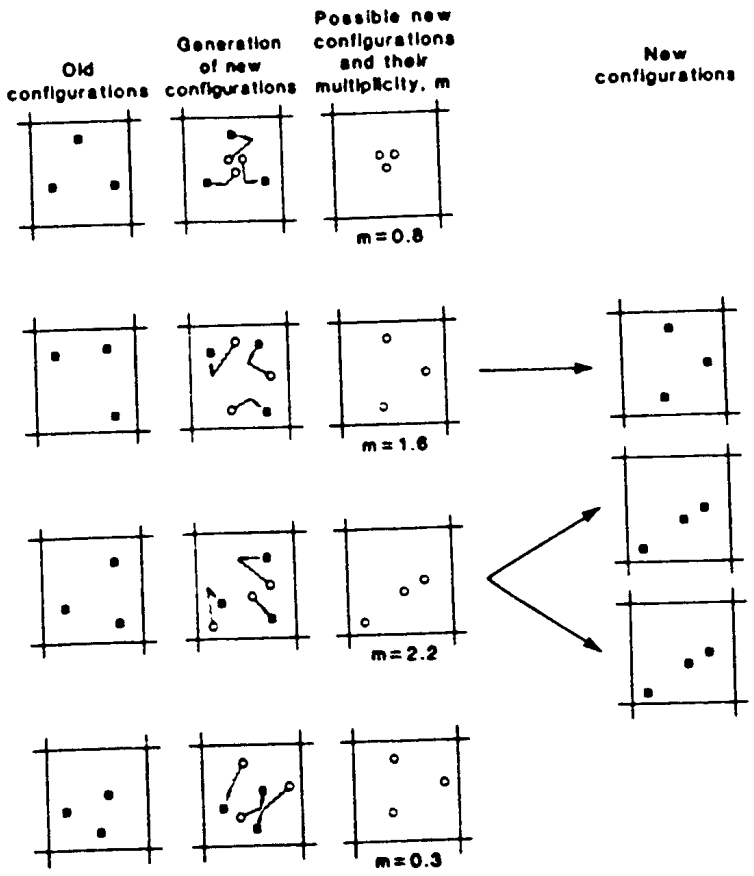
- We have three terms in the evolution equation. Trotter's formula applies.

Schematic of DMC

Ensemble evolves according to

- Diffusion
- Drift
- branching

ensemble



Fermions?

- How can we do fermion simulations? The initial condition can be made real but not positive (for more than 1 electron in the same spin state)

- In *transient estimate* or *released-node* methods one carries along the sign as a weight and samples the modulus.

$$\phi(t) = e^{-(\hat{H}-E_T)t}\,\mathrm{sign}(\phi(R,0))\,|\,\phi(R,0)\,|$$

- Do not forbid crossing of the nodes, but carry along sign when walks cross.
- What's wrong with node release:
  - Because walks don't die at the nodes, the computational effort increases (bosonic noise)
  - The signal is in the cancellation which dominates

Monte Carlo can add but not subtract

Fixed-node method

- Initial distribution is a pdf.
  It comes from a VMC simulation.
- Drift term pushes walks away from the nodes.
- Impose the condition:
- This is the fixed-node BC
- Will give an upper bound to the exact energy, the best upper bound consistent with the FNBC.

$$f(R,0) = |\psi_T(R)|^2$$

$$\phi(R) = 0 \quad \text{when} \quad \psi_T(R) = 0.$$

$$E_{FN} \geq E_0$$

$$E_{FN} = E_0 \quad \text{if} \quad \phi_0(R)\psi(R) \geq 0 \quad \text{all } R$$

- f(R,t) has a discontinuous gradient at the nodal location.
- Accurate method because Bose correlations are done exactly.
- Scales well, like the VMC method, as N^3. Classical complexity.
- Can be generalized from the continuum to lattice finite temperature, magnetic fields, …
- One needs trial functions with accurate nodes.

Fixed-Phase method

- Generalize the FN method to complex trial functions: $\Psi(R) = e^{-U(R)}$
- Since the Hamiltonian is Hermitian, the variational energy is real:

$$E_V = \frac{\int dR\, e^{-2\Re U(R)} \left[V(R) + \lambda \nabla^2 U(R) - \lambda \left[\Re \nabla U(R) \right]^2 + \lambda \left[\Im \nabla U(R) \right]^2 \right]}{\int dR\, e^{-2\Re U(R)}}$$

- We see only one place where the energy depends on the phase of the wavefunction.
- If we fix the phase, then we add this term to the potential energy. In a magnetic field we get also the vector potential.

$$\text{effective potential} = V(R) + \sum_i \lambda_i \left[A(r_i) + \Im \nabla_i U(R) \right]^2$$

- We can now do VMC or DMC and get upper bounds as before.
- The imaginary part of the local energy will not be zero unless the right phase is used.
- Used for twisted boundary conditions, magnetic fields, vortices, phonons, spin states, …

Problem with core electrons

- Bad scaling in both VMC and DMC
- In VMC, energy fluctuations from core dominate the calculation
- In DMC, time step will be controlled by core dynamics
- Solution is to eliminate core states by a pseudopotential

- Conventional solution: semi-local form

$$\left\langle r\left|\hat{v}_{e-core}\right|r'\right\rangle = v_{local}(r)\delta(r-r') + \sum_{l} v_l(r) P_l(\cos(r\cdot r'))$$

- Ensures that valence electrons go into well defined valence states with the wavefunction and energy for each angular momentum state prescribed.
- PP is non-local: OK for VMC. Leads to an extra MC integral. But DMC uses a locality approximation and needs good trial functions. Extra approximation.

Defect Calculations with QMC

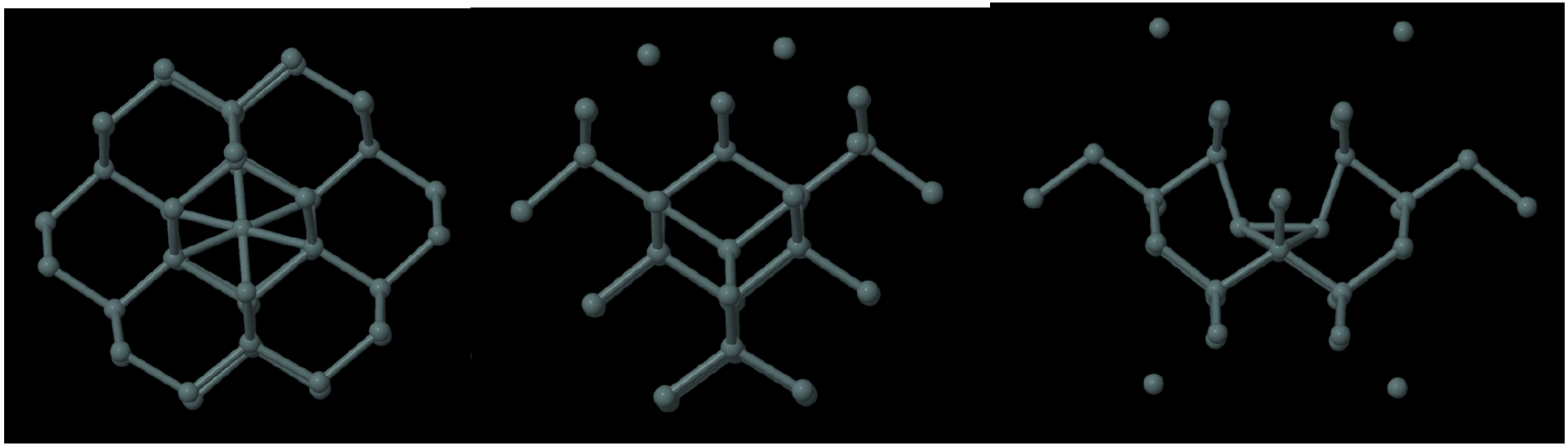
			energy	DFT			*GW*	DMC	exp.	Refs.
				LDA	GGA	hybrid				
C	diamond vacancy		formation	6.98	7.51	-	-	5.96(34)	-	[49, 50]
			migration	2.83	-	-	-	4.40(36)	2.3(3)	
MgO	Schottky defect		formation	5.97, 6.99, 6.684	-	-	-	7.50(53)	5–7	[51, 52]
Si	self-interstitial defect	X		3.31	3.64	4.69	4.40	5.0(2), 4.94(5)	-	[26, 24]
		T	formation	3.43	3.76	4.95	4.51	5.5(2), 5.13(5)	-	
		H		3.31	3.84	4.80	4.46	4.7(2), 5.05(5)	-	

Parker, Wilkins and Hennig, Phys. Status Solidi B 248, No. 2, 267–274 (2011)

1999 Leung PRL 83 2351	Si Interstitial
2003 Hood PRL 91 076403	C Vacancy
2005 Alfe PRB 71 220101	MgO Charged Schottky Defect
2006 Batista PRB 74 121102	Si Interstitial
2010 Parker Physica St. Sol.,	Si Interstitial
2012 Krogel, Kim, Ceperley	Ge interstitials.

Germanium & Silicon interstitials

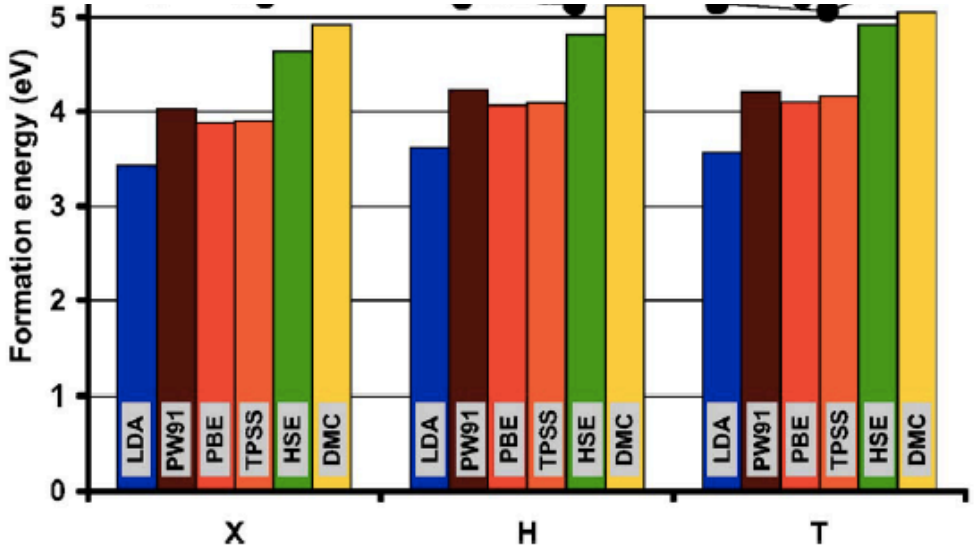
H T X



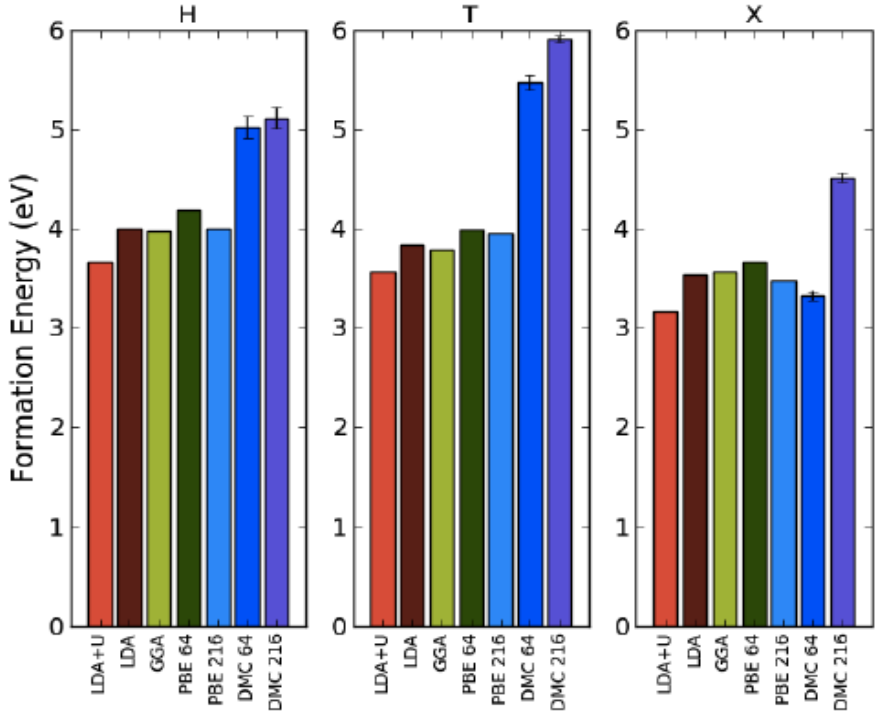
Batista et al. PRB 74, 121102R 2006

- QMC energies are larger than DFT by 1eV
- Problems?
  - Finite size
  - Pseudopotential
  - Geometry
  - Nodes

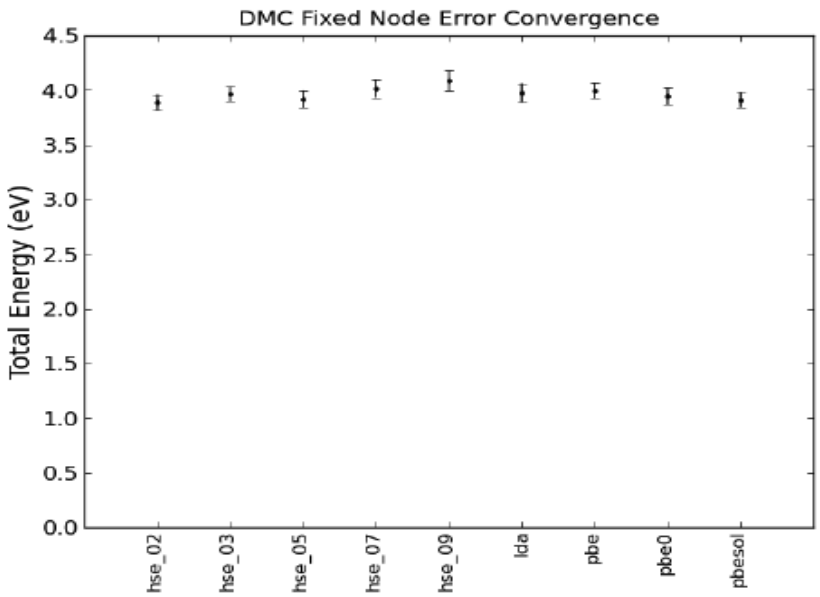
Si

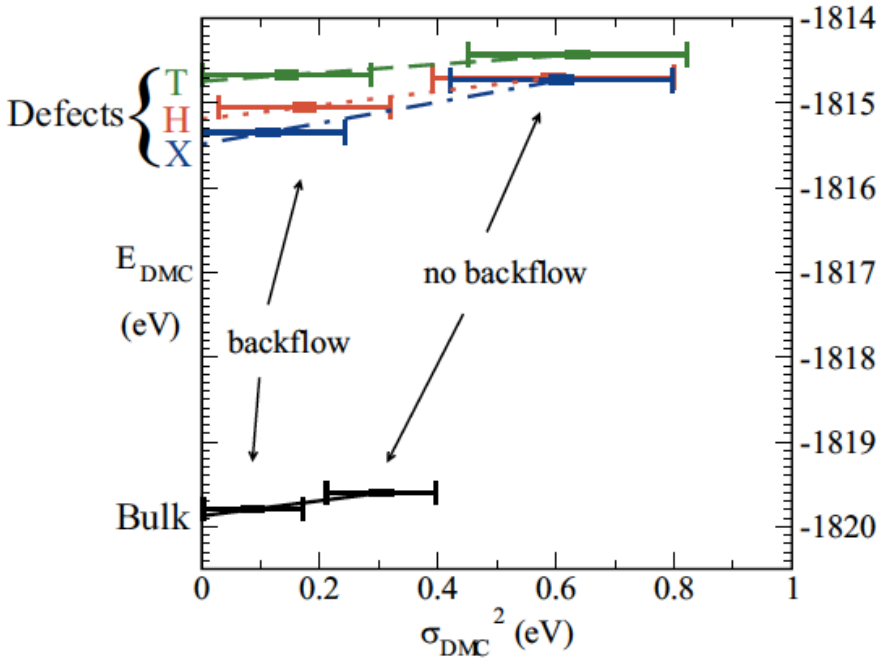


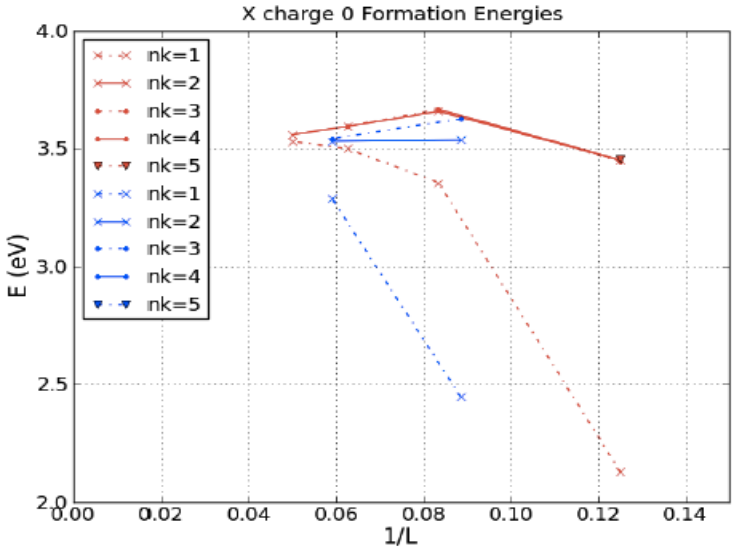
Ge



Krogel, Kim, DMC 2012







- PBE relaxations & scf at a_{latt}=5.66 A
- red→fcc cells
  blue→sc cells
  nk→twist grid
- cubic converges fastest
- 2x2x2 kpoints near convergence for 64/216
- twist averaging increases formation energies

Prospects for QMC defect studies

- We can do 200,000 bosons with PIMC, matching experiment on cold atoms
- Brute force approach for defects has a scaling problem
- Can we scale better with
  - Energy density
  - Local trial functions, sparse matrices
  - GSPIMC methods for focusing CPU time
- Some advantages of QMC
  - Benchmark quality (with bounds)
  - Ability to treat disorder
  - Non-zero temperature with PIMC

Scaling of defect calculation

- Assume variance proportional to energy error (extensive)

$$total\_error = \left(\frac{N(E_V - E_0)}{\tau M}\right)^{1/2}$$

$$CPU\_time = \left(N^3 + aN^2 + bN + c\right) M \,/\, flops$$

$$M = total\_number\_steps$$

$$CPU\_time \propto \frac{\left(E_V - E_0\right) N^4}{error^2}$$

- Energy density method reduces it to N^3
- Localized orbitals can give another factor of N.
- GS-PIMC can give another factor of N.

Compute energy density at a given atom

Properties of $\hat{\mathcal{E}}_r$

- normalization: $\int dr \hat{\mathcal{E}}_r = \hat{H}$
- $\hat{H}$ symmetries: particle addition, exchange, translation, etc.
- particles carry the energy, not field
- neutral subsystems transfer energy slower than $1/r$
- no left acting derivatives (no derivatives of DMC projector)

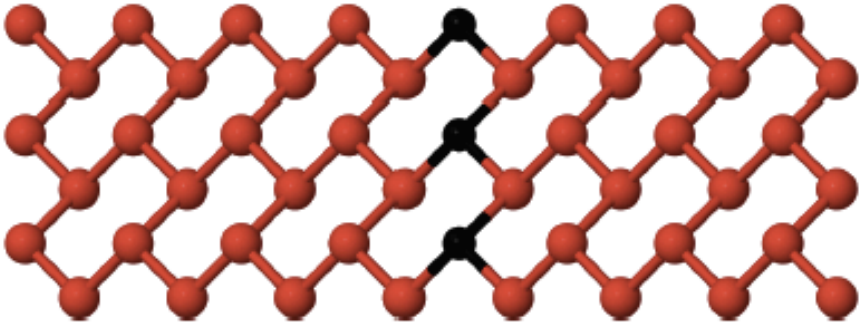
Unique form for $\hat{\mathcal{E}}_r$

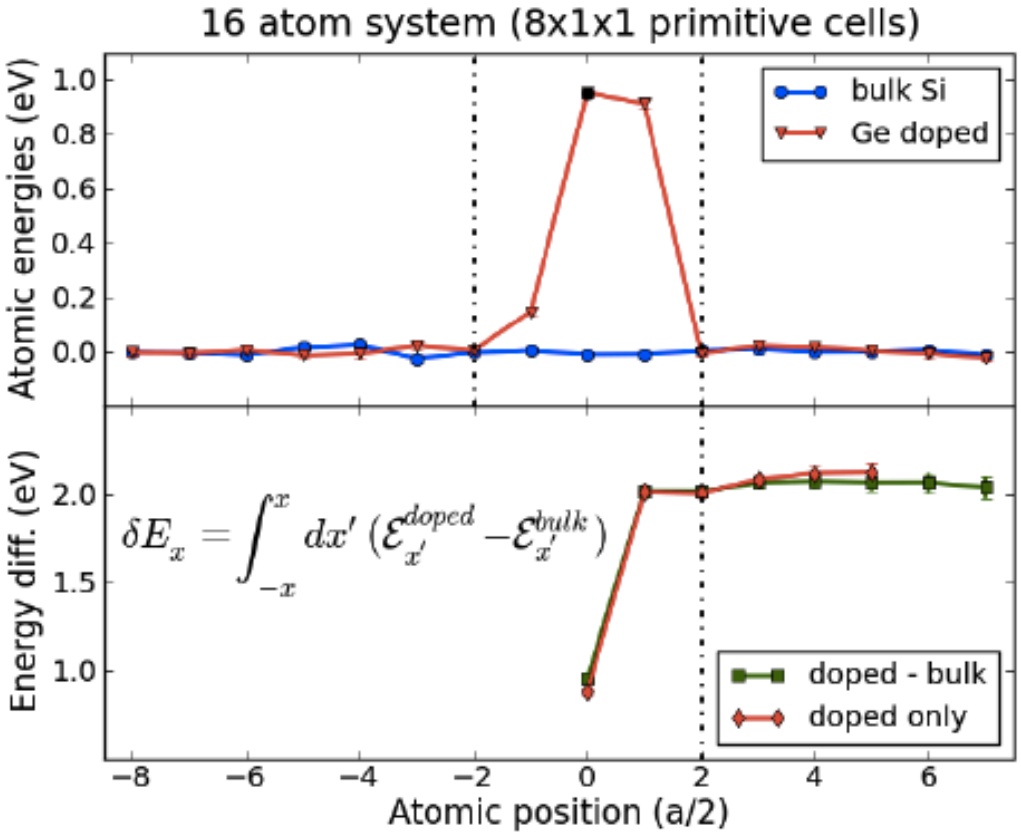
$$\hat{\mathcal{E}}_r = \sum_i \delta_{rr_i} \frac{\hat{p}_i^2}{2m_e} + \sum_{i<j} \frac{\delta_{rr_i} + \delta_{rr_j}}{2} \hat{v}_{ij}^{ee} + \sum_{i\mu} \frac{\delta_{rr_i} + \delta_{rr_\mu}}{2} \hat{v}_{i\mu}^{el}$$

- $\hat{\mathcal{E}}_r$ equivalent to $(\hat{\mathcal{E}}_r + \hat{\mathcal{E}}_r^\dagger)/2$ if Ψ real or k-points invertable

- Divide up space into Voronoi cells to record E_r
- Lost zero variance property of QMC

Efficient QMC Energy Differences: δ-doped Si





Formation Energy

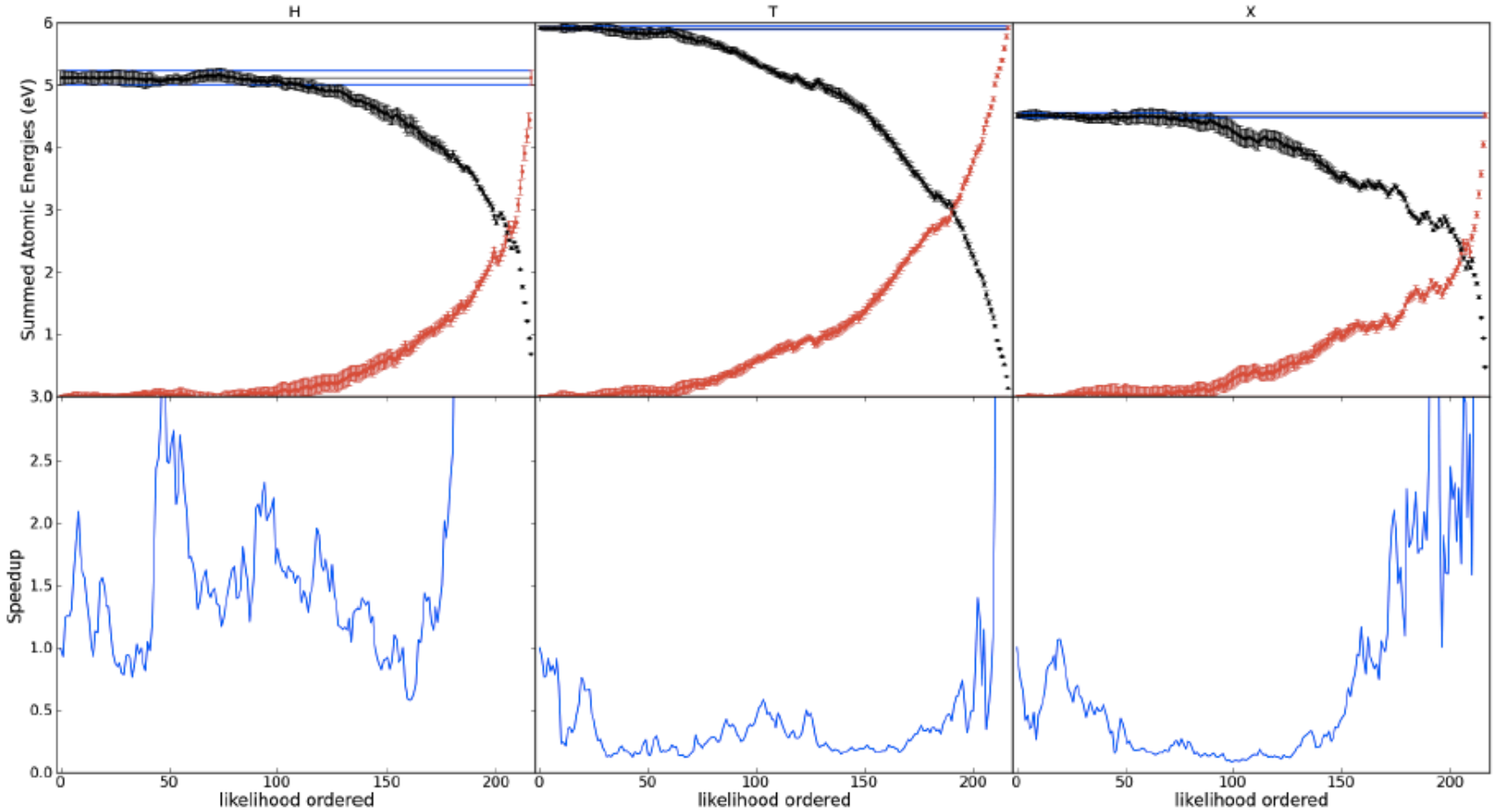
$$E_F = (E_{tot}^{doped} - E_{tot}^{bulk}) + (\mu_{Ge} - \mu_{Si})\delta N$$

Speedup

- ratio of walltimes at fixed error bar
- total energy difference vs. energy density

doped-bulk	9.7
doped only	20.5

Germanium Interstitials



- Speed-up not evident for interstitial
- Need a larger cell

Ground State Path Integrals

- Project a trial function as in DMC
- Similar to VMC
- Can concentrate moves to neighborhood of defect.

$$\Psi(\beta) = e^{-\frac{\beta}{2}H}\Psi$$

ψ(β) converges to the exact ground state as a function of imaginary time.

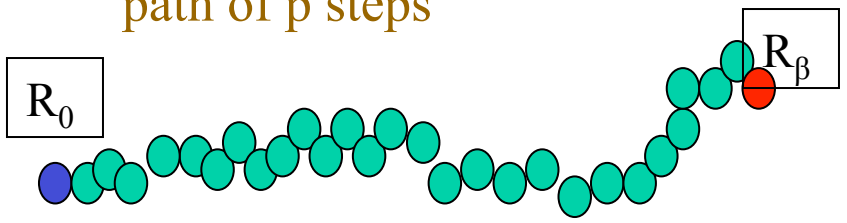
$$E(\beta) = \frac{\langle \Psi(\beta) H \Psi(\beta) \rangle}{\langle \Psi(\beta)\Psi(\beta) \rangle} = \langle E_L(R_0) \rangle_\beta$$

E(β) is an upper bound converging to the exact answer monotonically

$$Z(\beta) = \langle \Psi(\beta)\Psi(\beta) \rangle = \langle \Psi e^{-\beta H}\Psi \rangle = \int dR_0 \ldots dR_p \Psi(R_0) \langle R_0 e^{-\tau H} R_1 \rangle \ldots \langle R_{p-1} e^{-\tau H} R_p \rangle \Psi(R_p)$$

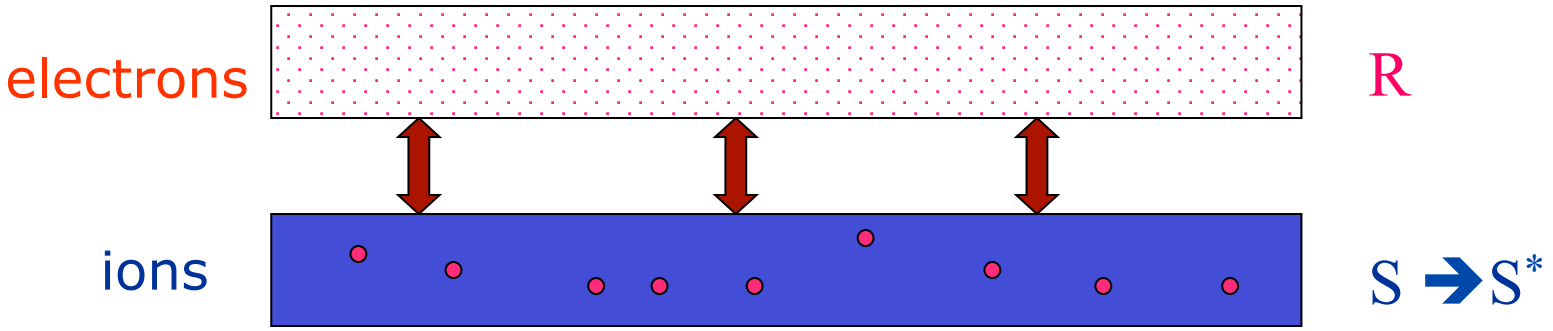
$$\tau = \frac{\beta}{p} = \text{timestep}$$

Do Trotter break-up into a path of p steps



Coupled Electron-Ionic Monte Carlo:CEIMC

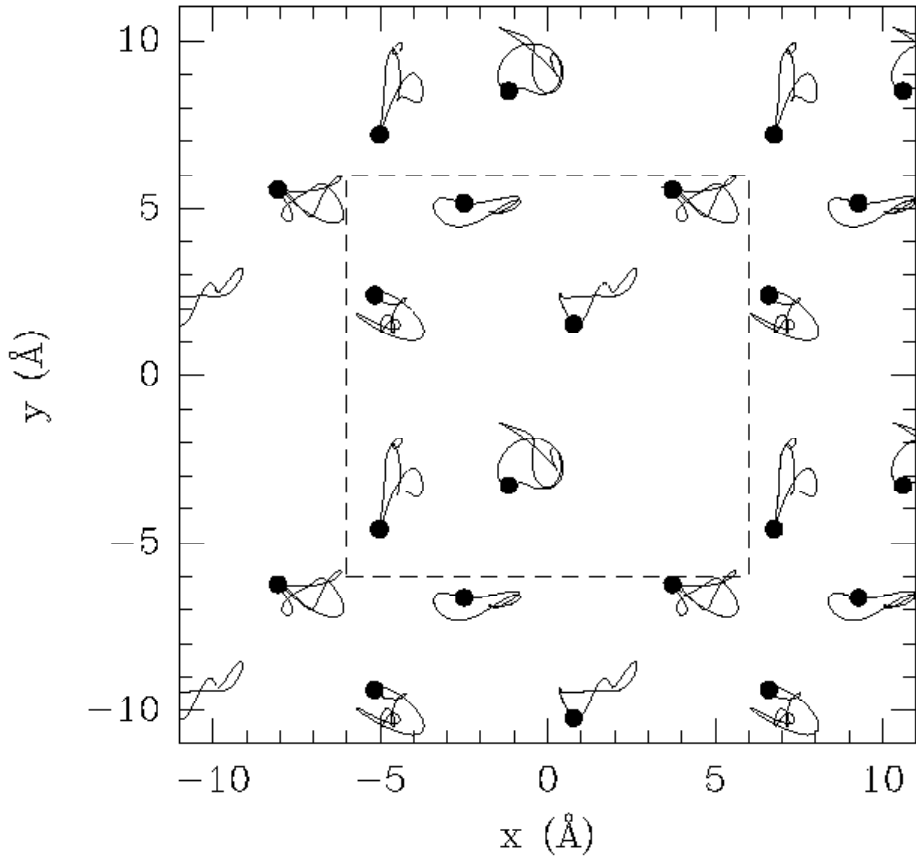
1. Do Classical MC for the ions at T>0.
2. Let electrons be at zero temperature, a reasonable approximation for room temperature simulations.
3. Use Metropolis MC to accept/reject moves based on QMC computation of electronic energy



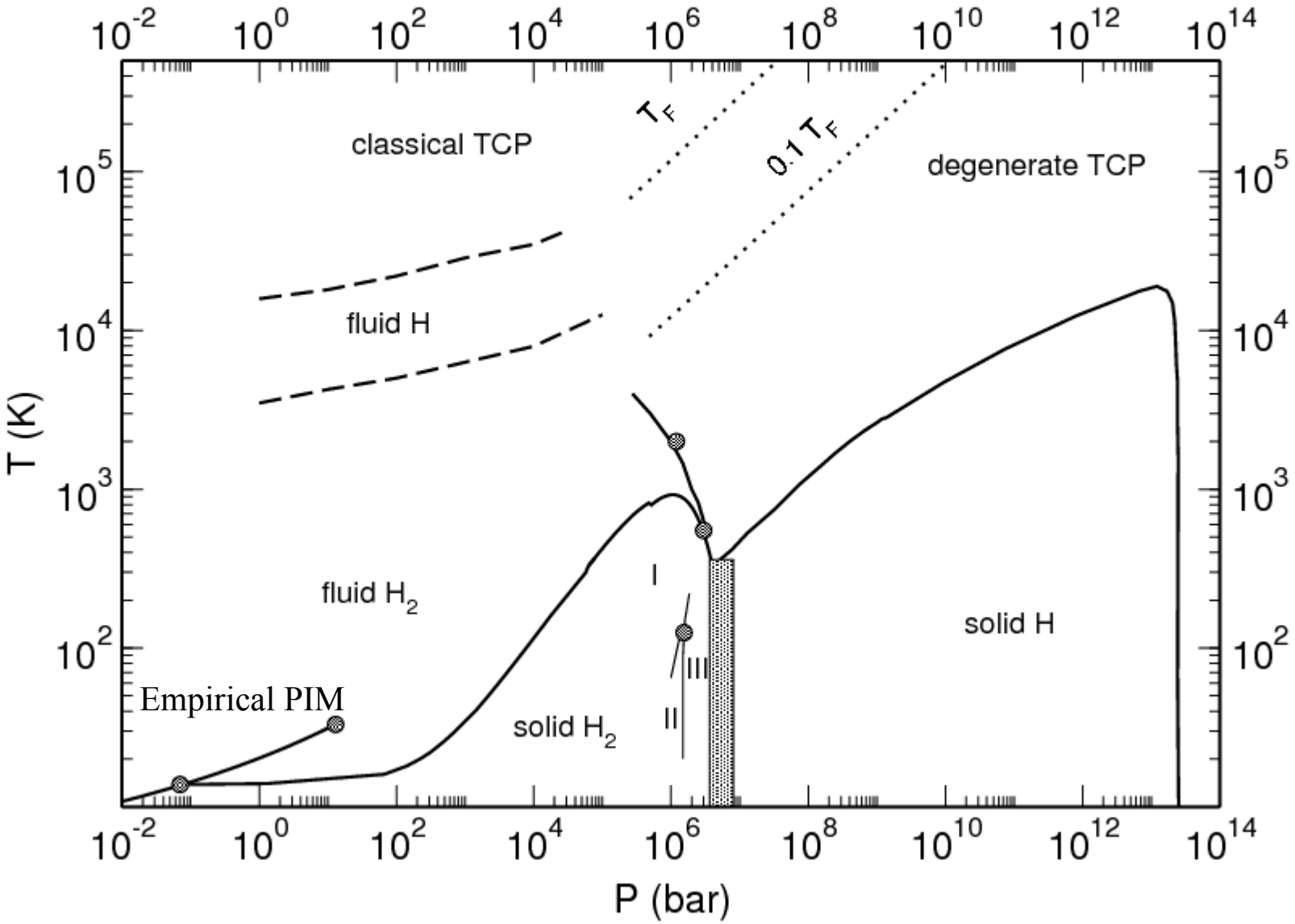
The "noise" coming from electronic energy can be treated without approximation using the penalty method.

Path Integral MC (Distinguishable statistics)

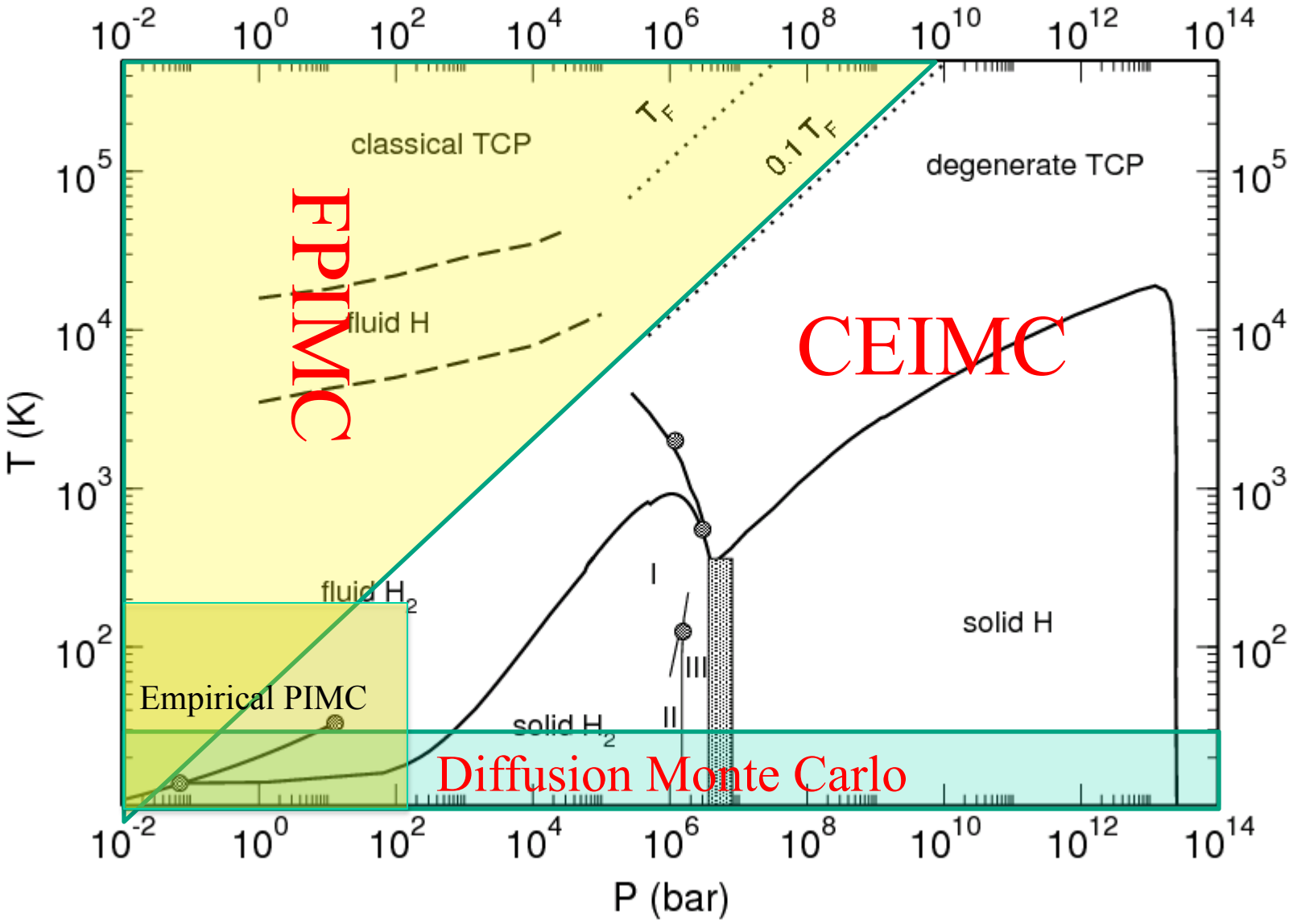
- Each particle is a ring polymer; an exact representation of a quantum wavepacket in imaginary time.
- Integrate over all paths
- The dots represent the “start” of the path. (but all points are equivalent)
- The lower the real temperature, the longer the “string” and the more spread out the wavepacket.
- Quantum statistics: all crosslinking of paths
- Restricted fermion paths: allow only those keeping density matrix positive.
- Used to calculate defect energies in quantum crystals: helium and wigner.



dense H


classical TCP
degenerate TCP
T_F
$0.1\,T_F$
fluid H
fluid H_2
solid H_2
solid H
I
II
III
Empirical PIM
T (K)
P (bar)
10^{-2}
10^{0}
10^{2}
10^{4}
10^{6}
10^{8}
10^{10}
10^{12}
10^{14}
10^{2}
10^{3}
10^{4}
10^{5}

Regimes for Quantum Monte Carlo



COMPARISON OF QMC METHODS

they are all based on thermal density matrix

they share features: the action, the sampling

DMC

- State space: P walkers
- Action/Dynamics determined by H
- Observables have mixed estimator problem-need to do forward walking
- No permutations
- Trial function & trial energy
- Population bias
- Energy is "zero variance""

GS-PIMC

- State space: reptile & direction
- Action fixed/Dynamics arbitrary
- Single particle moves more difficult
- Observables from middle of path
- Trial function
- Projection bias
- Energy is "zero variance"

$$\left\langle R \left| e^{-\tau H} \right| R' \right\rangle$$

PIMC

- State space: closed path & permutation
- Action fixed/Dynamics arbitrary
- Observables anywhere
- Particle statistics=Permutations
- Temperature>0! T=0 expensive.

Disorder in Monte Carlo

Extra averaging is free! (almost)

Types of averaging we use:

1. Path Integrals for ions (for protons or light ions)

   (M_1 time slices to average over.)
2. k-point sampling (integrate over Brillouin zone of supercell). Twist averaged boundary conditions converge much faster than periodic boundary conditions for metals. (M_2 k-points)
3. Disorder averaging

- In explicit methods such as AIMD these extra variables will increase the CPU time by $M_1 M_2$.
- With QMC there will be little increase in time since imaginary time and/or k are simply new variables to average over. Increase sample space.
- Increase in parallelization

New QMC Techniques

- Algorithms (e.g. reptation)
- Better Finite-Size scaling methods
  - Twist averaging for kinetic energy
  - Coulomb corrections for potential energy
- Better trial wavefunctions -> better nodes
  - Backflow
  - Direct coupling to DFT
- Coupled Electron-Ion Monte Carlo
- Optimization of trial function parameters
- Computers/parallelization: huge increase in available resources. Factors of 1000 / decade.

Limitations of QMC

- Technical (programming) issues
- Approximations that are controlled:
  - Iterations of Markov chain.
  - Size of supercell
  - Time step
- Geometries (forces) come from other methods
- Nodes (phases) of density matrix or trial function
- Pseudopotentials come from other methods

- QMC is valuable as benchmark.
- Unique capabilities for disorder, finite temperature…