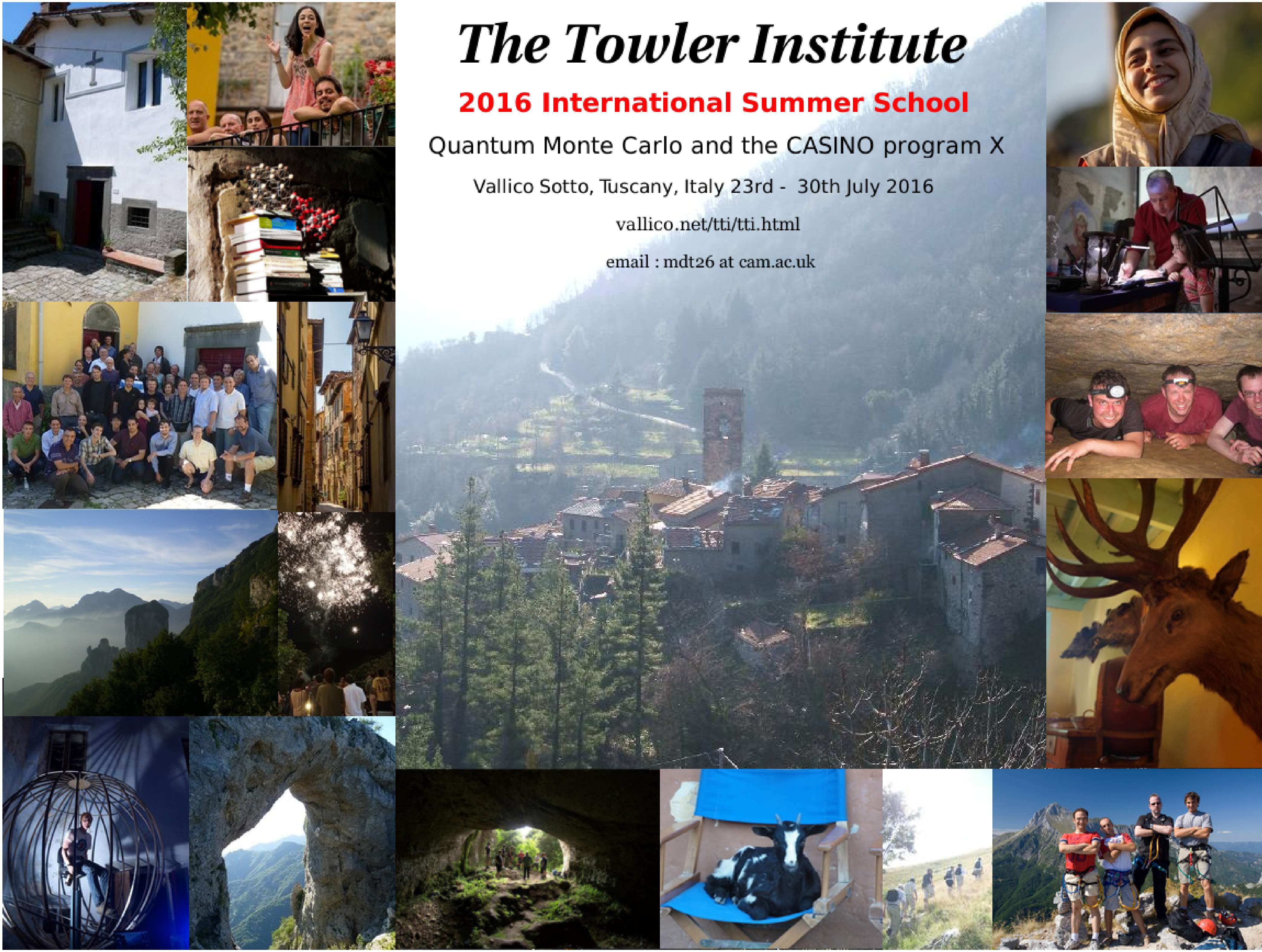


The Towler Institute

2016 International Summer School

Quantum Monte Carlo and the CASINO program X

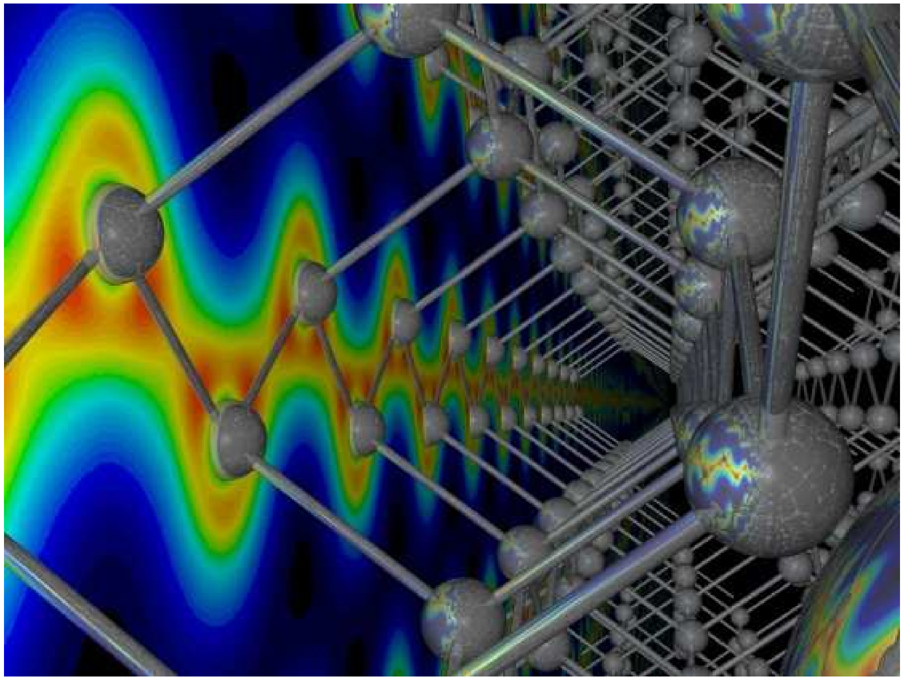
Vallico Sotto, Tuscany, Italy 23rd - 30th July 2016

vallico.net/tti/tti.html

email : mdt26 at cam.ac.uk

Quantum Monte Carlo

A practical solution to the correlation problem in electronic structure calculations



Mike Towler

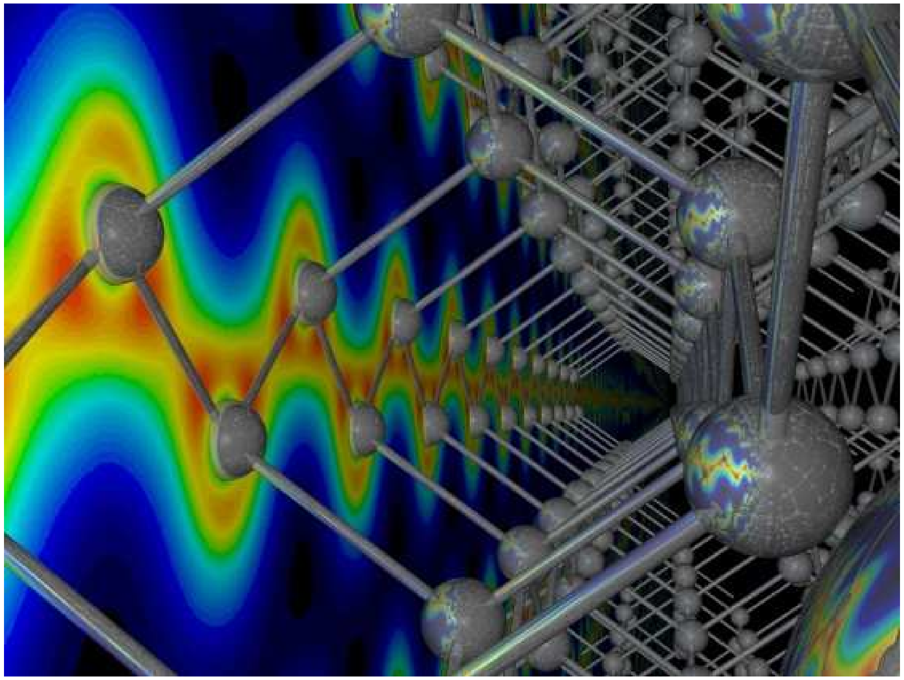
TCM Group, Cavendish Laboratory, University of Cambridge

QMC web page: `vallico.net/casinoqmc`

Email: `mdt26@cam.ac.uk`

Quantum Monte Carlo

A practical solution to the correlation problem in electronic structure calculations



Mike Towler

~~*TCM Group, Cavendish Laboratory, University of Cambridge*~~
UNEMPLOYED!

QMC web page: `vallico.net/casinoqmc`

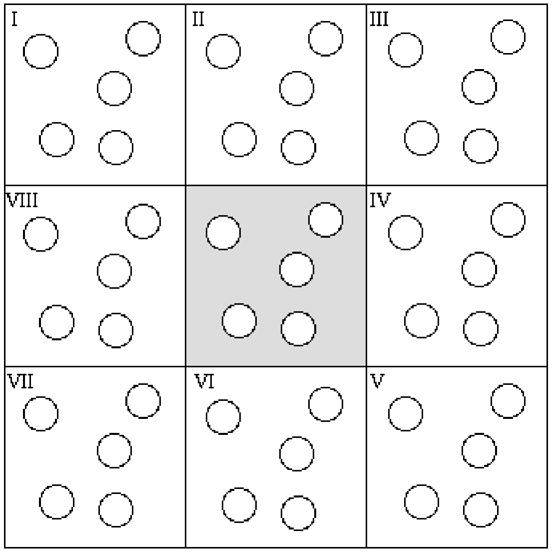
Email: `mdt26@cam.ac.uk`

Keil Art Bronze

Programme for Sunday

8.30am	Lecture 1	Basic theory: variational and diffusion Monte Carlo
11.00am	BREAK	
11.15am	Lecture 2	Introduction to the CASINO program
12.15am	Lecture 3	Distribution, setup, and compilation of CASINO. Local and remote computer resources. Applying for computer time. Web resources.
12.45	LUNCH	Village bar
1.45pm	QMC Q+A	Ask Neil and Mike. What do you want to do with QMC?
4pm	EXCURSION	Alta Matanna (3pm for Monte Procinto)
7.30pm	DINNER	
c. 10.30pm		Mike and Sam will be available in the church for help with CASINO setup on personal machines.

FIRST PRINCIPLES CALCULATIONS



Atomic numbers

↓

Solve the quantum mechanical equations for the electrons

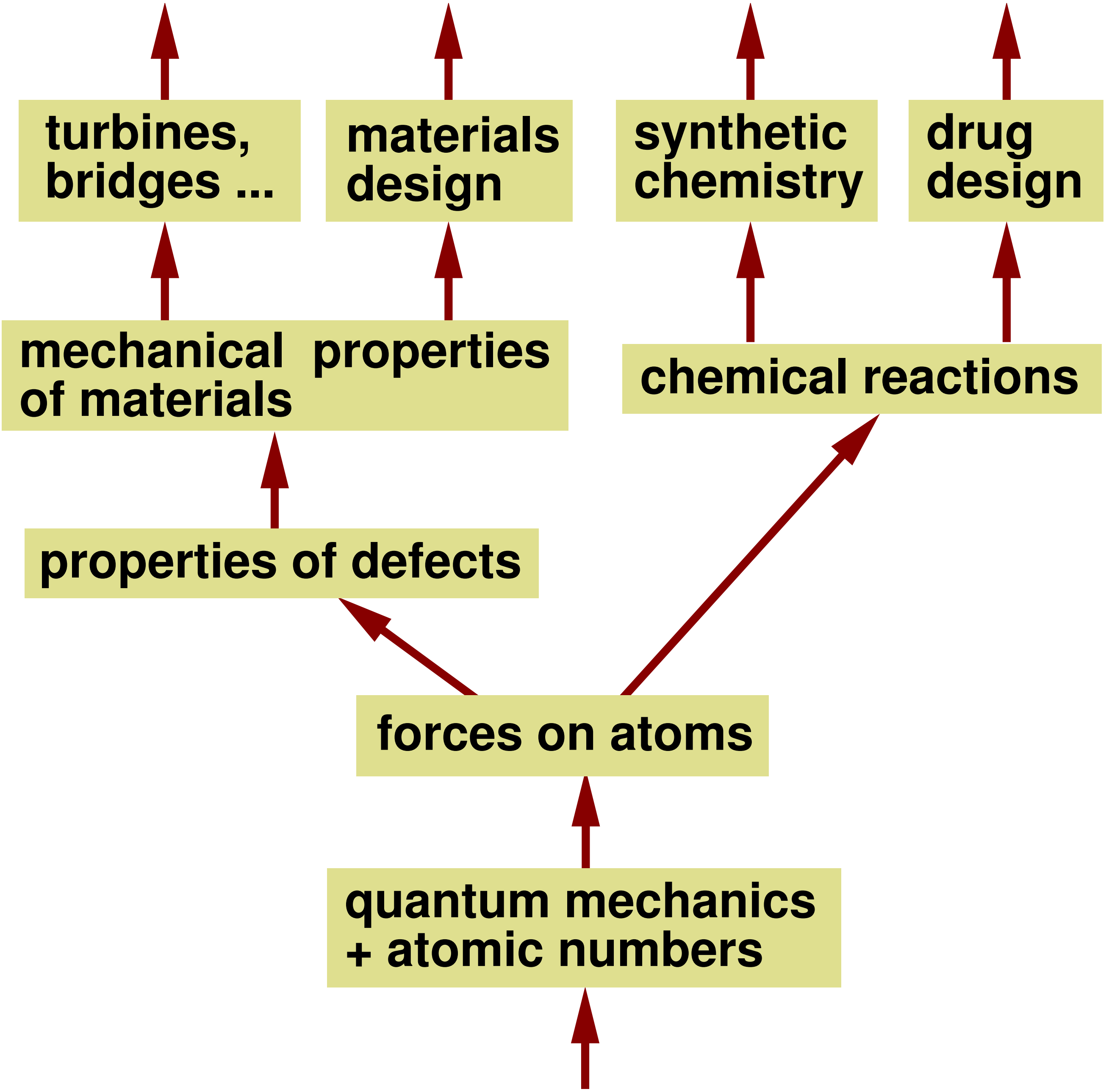
↓

(Follow the time evolution of the nuclei)

↓

Understand and predict physical and chemical properties of the system

Why?



turbines, bridges ...

materials design

synthetic chemistry

drug design

mechanical properties of materials

chemical reactions

properties of defects

forces on atoms

quantum mechanics + atomic numbers

Electronic Structure Theory

Justification for the existence of the Cambridge TCM group

Ab initio electronic structure calculations of materials *complement* experiments, by

- helping to understand the results of real experiments
- calculating things that are experimentally inaccessible or cost money
- providing an atomic scale resolution beyond most experiments
- calculating things which don't exist (in experiments we are usually stuck with the lowest free energy phase unless there is a large activation energy so that metastable phases can be studied)

This is generally done by 'solving the Schrödinger equation' entirely from first principles with no adjustable parameters.

Electronic Structure Theory

Justification for the existence of the Cambridge TCM group

Ab initio electronic structure calculations of materials *complement* experiments, by

- helping to understand the results of real experiments
- calculating things that are experimentally inaccessible or cost money
- providing an atomic scale resolution beyond most experiments
- calculating things which don't exist (in experiments we are usually stuck with the lowest free energy phase unless there is a large activation energy so that metastable phases can be studied)

This is generally done by 'solving the Schrödinger equation' entirely from first principles with no adjustable parameters.

PLUS since we can only do this approximately, we try to

- improve the approximations to get better accuracy
- improve the speed and scaling behaviour of the algorithms
- work out how to use the theory to calculate new things

Electronic Structure Theory

Justification for the existence of the Cambridge TCM group

Ab initio electronic structure calculations of materials *complement* experiments, by

- helping to understand the results of real experiments
- calculating things that are experimentally inaccessible or cost money
- providing an atomic scale resolution beyond most experiments
- calculating things which don't exist (in experiments we are usually stuck with the lowest free energy phase unless there is a large activation energy so that metastable phases can be studied)

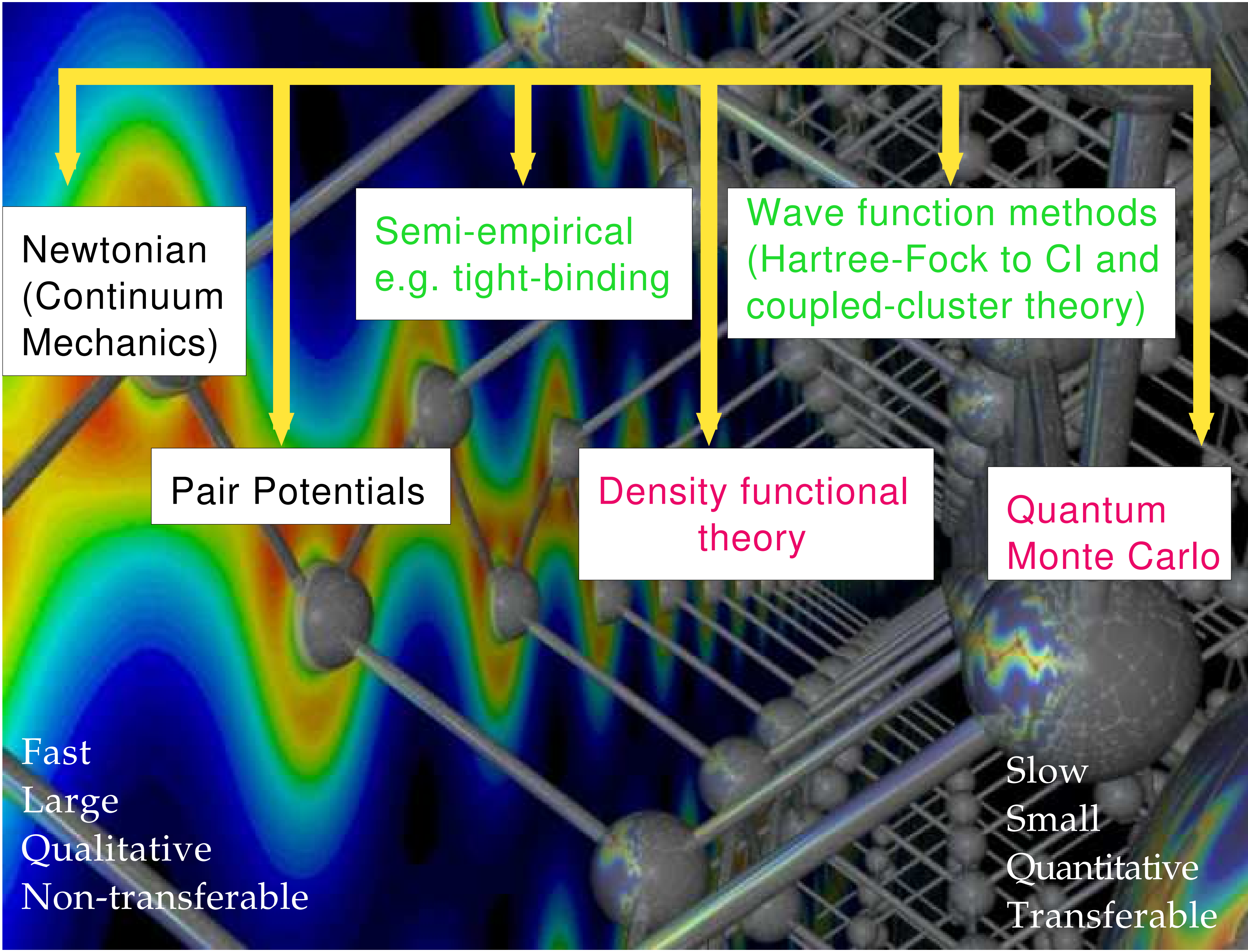
This is generally done by 'solving the Schrödinger equation' entirely from first principles with no adjustable parameters.

PLUS since we can only do this approximately, we try to

- improve the approximations to get better accuracy
- improve the speed and scaling behaviour of the algorithms
- work out how to use the theory to calculate new things

PLUS, as a service to the scientific community,

- we write generally applicable computer codes which do all of the above.



Newtonian (Continuum Mechanics)

Pair Potentials

Semi-empirical e.g. tight-binding

Density functional theory

Wave function methods (Hartree-Fock to CI and coupled-cluster theory)

Quantum Monte Carlo

Fast
Large
Qualitative
Non-transferable

Slow
Small
Quantitative
Transferable

The cliché of Paul Dirac

“The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.”

The cliché of Paul Dirac

“The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.”.

NONSENSE!

That was in 1928. We can do better now - because we have quantum Monte Carlo. And one of these..

What is quantum Monte Carlo?

QMC is, essentially, '*accurate calculation of the properties of assemblies of interacting quantum particles using* **random sampling** *of the full many-body Schrödinger wave function..*' For my purposes, it is applied to 'continuum fermion' cases i.e. realistic atoms, molecules, and crystalline solids rather than bosonic systems or lattice models.

Why use wave functions? We work directly with the many-body wave function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ rather than the density $\rho(\mathbf{x})$ since we can write down exactly the equation that Ψ obeys (a very helpful thing to do when requiring *accuracy*. . .):

$$\boxed{\hat{H}\Psi = E\Psi}$$

- *Variational Monte Carlo* (VMC) - a cheaper less accurate technique which directly computes expectation values such as the one below using standard Monte Carlo numerical integration. May also involve optimization of Ψ by direct variation of any parameters upon which it depends. Limited by flexibility of parameterization..

$$E_{\mathrm{VMC}} = \frac{\int \Psi \hat{H} \Psi \, \mathrm{d}\mathbf{x}}{\int \Psi^2 \, \mathrm{d}\mathbf{x}}$$

- *Diffusion Monte Carlo* (DMC) - a more sophisticated, accurate (and expensive) method which - in principle - projects out the exact ground state wave function from a given 'starting guess' and then computes expectation values as above.

Large systems

Anyone can solve the Schrödinger equation as accurately as you like for a few particles - that's a given. The key point is *scaling* with system size - in the modern age N^7 will not do. How can one address the problem for a huge molecule or a crystalline solid?

$$\hat{H}\Psi = \sum_{i=1}^{6\times 10^{23}} \left(-\frac{\hbar^2}{2m}\nabla_i^2\Psi - Ze^2\sum_{\mathbf{R}}\frac{1}{|\mathbf{r}_i - \mathbf{R}|}\Psi \right) + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}\Psi = E\Psi$$

In Ashcroft and Mermin's *Solid State Physics* (p. 330) it says"*One has no hope of solving an equation such as [this]. Further progress requires some simplifying physical idea.*" Thus introduce one-electron orbitals and develop Hartree-Fock approximation.

In fact, to do calculations with many-body wave functions in a solid, just need *periodic boundary conditions*. Largest QMC calculations done so far: e.g. liquid and solid iron (1500-2000 electrons per cell) by Dario Alfè's group at UCL.

Walter Kohn's objection: "*..the many-electron wave function is not a legitimate scientific concept for more than about $N = 1000$ particles..*" [Nobel prize speech: RMP **71**, 1253 (1999)]

Oh dear! Says this because overlap of approximate Ψ with exact Ψ tends exponentially to zero as N increases, unless one uses number of parameters exponentially increasing with N. Not relevant to QMC - just need good description of relatively low-order correlation functions (e.g. PCFs) to get total energy correct. And Kohn's argument surely applies to determinant of approximate Kohn-Sham orbitals as well?

Good things about DMC

- ‘Chemical accuracy’ (1 kcal/mol or 0.04 eV) and beyond readily achievable. Can feed results of standard DFT/HF/etc.. calculations into a QMC code and greatly ‘improve’ the answer. Provides *reliable benchmarking* comparable to CCSD(T).
- Required computer time scales as N^3 with system size (improvable to N^2 with some extra complication). Very favourable compared to standard correlated wave function methods (up to N^7 for similar accuracy).
- QMC algorithm intrinsically parallel. With certain caveats speed of calculation increases linearly with number of processors (tested in practice to more than half a million cores). Can therefore fully exploit top-of-the-range supercomputers and other modern hardware including - in principle - GPUs. Most standard methods cannot exploit more than a few thousand processors..
- ‘Natural’ description of electron correlation. Uses arbitrarily complex many-electron Ψ with no need for analytic integrability. No BSSE or size consistency problems. Not overly dependent on basis set quality - Ψ not represented by basis set.
- Can calculate ground states, excited states, chemical reaction barriers and other properties within single unified methodological framework to high accuracy, based solely on the variational principle. Works for finite or periodic systems.

QMC is essentially the only highly accurate method whose cost can be made to scale reasonably with system size without significantly degrading the quality of the answer.

Why use QMC when we can use DFT?

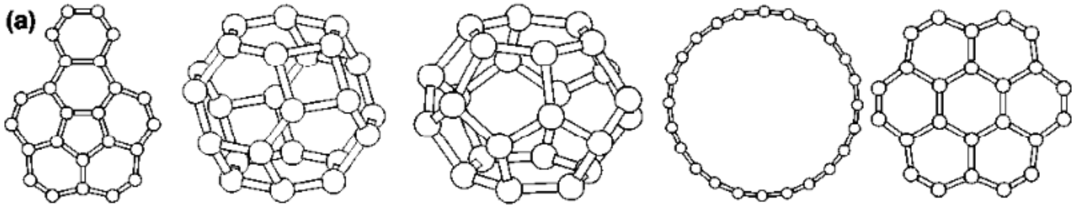
"Ψ is a high-dimensional object, dependent on $3N$ electron Cartesian coordinates, N electron spin coordinates, and $3N_n$ nuclear coordinates. For a molecule such as benzene, with 12 nuclei and 42 electrons, this object exists in $162 - 6 = 156$-dimensional Cartesian space! The determinant has $42! = 1.4 \times 10^{51}$ terms! We can't even make a picture of it. If we were to store it as a numerical object, with a resolution of 0.1 au out to 10 au, we would need 10^{312} numbers to store, at single precision (4 bytes/number), at 2 Gbytes/cm^2 (counting electrical connections access), we would need more than 10^{293} km^2 of surface to store that information. The Earth has a surface area of less than 10^9 km^2. The promised knowledge hidden in the Schrödinger equation is not quite easily accessible! We must make do with much much less. How much less can we do with?" [**from an online DFT course**]

DFT texts always begin by saying it is better to use the density - which depends on 3 variables - instead of Ψ - which depends on $3N$ variables - and thus we use DFT. Fine - who has 10^{284} planets after all? - but (a) in QMC we don't represent Ψ everywhere in configuration space - we just sample it, and (b) unfortunately the *equation* satisfied by the density is simply not known:

$$\boxed{E[\rho] = T[\rho] + \int V_{ext}(\mathbf{x})\rho(\mathbf{x})\,\mathrm{d}x + V_H[\rho] + ??E_{xc}[\rho]??}$$

Exchange-correlation functionals are *serious uncontrolled approximations*..

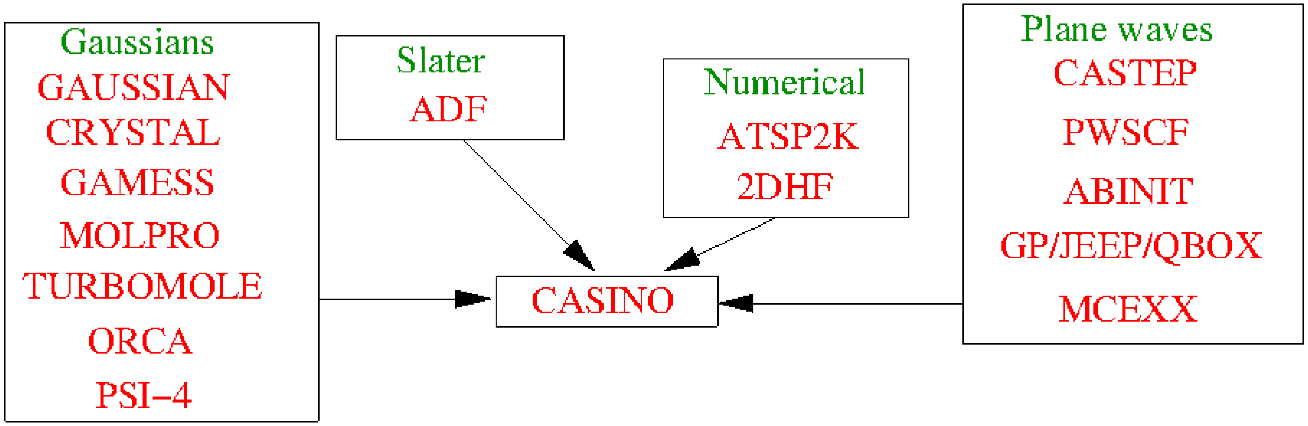
- There are large classes of problem for which DFT gives qualitatively the wrong answer (weak interactions, strongly-correlated systems, energy differences between different types of system, etc.)
- Huge dependence on XC functional. Answers not *reliable* i.e. even for functionals which give better results *on average*, there is no guarantee it gives better answers for any single system.
- Many DFT users are unaware of this, and consider their results to be the 'quantum' answer. Until recently - particularly in periodic systems - no benchmark method available to check accuracy.



Why use QMC instead of proper quantum chemistry methods?

- Usually takes serious computational effort to be accurate *enough* (e.g. the 'gold standard' CCSD(T) extrapolated to complete basis limit).
- *In principle* can solve the Schrödinger equation as accurately as you like - just do full CI with a complete basis set. This is not the point. The key point is *scaling of computer time with system size* - in the modern age N^7 will not do - as such calculations quickly become impossible. We need to scale at most as a low-order polynomial to do most problems of technological interest. QMC can do this.
- In general, you can do higher-order quantum chemistry techniques for atoms and molecules but not for *periodic* systems such as crystalline solids.
- QMC is not restricted to the regular paradigm of expanding the many-electron wave function in a basis set of differently occupied many-electron determinants constructed from orbitals expanded in analytically integrable Gaussian basis sets. If a mathematical expansion requires millions or billions of terms to converge, this is normally taken to mean that the basis set is very poor. Why should the quantum chemistry case be any different?
- Note that real wave functions have a cusp (gradient discontinuity) as an electron passes through a nuclear position - this is impossible to represent with a set of Gaussian functions all with zero gradient at the nucleus.
- Difficult to efficiently exploit many processors on large computers with QC methods.

What's wrong with QMC?



Gaussians
GAUSSIAN
CRYSTAL
GAMESS
MOLPRO
TURBOMOLE
ORCA
PSI–4

Slater
ADF

Numerical
ATSP2K
2DHF

Plane waves
CASTEP
PWSCF
ABINIT
GP/JEEP/QBOX
MCEXX

CASINO

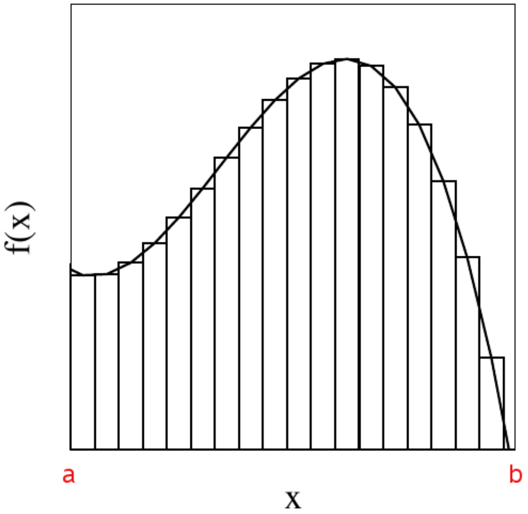
- Need to do a preparatory calculation with someone else's HF/DFT code, and your QMC software must have an explicit interface to that code. Often people find that their HF/DFT code of choice is not supported, so they give up before they start.
- DMC is computationally expensive. Even though the scaling with system size is good, there is a large pre-factor.. (c. 1000× slower than DFT).
- Difficult to compute forces and hence dynamics. Usually zero temperature.
- You will have no friends (both in the sense of being lonely - because to a first approximation no-one else does it - and also because everybody hates a smartarse..)

QMC is not being advocated as a *replacement* for any of these other techniques - which are all very useful in the spheres where they work; it should be considered as the *final building block* in our atomic scale micro-laboratory of multiple techniques that we use to simulate small parts of the real world.

So how does QMC work? Some technical details..

Monte Carlo integration

Alternative to traditional fixed-grid quadrature methods for evaluation of integrals, the main difference being that the sampling points are chosen *at random*.



- $I = \int_a^b f(x)\,dx = (b-a)\bar{f}$ with $\bar{f}$ mean value of $f(x)$.
- Unbiased estimate of $\bar{f}$ is $\langle f \rangle = \frac{1}{M}\sum_{i=1}^{M} f(x_i)$ with the x_i randomly selected from the interval $a \leq x_i \leq b$.
- Statistical uncertainty in $\langle f \rangle$ is given by $\sigma = \frac{\sigma_{\text{sample}}}{\sqrt{M}}$ with $\sigma_{\text{sample}} = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$. If M large enough, estimate of the mean $\bar{f}$ is *normally* distributed ('Central Limit Theorem').

- For Monte Carlo integration, error decreases as the square root of the number of sampling points (i.e. as $\frac{1}{\sqrt{M}}$) *irrespective of the dimensionality* d *of the integral.*
- For a standard grid method such as the trapezium rule the error decreases as $\mathcal{O}(M^{-\frac{2}{d}})$. Monte Carlo wins in more than four dimensions.
- To make the estimate of a 100-dimensional integral ten times more accurate requires 100 times more work with Monte Carlo integration. With the trapezium rule it would require 10^{50} times more work!
- For N particle system we must do $3N$-dimensional integrals. For high-dimensional numerical integration there is effectively *no alternative* to Monte Carlo methods.

How to sample things from probability distributions

In previous example, implicitly selected random points from *uniform probability distribution* (i.e. from anywhere in interval with equal probability). Not sensible for integral like $\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x$ - few if any points would lie in region where integrand is finite. Thus required to sample points from *non-uniform probability distributions*.

Properties of probability distributions

- $p(x)$ = probability density; $p(x)dx$ = probability to be in interval $(x, x+\mathrm{d}x)$.
- $p(x)$ must be positive and normalized to unity: $\int p(x)dx = 1$
- Average values $\langle f(x)\rangle = \int p(x)f(x)\mathrm{d}x$ with $\sigma^2 = \int p(x)f(x)^2\mathrm{d}x - \langle f(x)\rangle^2$.

To do Monte Carlo integration with a non-uniform distribution, generate set of M points x_i distributed according to desired $p(x)$. This can be done using a *random walk* moved according to the *Metropolis algorithm*. Then $\langle f(x)\rangle \simeq \frac{1}{M}\sum_i f(x_i)$ with $\sigma^2 \simeq \frac{1}{M}\sum_i f(x_i)^2 - \left(\frac{1}{M}\sum_i f(x_i)\right)^2$.

What is best $p(x)$ to use? The one that minimizes the variance: $p_{best}(x) = \frac{|f(x)|}{\int |f(x')|\,\mathrm{d}x'}$. Thus concentrate sampling points in regions where absolute value of integrand large - *importance sampling*. Don't in general know normalization though..

How does Metropolis work? Random walk moving from $\mathbf{r}$ to $\mathbf{r}'$ with prob $T(\mathbf{r} \longrightarrow \mathbf{r}')$ (e.g. Gaussian). Accept move with probability: $a(\mathbf{r} \longrightarrow \mathbf{r}') = \min\left[1, \frac{T(\mathbf{r}' \longrightarrow \mathbf{r})p(\mathbf{r}')}{T(\mathbf{r} \longrightarrow \mathbf{r}')p(\mathbf{r})}\right]$ i.e. occasionally reject moves to regions of lower probability. *Equilibration* required.

Many-electron wave functions

- Hartree product: $\Psi_{\mathcal{H}} = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\ldots\psi_N(\mathbf{x}_N)$
- Single determinant:

$$\Psi_{\mathcal{D}} = \frac{1}{\sqrt{N_e}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \ldots & \psi_1(\mathbf{x}_{N_e}) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & & \\ \vdots & & \ddots & \\ \psi_{N_e}(\mathbf{x}_1) & & & \psi_{N_e}(\mathbf{x}_{N_e}) \end{vmatrix}$$

- Single determinant Slater-Jastrow function: $\Psi_{\mathcal{DJ}} = \Psi_{\mathcal{D}}\Psi_{\mathcal{J}}$
- Multi-determinant Slater-Jastrow function: $\Psi_{n\mathcal{DJ}} = (\sum_i^n c_i \Psi_{\mathcal{D}_i})\Psi_{\mathcal{J}}$
- Multi- or single determinant Slater-Jastrow function with *backflow*

Jastrow factor $\Psi_{\mathcal{J}} = \exp(\mathcal{J})$ is optimizable functional form for pair correlation, e.g.:

$$\mathcal{J} = \sum_{i \neq j}^{N_e} \left[-U_0(r_{ij}) - U(r_{ij}) + \sum_n^{N_n} S^n(\mathbf{r}_i, \mathbf{r}_j, r_{ij}) \right] \text{ and } U_0(r_{ij}) = \frac{A}{r_{ij}} \left(1 - \exp\left(-\frac{r_{ij}}{F} \right) \right)$$

with U power series expansion in e-e separation and S^n set of atom-centred functions giving additional variational freedom in description of pair correlation near nuclei.

Backflow: replace coords $\mathbf{x}_i$ in $\Psi_{\mathcal{D}}$ with *collective coords* $\mathbf{r}_i(\{\mathbf{x}_j\})$ given by $\mathbf{r}_i = \mathbf{x}_i + \xi_i(\{\mathbf{x}_j\})$ with ξ_i configuration-dependent *backflow displacement* of particle i.

Pair correlation function

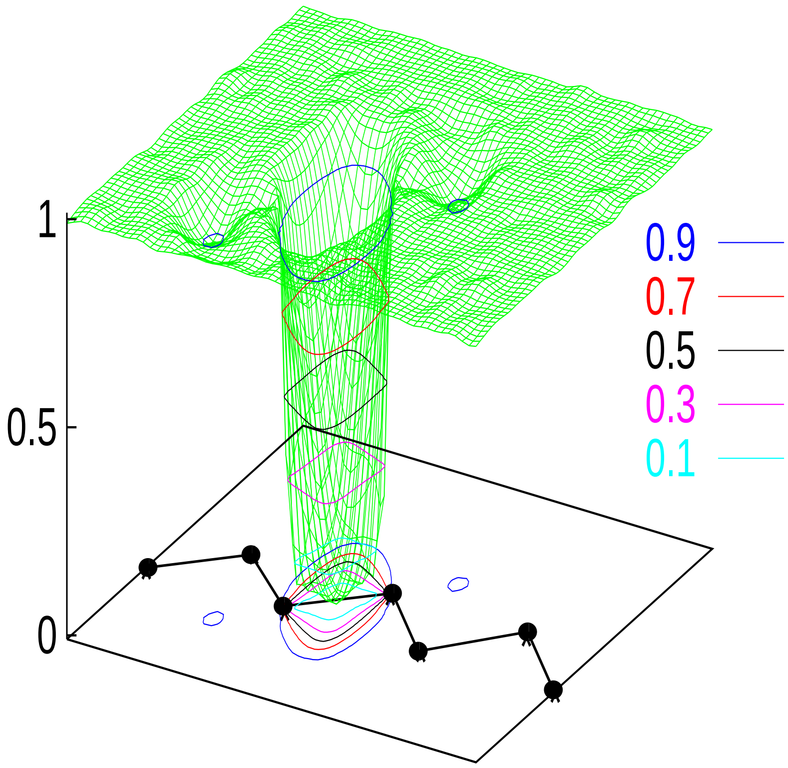
Silicon with Slater-Jastrow wave function

VMC

$\bar{g}_{\uparrow\uparrow}(\mathbf{r},\mathbf{r}';[n])$

$\mathbf{r}$ at bond center

$\mathbf{r}'$ in (110) plane

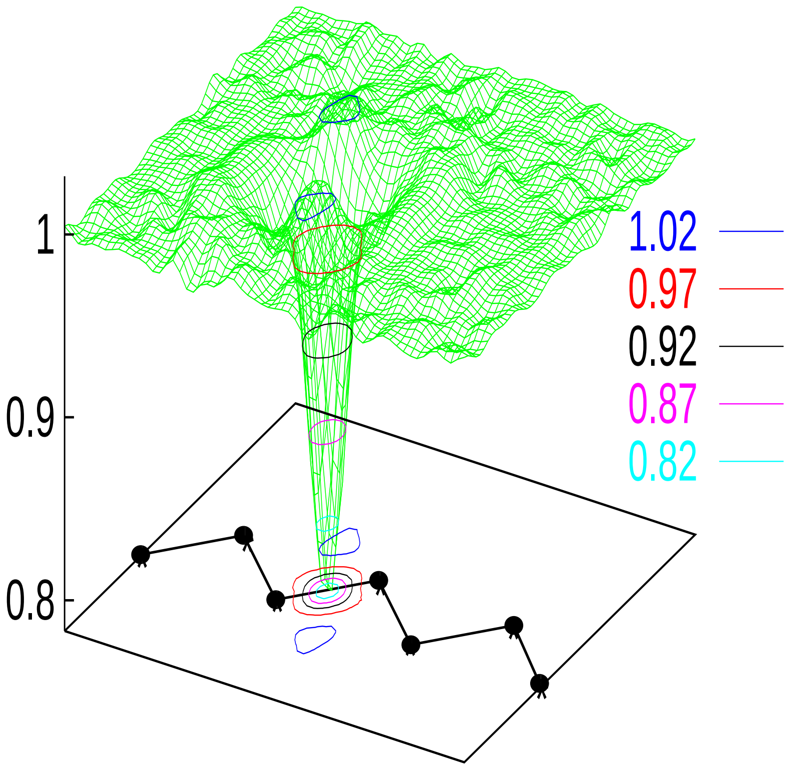


VMC

$\bar{g}_{\uparrow\downarrow}(\mathbf{r},\mathbf{r}';[n])$

$\mathbf{r}$ at bond center

$\mathbf{r}'$ in (110) plane

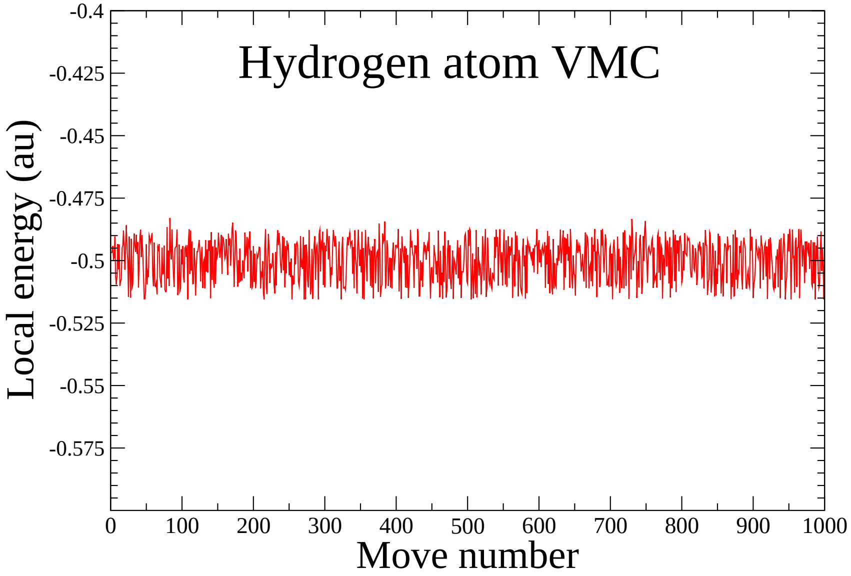


Variational Monte Carlo

Stochastic integration method for evaluating expectation values for fixed trial many-body wave function $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N) \equiv \Psi(\mathbf{x})$:

$$E_{\mathrm{VMC}} = \frac{\int \Psi^* \hat{H} \Psi \, \mathrm{d}\mathbf{x}}{\int |\Psi|^2 \, \mathrm{d}\mathbf{x}} = \frac{\int |\Psi|^2 \left(\frac{\hat{H}\Psi}{\Psi}\right) \mathrm{d}\mathbf{x}}{\int |\Psi|^2 \, \mathrm{d}\mathbf{x}} \left(\equiv \int p(x) f(x) \, \mathrm{d}x \right)$$

where $f(x) = \frac{\hat{H}\Psi}{\Psi} \equiv E_{\mathrm{L}} =$ 'local energy' and $p(x) = |\Psi|^2 / \int |\Psi|^2 \, \mathrm{d}\mathbf{x}$.



- Generate M points distributed as $|\Psi|^2$ and average local energies $E_{\mathrm{VMC}} \simeq \frac{1}{M} \sum_i E_{\mathrm{L}}(\mathbf{x}_i)$.
- Optimal probability density $p(x) = |f(x)| / \int |f(x')| \, \mathrm{d}x' = |\Psi^* \hat{H} \Psi| / \int \Psi^* \hat{H} \Psi \, \mathrm{d}\mathbf{x}$ but this tends to $|\Psi|^2 / \int |\Psi|^2 \, \mathrm{d}\mathbf{x}$ as Ψ tends to exact wave function when $\hat{H}\Psi = E\Psi$.
- As Ψ tends to exact wave function, fluctuations in E_{L} tend to zero (the '*zero variance principle*').

Effectively still a *basis set method* - you get out what you put in.

Wave function optimization

Trial wave functions contain parameters (Jastrow, backflow, det coeffs etc.) that must be *optimized* in VMC so Ψ has optimal functional form. Need *objective function* to minimize with respect to parameter set $\{\alpha\}$ - generally choose *energy* or *variance*:

$$E_{\mathrm{V}} = \frac{\int \Psi^2(\alpha) E_{\mathrm{L}}(\alpha)\,\mathrm{d}\mathbf{x}}{\int \Psi^2(\alpha)\,\mathrm{d}\mathbf{x}} \quad \text{or} \quad \mathrm{var}(E) = \frac{\int \Psi^2(\alpha)[E_{\mathrm{L}}(\alpha) - E_{\mathrm{V}}(\alpha)]^2\,\mathrm{d}\mathbf{x}}{\int \Psi^2(\alpha)\,\mathrm{d}\mathbf{x}}$$

Until recently more usual to minimize variance, because:

- It has a known lower bound of zero.
- It can be applied to excited states.
- Efficient algorithms known for minimizing objective functions expressible as sum of squares (for many years energy minimization was considered to be a difficult numerical problem).
- Varmin-optimized trial-functions exhibits greater numerical stability than energy minimized ones in DMC (there is a smaller tendency for population explosions).

Still default choice if want DMC energies with Slater-Jastrow trial function.

Now also common to minimize energy, because:

- Efficient, numerically-stable algorithms for doing so have been discovered.
- Since trial wave functions cannot represent eigenstate exactly, the energy and variance minima don't coincide. Energy minimization should therefore produce lower VMC energies (and consequently better DMC wave functions - see later).
- Better at optimizing parameters that change *nodal surface* - significance will become clear shortly.
- Resulting wave functions give better forces (and some other properties).

Practical optimization

Rewrite energy and variance for correlated sampling

$$E_{\mathrm{V}} = \frac{\int \Phi^2(\alpha_0)\ w(\alpha)\ E_{\mathrm{L}}(\alpha)\, d\mathbf{R}}{\int \Phi^2(\alpha_0)\ w(\alpha)\, d\mathbf{R}}$$

$$\mathrm{Var}(E) = \frac{\int \Phi^2(\alpha_0)\ w(\alpha)\ [E_{\mathrm{L}}(\alpha) - E_{\mathrm{V}}(\alpha)]^2\, d\mathbf{R}}{\int \Phi^2(\alpha_0)\ w(\alpha)\, d\mathbf{R}}$$

$$w(\alpha) = \frac{\Phi^2(\alpha)}{\Phi^2(\alpha_0)} \quad \text{Weight}$$

Average over M configurations drawn from $\Phi^2(\alpha_0)$

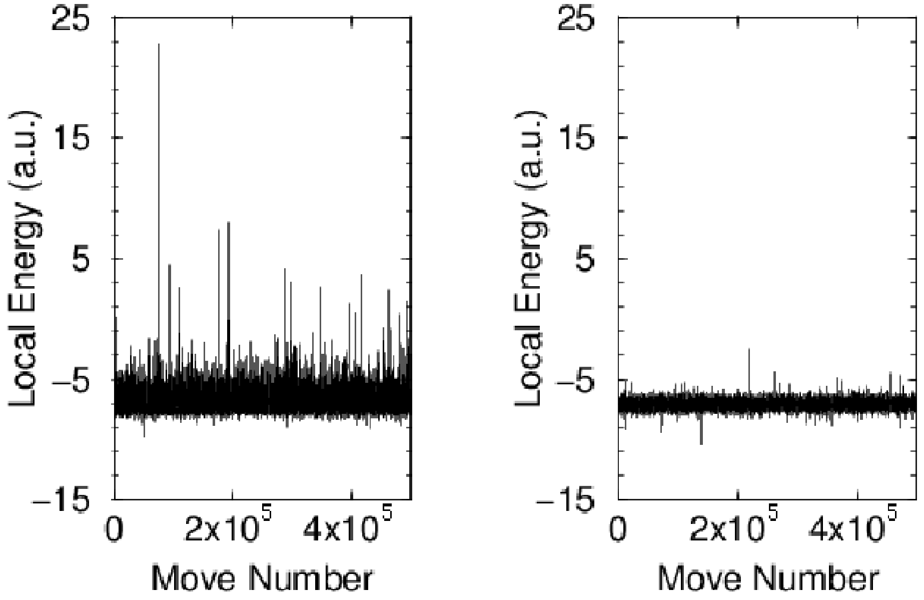
$$E_{\mathrm{V}} \simeq \frac{\sum_i^M w(\mathbf{R}_i;\alpha) E_{\mathrm{L}}(\mathbf{R}_i;\alpha)}{\sum_i^M w(\mathbf{R}_i;\alpha)}$$

$$\mathrm{Var}(E) \simeq \frac{\sum_i^M w(\mathbf{R}_i;\alpha)[E_{\mathrm{L}}(\mathbf{R}_i;\alpha) - E_{\mathrm{V}}(\{\mathbf{R}_i\};\alpha)]^2}{\sum_i^M w(\mathbf{R}_i;\alpha)}$$

- Eigenstates of $\hat{H}$ give zero variance for *any* set of configurations.
- Eigenstates of $\hat{H}$ give zero variance for *any* set of (positive) weights.
- In practice, *unreweighted minimization* is normally used! See later.

Results of a VMC simulation

No Jastrow



Optimized Jastrow

- Answer given as $E_{\mathrm{VMC}} \pm \sigma$, where σ is some statistical error bar.
- Better Ψ give lower energy (variational principle) and smaller error bar.
- Statistical analysis required to eliminate serial correlation and get correct error bar.

Example results: cohesive energies

Method	Si	Ge	C	BN
LSDA	5.28	4.59	8.61	15.07
VMC	4.48±0.01	3.80±0.02	7.36±0.01	12.85±0.03
Exp.	4.62	3.85	7.37	12.9

Not bad but not brilliant either..

VMC is not enough

VMC results are OK (might recover $75-85\%$ *of correlation energy) - not really good enough to be worth the effort. Results highly dependent on trial wave function and basis set quality. Problem is one of* **complexity** *- not clear how to make arbitrarily accurate wave functions this way.*

However, there is a class of methods - collectively called *projector Monte Carlo* - which in principle solve quantum problems exactly. Idea is to project out ground state by repeated iteration of a projection operator. After many iterations the excited state $|\Psi_i\rangle$ contributions are filtered out and the ground state is recovered.

Such methods depend upon guessed properties of many-body Ψ only in computational efficiency, if at all. *Diffusion Monte Carlo* (DMC) is one such method.

Characteristics of diffusion Monte Carlo

- Systematically improves Ψ through magic process: *stochastic propagation in imaginary time*.
- Extremely weakly dependent on basis set quality.
- Scales as N^3 or better with system size.
- Taken to limit, comparable in accuracy with benchmark quantum chemistry correlated methods.

The main use of VMC in practical calculations is to prepare the input for a diffusion Monte Carlo simulation.

Projector methods

First suggestion of Monte Carlo solution of Schrödinger equation dates back to Fermi. Showed that a solution of the stationary state equation,

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R})$$

could be obtained by introducing a time-dependent wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This will obey the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \frac{1}{2}\nabla^2\Psi(\mathbf{R},\tau) - V(\mathbf{R})\Psi(\mathbf{R},\tau)$$

and in the long time limit, look like the ground state solution to the first equation.

Projector methods

First suggestion of Monte Carlo solution of Schrödinger equation dates back to Fermi. Showed that a solution of the stationary state equation,

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R})$$

could be obtained by introducing a time-dependent wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This will obey the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \frac{1}{2}\nabla^2\Psi(\mathbf{R},\tau) - V(\mathbf{R})\Psi(\mathbf{R},\tau)$$

and in the long time limit, look like the ground state solution to the first equation.

- Ignore 2nd term on RHS, equation is isomorphic with a diffusion equation, which can be simulated by a random walk where random walkers diffuse in an $\mathbf{R}$-dimensional space.

Projector methods

First suggestion of Monte Carlo solution of Schrödinger equation dates back to Fermi. Showed that a solution of the stationary state equation,

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R})$$

could be obtained by introducing a time-dependent wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This will obey the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \frac{1}{2}\nabla^2\Psi(\mathbf{R},\tau)$$

and in the long time limit, look like the ground state solution to the first equation.

- Ignore 2nd term on RHS, equation is isomorphic with a diffusion equation, which can be simulated by a random walk where random walkers diffuse in an $\mathbf{R}$-dimensional space.

Projector methods

First suggestion of Monte Carlo solution of Schrödinger equation dates back to Fermi. Showed that a solution of the stationary state equation,

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R})$$

could be obtained by introducing a time-dependent wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This will obey the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \qquad\qquad\qquad\qquad - V(\mathbf{R})\Psi(\mathbf{R},\tau)$$

and in the long time limit, look like the ground state solution to the first equation.

- Ignore 2nd term on RHS, equation is isomorphic with a diffusion equation, which can be simulated by a random walk where random walkers diffuse in an $\mathbf{R}$-dimensional space.
- Ignore 1st term, have first order kinetics equation with position-dependent rate constant $V(\mathbf{R})$ (interpret as stochastic survival probability).

Projector methods

First suggestion of Monte Carlo solution of Schrödinger equation dates back to Fermi. Showed that a solution of the stationary state equation,

$$-\frac{1}{2}\nabla^2\Psi(\mathbf{R}) = E\Psi(\mathbf{R}) - V(\mathbf{R})\Psi(\mathbf{R})$$

could be obtained by introducing a time-dependent wave function of the form $\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R})e^{-E\tau}$. This will obey the equation

$$\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = \frac{1}{2}\nabla^2\Psi(\mathbf{R},\tau) - V(\mathbf{R})\Psi(\mathbf{R},\tau)$$

and in the long time limit, look like the ground state solution to the first equation.

- Ignore 2nd term on RHS, equation is isomorphic with a diffusion equation, which can be simulated by a random walk where random walkers diffuse in an $\mathbf{R}$-dimensional space.
- Ignore 1st term, have first order kinetics equation with position-dependent rate constant $V(\mathbf{R})$ (interpret as stochastic survival probability).

A numerical simulation in which random walkers diffuse through $\mathbf{R}$*-space, reproduce in regions of low potential, and die in regions of high potential leads to a stationary distribution proportional to* $\Psi(\mathbf{R})$*, from which expectation values can be obtained.*

Projector methods

General idea is to project out a state of the Hamiltonian by iteration of a projection operator $\hat{P}$. Assume ground state for simplicity.

$$\lim_{i \longrightarrow \infty} \hat{P}^i |\Psi_T\rangle \approx |\Psi_0\rangle$$

After sufficient iterations i, the contribution of all excited states $|\Psi_i\rangle$ will be filtered out and only the ground state is recovered.

If Ψ_T is a vector and P is a matrix then the procedure implied by this is just a standard piece of matrix algebra called the power method.

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$q(0) = (1, 0, 0)$

$\lambda(A) = (10, 4, 3)$

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191

Power method

•Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$

•$X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors

•$q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

•Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX =$ diag $(\lambda_1, \lambda_2, \dots, \lambda_n)$
- $X = [x_1, x_2, \dots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \dots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633
5	10.0198

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633
5	10.0198
6	10.0063

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = $[q(k)]^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633
5	10.0198
6	10.0063
7	10.0020

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = $[q(k)]^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633
5	10.0198
6	10.0063
7	10.0020
8	10.0007

Power method

- Diagonalizable $n \times n$ matrix A i.e. $X^{-1}AX = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$
- $X = [x_1, x_2, \ldots, x_n]$ matrix of eigenvectors
- $q = (1, 0, 0, 0, \ldots)$ unit norm vector

```
do k = 1,2,...
z(k) = matmul[A,q(k-1)]
q(k) = z(k) / norm[z(k)]
$\lambda$(k) = [q(k)]$^{-1}$ A q(k)
enddo
```

$$A = \begin{pmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{pmatrix}$$

$$q(0) = (1, 0, 0)$$

$$\lambda(A) = (10, 4, 3)$$

- Only dominant eigenvector survives after large number of iterations.

k	λ(k)
1	13.0606
2	10.7191
3	10.2703
4	10.0633
5	10.0198
6	10.0063
7	10.0020
8	10.0007
9	10.0002

The concept of a propagator

Let's say we wish to integrate the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial\Psi(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(\mathbf{x},t) + V(\mathbf{x},t)\Psi(\mathbf{x},t) = \hat{H}\Psi(\mathbf{x},t)$$

where $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. Usually use atomic units: $\hbar = 1, m = 1$.

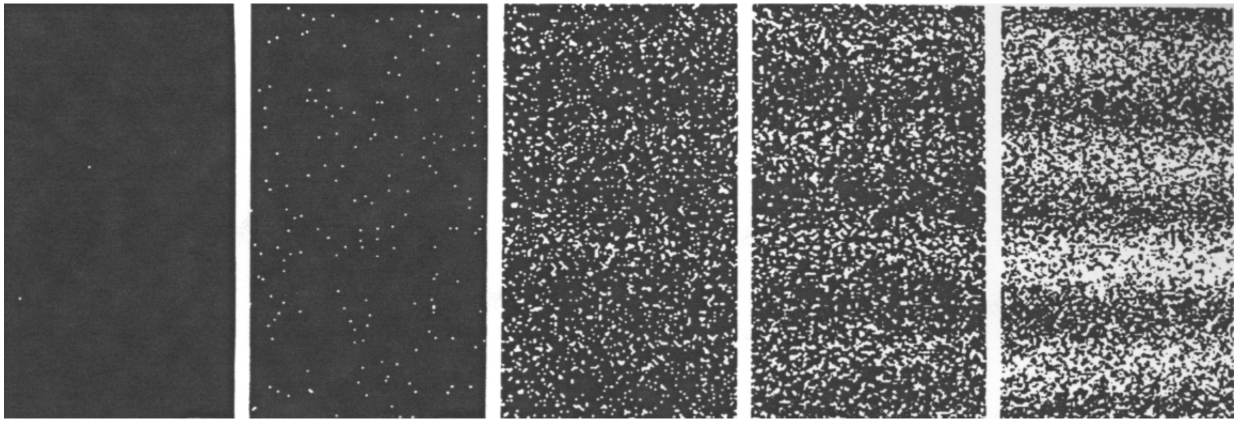
- Inverse of this differential equation is an integral equation involving the *propagator*:

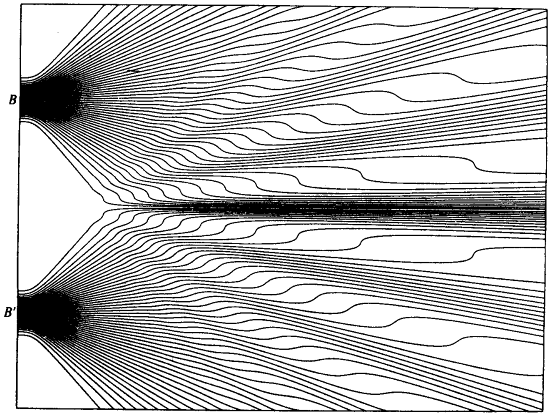
$$\Psi(\mathbf{x},t) = \int K(\mathbf{x},t;\mathbf{x}',t')\Psi(\mathbf{x}',t')\,\mathrm{d}x'$$

The propagator is the *probability amplitude for a particle to travel from one place to another in a given time*. It is a *Green's function* for the Schrödinger equation.

- Ordinarily one might use the *Feynman path-integral* formulation of quantum mechanics to calculate the propagator. For given start and end points, overall amplitude given by summing contributions of infinite number of all possible *histories* which include those points. Amplitude contributed by a particular history proportional to $e^{iS_{cl}/\hbar}$ where S_{cl} is classical *action* of that history i.e. time integral of classical Lagrangian $T - V$ along corresponding phase space path of system.

Pilot wave theory





Otherwise known as *de Broglie-Bohm theory* or *Bohmian mechanics*, this is an interpretation of QM where particles exist and have *trajectories* (most people believe not possible!). Invented in **1927** by de Broglie, but unfairly overruled by *logical positivists* of Copenhagen school (Bohr, Heisenberg, etc.).

- Particles *guided* by wave along streamlines of probability flow, following trajectory $\mathbf{v} = \nabla S/m$ where S is *phase* of Schrödinger wave function written in polar form $\Psi(x,t) = R(x,t)e^{iS(x,t)/\hbar}$.
- Whole thing follows from a single semantic change in the meaning of a word: $|\Psi|^2$ is *probability of **being** at* x rather than *probability of being* ***found there in a suitable measurement***. Under this assumption, measurement problem and all quantum 'paradoxes' simply disappear.

Quantum trajectory method

Pilot-wave theory not just interpretation: mathematical formulation of QM useful to compute propagator! Substitute Ψ polar form in Schrödinger equation. Separate real and imaginary parts. Get equations of *quantum hydrodynamics* - a 'quantum trajectory method' (analogy with fluid mechanics):

$\frac{\partial \rho}{\partial t} = -\rho \nabla \cdot \mathbf{v}$ (continuity equation, where $\rho = R^2$) and $\frac{\partial S}{\partial t} = L(t) = \frac{1}{2}mv^2 - (V + Q)$ (the quantum Hamilton-Jacobi equation - an equation for the phase involving the *quantum potential* Q).

Wave function synthesis along quantum trajectory

$$\boxed{\Psi = R(x,t)\mathrm{e}^{\frac{iS(\mathbf{x},t)}{\hbar}}}$$

Propagation of the amplitude

Along trajectory $\mathbf{x}(t)$ from $(\mathbf{x}_0, t_0)$ to $(\mathbf{x}_1, t_1)$ rate of change of density $\frac{\partial\rho}{\partial t} = -\rho\nabla\cdot\mathbf{v}$. Amplitude $R = \rho^{\frac{1}{2}}$ so $\frac{\partial R}{\partial t} = -\frac{R}{2}\nabla\cdot\mathbf{v}$. Integrate to get new R in terms of value at t_0.

$$\frac{\mathrm{d}R}{R} = -\frac{1}{2}\nabla\cdot\mathbf{v}\;\mathrm{d}t \quad \xrightarrow{\text{integrate}} \quad \ln R + c = -\frac{1}{2}\int_{t_0}^{t_1}\nabla\cdot\mathbf{v}\;\mathrm{d}t \quad \xrightarrow{\text{exponentiate}} \quad A\exp(\ln R) = \exp\left[-\frac{1}{2}\int_{t_0}^{t_1}\nabla\cdot\mathbf{v}\;\mathrm{d}t\right]$$

$$R(\mathbf{x}_1, t_1) = \exp\left[-\frac{1}{2}\int_{t_0}^{t_1}(\nabla\cdot\mathbf{v})_{\mathbf{x}(t)}\;\mathrm{d}t\right]R(\mathbf{x}_0, t_0)$$

To propagate R we integrate the *divergence of the velocity field* along the trajectory.

Propagation of the exponential of the phase

Quantum H-J eqn is $\frac{\partial S}{\partial t} = L(t) = \frac{1}{2}mv^2 - (V+Q)$. From $S(t_1) = S(t_0) + \int_{t_0}^{t_1}\frac{\partial S}{\partial t}\;\mathrm{d}t$ $\xrightarrow{\times\frac{i}{\hbar}\text{ and exponentiate}}$

$$\mathrm{e}^{\frac{iS(\mathbf{x}_1,t_1)}{\hbar}} = \exp\left[\frac{i}{\hbar}\int_{t_0}^{t_1}L(t)\;\mathrm{d}t\right]\mathrm{e}^{\frac{iS(\mathbf{x}_0,t_0)}{\hbar}}$$

To propagate $e^{\frac{iS(\mathbf{x},t)}{\hbar}}$ we integrate the *quantum Lagrangian* along the trajectory.

Wave function synthesis along quantum trajectory

We multiply the expressions for the R propagator and the $e^{\frac{iS(\mathbf{x},t)}{\hbar}}$ propagator to obtain an expression for updating the full wave function along the trajectory:

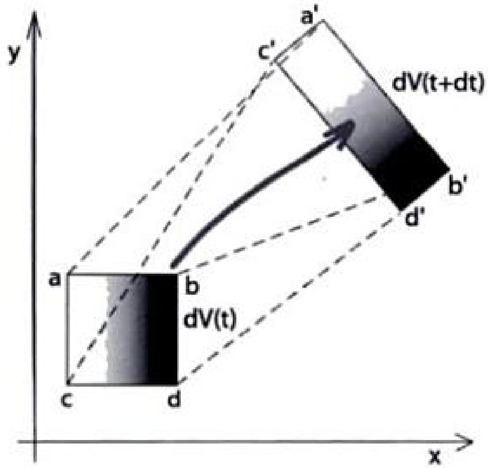
$$\boxed{\Psi(\mathbf{x}_1,t_1) = \exp\left[-\frac{1}{2}\int_{t_0}^{t_1}(\nabla\cdot\mathbf{v})_{\mathbf{x}(t)}\,\mathrm{d}t\right]\exp\left[\frac{i}{\hbar}\int_{t_0}^{t_1}L(t)\,\mathrm{d}t\right]\Psi(\mathbf{x}_0,t_0)}$$

Hydrodynamical wave function propagator (HWFP) $K^Q(\mathbf{x}_1,t_1;\mathbf{x}_0,t_0)$

What does $\nabla\cdot\mathbf{v}$ mean here?

At time t have volume element $\mathrm{d}V(t)$. Element corners defined by trajectory positions $\{a,b,c,d\}$. Increment time by $\mathrm{d}t$ and equations of motion shift corners to $\{a',b',c',d'\}$ and volume element changes to $\mathrm{d}V(t+\mathrm{d}t)$. Ratio of new to old volumes is the *Jacobian*: $\mathrm{d}V(t+\mathrm{d}t) = J(t+\mathrm{d}t,t)\mathrm{d}V(t)$. Can be shown Jacobian is:

$$J(t_1,t_0) = \exp\left[\int_{t_0}^{t_1}\nabla\cdot\mathbf{v}\,\mathrm{d}t\right].$$



- Implies if velocity field has positive divergence (velocity vectors 'point away from each other') then Jacobian increasing and local volume element *expanding* along flow. So velocity divergence locally measures rate of change of geometric quantity.
- Note the R-propagator $\exp\left[-\frac{1}{2}\int_{t_0}^{t_1}\nabla\cdot\mathbf{v}\,\mathrm{d}t\right]$ is thus just $J(t)^{-\frac{1}{2}}$.

An interesting comparison

Quantum trajectories and Feynman path integrals

- In the expression $\Psi(\mathbf{x}_1,t_1) = K^Q(\mathbf{x}_1,t_1;\mathbf{x}_0,t_0)\Psi(\mathbf{x}_0,t_0)$ that propagates the wave function along the quantum trajectory, the propagator K^Q may - expressing the R-propagator in terms of the Jacobian - be written as

$$K^Q(\mathbf{x}_1,t_1;\mathbf{x}_0,t_0) = \frac{1}{J(t)^{\frac{1}{2}}} \exp\left[\frac{i}{\hbar}\int_{t_0}^{t_1} L(t)\,\mathrm{d}t\right].$$

- In *Feynman's path integral formulation of quantum mechanics* the equivalent propagator may be written as

$$K^F(\mathbf{x}_1,t_1;\mathbf{x}_0,t_0) = N \sum_{\text{all paths}} \exp\left[\frac{i}{\hbar}\int_{t_0}^{t_1} L_{cl}(t)\,\mathrm{d}t\right].$$

  Here propagator linking two spacetime points calculated by linearly superposing amplitudes $e^{iS/\hbar}$ (obtained by integrating **classical** Lagrangian $L_{cl}(t) = \frac{1}{2}mv^2 - V$) associated with infinite number of **all possible paths** connecting the points.

In pilot-wave approach, achieve same effect by integrating the **quantum** *Lagrangian* $L(t) = \frac{1}{2}mv^2 - (V+Q)$*along precisely* **one** *path. Bet you didn't know that..*

Quantum trajectories and quantum Monte Carlo

What connection, if any, is there between quantum trajectory methods and state-of-the-art techniques like **quantum Monte Carlo** *that accurately solve the time-independent Schrödinger equation sampling the full many-electron* Ψ*?*

Diffusion Monte Carlo

The most highly-evolved QMC variant with broad scope is *diffusion Monte Carlo* (DMC). It is probably the most accurate method known for solving the many-electron Schrödinger equation that also scales reasonably with the number of particles. It remains tractable (and highly accurate) for large system sizes; simulations of periodic systems with over 2000 electrons per cell have been reported.

It propagates an arbitrary starting wave function using a (Green's function) propagator just like the ones we have been discussing. The main difference is that the propagation occurs in *imaginary time* τ as opposed to real time t. This has the effect of 'improving' the wave function i.e. making it look more like the ground state as imaginary time passes (see later).

$$\Psi(\mathbf{x},\tau+\delta\tau)=\int K(\mathbf{x},\mathbf{x}',\delta\tau)\Psi(\mathbf{x}',\tau)\,d\mathbf{x}'$$

Evolving wave function represented by distribution in space and time of randomly-diffusing electron positions over an ensemble of copies of the system. From pilot-wave perspective, this is something like calculating expectation values by 'sampling trajectories' (from ensemble of different launch points).

Further reading

Quantum Monte Carlo simulations of solids, W.M.C. Foulkes, L. Mitas, R.J. Needs and G. Rajagopal, Rev. Mod. Phys **73**, 33 (2001).
The quantum Monte Carlo method, M.D. Towler (cough), Phys. Stat. Solidi **243**, 2573 (2006).

Why do we propagate Ψ in imaginary time in DMC?

Consider Schrödinger equation (in a.u.) with constant offset E_T to zero of potential:

$$-\frac{\partial\Psi(\mathbf{x},t)}{i\partial t} = \left(\hat{H} - E_T\right)\Psi(\mathbf{x},t).$$

For eigenstate, general solution is clearly

$$\phi(\mathbf{x},t) = \phi(\mathbf{x},0)\mathrm{e}^{-i(\hat{H}-E_T)t}.$$

Then expand an *arbitrary* ('guessed') $\Psi(\mathbf{x},t)$ in complete set of eigenfunctions of $\hat{H}$.

$$\Psi(\mathbf{x},t) = \sum_{n=0}^{\infty} c_n\phi_n(\mathbf{x})\mathrm{e}^{-i(E_n-E_T)t}$$

Substitute it with *imaginary time* $\tau = it$. Oscillatory behaviour becomes exponential.

$$\Psi(\mathbf{x},\tau) = \sum_{n=0}^{\infty} c_n\phi_n(\mathbf{x})\mathrm{e}^{-(E_n-E_T)\tau}$$

Get imaginary time independence by choosing constant E_T to be ground state eigenvalue E_0. As $\tau \to \infty$, Ψ comes to look more and more like the ground state ϕ_0.

$$\Psi(\mathbf{x},\tau) = c_0\phi_0 + \sum_{n=1}^{\infty} c_n\phi_n(\mathbf{x})\mathrm{e}^{-(E_n-E_0)\tau}$$

How do we propagate Ψ in imaginary time in DMC?

- We use a Green's function propagator $K(\mathbf{x},\mathbf{x}',\delta\tau)$:

$$\Psi(\mathbf{x},\tau+\delta\tau) = \int K(\mathbf{x},\mathbf{x}',\delta\tau)\Psi(\mathbf{x}',\tau)\,d\mathbf{x}'$$

- How do we find an expression for the propagator K? Consider imaginary-time Schrödinger equation in two parts:

$$\begin{aligned} \frac{\partial\Psi}{\partial\tau} &= \frac{1}{2}\nabla^2_{\mathbf{x}}\Psi \quad \text{(diffusion equation)} \\ \frac{\partial\Psi}{\partial\tau} &= -V\Psi \quad \text{(rate equation)} \end{aligned}$$

- Propagator for diffusion equation well-known: it is a $3N$-dimensional Gaussian with variance $\delta\tau$ in each dimension. Propagator for rate equation known - gives so-called 'branching factor' which can be interpreted as a position-dependent *weight* or *stochastic survival probability* for a member of an ensemble.

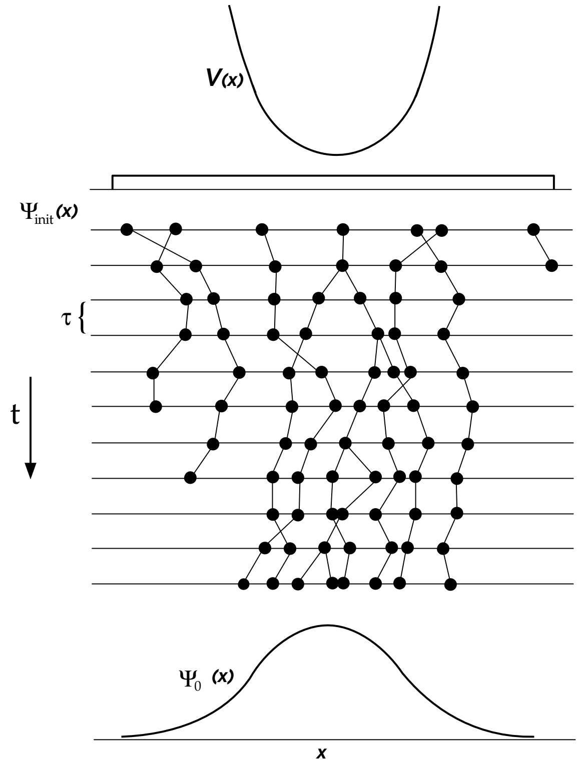
$$K(\mathbf{x},\mathbf{x}',\delta\tau) = \frac{1}{(2\pi\delta\tau)^{\frac{3N}{2}}}\exp\left(-\frac{|\mathbf{x}-\mathbf{x}'|^2}{2\delta\tau}\right) \times \exp\left[-\delta\tau\left(\frac{V(\mathbf{x})+V(\mathbf{x}')-2E_T}{2}\right)\right]$$

Multiplying the two together to get the propagator for the imaginary-time Schrödinger equation is an *approximation* valid in the limit of small $\delta\tau$.

A Diffusion Monte Carlo simulation

- Interpret Ψ as a *probability density*, then diffusion equation $\frac{\partial\Psi}{\partial\tau} = \frac{1}{2}\nabla^2_{\mathbf{x}}\Psi$ represents movement of N diffusing particles. Turning this around, can represent $\Psi(\mathbf{x}, \tau)$ by *ensemble* of such sets of particles. Member of ensemble called a ‘configuration’.
- Interpret propagator $K(\mathbf{x}, \mathbf{x}', \delta\tau)$ as probability of configuration moving from $\mathbf{x}'$ to $\mathbf{x}$ in a time $\delta\tau$. Branching factor determines population of configurations: in regions of high V configurations will be killed off; in low V regions configurations will multiply. It is this that ‘changes the shape of the wave function’ as it evolves.
- Propagate distribution in imaginary time, and after sufficiently long time excited states will have decayed away to leave the ground-state Ψ. Can then continue propagation and accumulate averages of observables.

Guess that the ground-state wave function for a single electron in a harmonic potential well is a constant over some range (stupid!). Start with seven copies of the system over which ensemble the electrons are distributed according to this constant probability distribution. Propagate the particle distribution in imaginary time according to the prescription above, and watch the electrons become distributed according to the proper Gaussian shape of the exact ground-state wave function. The change in shape is produced by the branching factor occasionally eliminating configurations in high V regions and duplicating ones in low V regions..

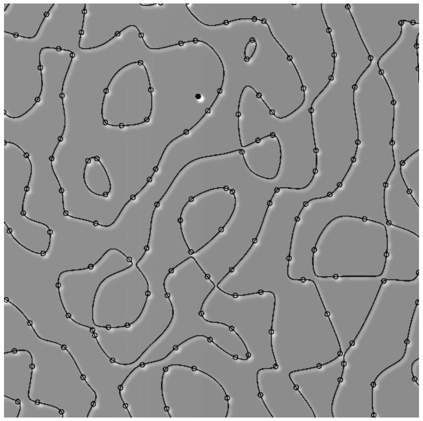


Importance sampling and the fixed-node approximation in DMC

The basic DMC algorithm sounds nice but *doesn't work* in practice. This is because:

- Ψ can only be a probability distribution if of one sign everywhere (e.g. H atom, boson system). For multi-particle fermion systems it necessarily has positive and negative bits. Obvious ways of getting round this (e.g. separate probabilities for different signs) fail on account of signal-to-noise problems ('*fermion sign problem*').
- Branching factor $\exp\left[-\frac{\delta\tau}{2}\left(V(\mathbf{x})+V(\mathbf{x}')-2E_{\mathrm{T}}\right)\right]$ contains potential V varying from $-\infty$ to $+\infty$. Sampling of points e.g. near nucleus leads to massive fluctuations in branching factor and significant numerical instabilities.

Fix problem by introducing *importance sampling* via a guessed trial function Φ_T (from a HF or DFT calculation, say). Require propagation to produce distribution $f(\mathbf{x},\tau)=\Psi(\mathbf{x},\tau)\Phi_T(\mathbf{x})$ - forced to be of one sign by demanding fixed Φ_T and variable Ψ functions have same *nodal surface* i.e. same zeroes ('*fixed-node approximation*').



New 'imaginary time Schrödinger equation' (Fokker-Planck)

$$-\frac{\partial f(\mathbf{x},\tau)}{\partial\tau}=-\frac{1}{2}\nabla_{\mathbf{x}}^2 f(\mathbf{x},\tau)+\nabla_{\mathbf{x}}\cdot\left[\mathbf{F}(\mathbf{x})f(\mathbf{x},\tau)\right]-\left(E_{\mathrm{L}}(\mathbf{x})-E_{\mathrm{T}}\right)f(\mathbf{x},\tau)$$

$\mathbf{F}(\mathbf{x})\equiv\frac{\nabla_{\mathbf{x}}\Phi_T}{\Phi_T}$ 'drift vector' $E_L=\frac{\hat{H}\Psi}{\Psi}$ 'local energy'

$$f(\mathbf{x},\tau+\delta\tau)=\int K(\mathbf{x}',\mathbf{x},\delta\tau)f(\mathbf{x},\tau)\,d\mathbf{x}'$$

Problem solved - more or less

Final propagator consists of diffusion, drift and branching processes:

$$K(\mathbf{x}',\mathbf{x},\delta\tau)=\frac{1}{(2\pi\delta\tau)^{\frac{3N}{2}}}\exp\left[-\frac{(\mathbf{x}'-\mathbf{x}-\delta\tau\mathbf{F}(\mathbf{x}))^2}{2\delta\tau}\right]\exp\left[-\frac{\delta\tau}{2}\left(E_{\mathrm{L}}(\mathbf{x})+E_{\mathrm{L}}(\mathbf{x}')-2E_{\mathrm{T}}\right)\right]$$

To be compared with the original propagator without importance sampling:

$$K(\mathbf{x},\mathbf{x}',\delta\tau)=\frac{1}{(2\pi\delta\tau)^{\frac{3N}{2}}}\exp\left(-\frac{(\mathbf{x}-\mathbf{x}')^2}{2\delta\tau}\right)\exp\left[-\frac{\delta\tau}{2}\left(V(\mathbf{x})+V(\mathbf{x}')-2E_T\right)\right]$$

- Mixed distribution $f=\Psi\Phi_T$ of one sign everywhere so no more sign problem (at cost of reducing flexibility by fixing the nodes).
- Branching term now contains local energy $E_L=\hat{H}\Psi/\Psi$ which fluctuates much less than the potential V (for an eigenstate, E_L is constant everywhere in configuration space). No more numerical instability in the branching.
- Importance sampling from drift term $\mathbf{F}(\mathbf{x})=\nabla\Phi_T/\Phi_T$ enhancing density of configs where Φ_T is large i.e. there is a *drift* or *osmotic velocity* directed towards large Φ_T on top of the random diffusion - like electric field on Brownian motion.

Fixed-node approximation most serious error in DMC, but still recover 95-100% of the total energy error made in a Hartree-Fock calculation (the ‘correlation energy’) for typical systems. Error typically increases with increasing atomic number.

Calculating expectation values

Calculate energy from the mixed estimator:

$$
\begin{aligned}
E_{\mathrm{DMC}} &= \frac{\int \Phi_T \Psi (\Phi_T^{-1} \hat{H} \Phi_T)\, d\mathbf{R}}{\int \Phi_T \Psi \, d\mathbf{R}} \\
&\simeq \frac{1}{M} \sum_i E_{\mathrm{L}}(\mathbf{R}_i)
\end{aligned}
$$

For expectation values of operators which do not commute with $\hat{H}$ can reduce bias by using e.g. "extrapolated estimation":

$$
\begin{aligned}
\frac{\int \Psi \hat{A} \Psi \, d\mathbf{R}}{\int \Psi \Psi \, d\mathbf{R}} &= 2\frac{\int \Phi_T \Psi (\Phi_T^{-1} \hat{A} \Phi_T)\, d\mathbf{R}}{\int \Phi_T \Psi \, d\mathbf{R}} \\
&\quad - \frac{\int \Phi_T \Phi_T (\Phi_T^{-1} \hat{A} \Phi_T)\, d\mathbf{R}}{\int \Phi_T \Phi_T \, d\mathbf{R}} + \mathcal{O}(\Psi - \Phi_T)^2
\end{aligned}
$$

or other techniques such as 'future walking' (see later).

Simple DMC algorithm

- Generate walker ensemble drawn from some initial distribution (normally Φ_T^2 calculated via VMC).
- Evaluate the drift vector $\mathbf{F}$ and local energy E_{L} for walker.
- Propagate configuration for a time step τ

$$\mathbf{R} = \mathbf{R}' + \tau\mathbf{F} + \chi$$

  χ is a $3N$-dimensional vector with components normally distributed (variance τ and zero mean).
- Check whether the configuration has crossed a node. If it has, reject the move.
- Accept-reject step. Accept move with probability

$$p = \min\left[\frac{|\Phi_T(\mathbf{R}')|^2 G(\mathbf{R}', \mathbf{R}, \tau)}{|\Phi_T(\mathbf{R})|^2 G(\mathbf{R}, \mathbf{R}', \tau)}, 1\right]$$

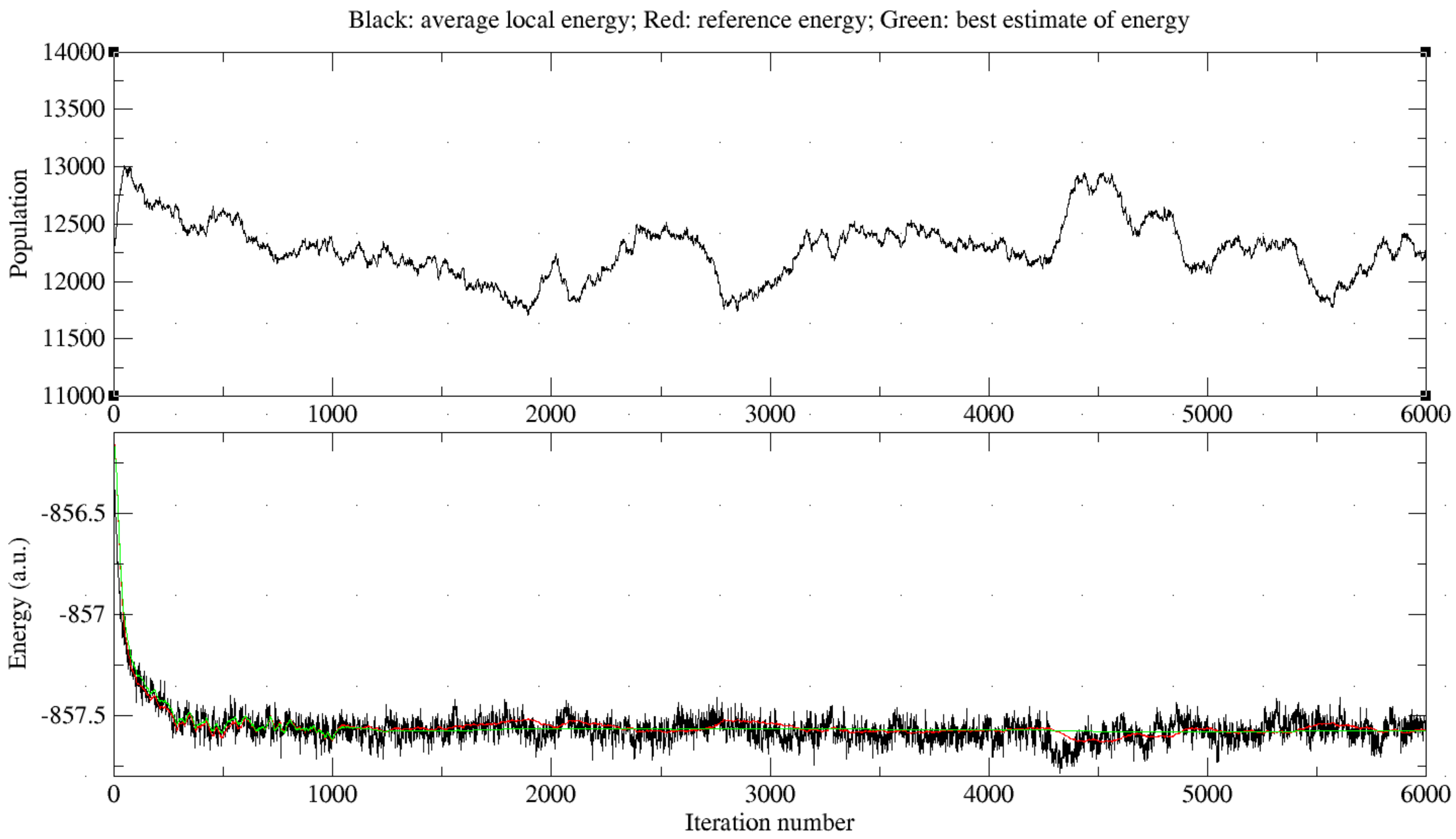
  Choose τ such that $\geq$99% of moves are accepted.
- Branching. Calculate the number of copies of this configuration which continue at the next step

$$M = \mathrm{INT}\left\{\eta + \exp\left[-\tau\left(\frac{E_{\mathrm{L}}(\mathbf{R}) + E_{\mathrm{L}}(\mathbf{R}')}{2} - E_{\mathrm{T}}\right)\right]\right\}$$

  η is a random number uniformly distributed in (0,1)
- Accumulate local energy E_{L}
- Repeat previous steps for each walker in the ensemble
- Repeat until the required accuracy is obtained. During the process E_{T} is occasionally adjusted so that the number of configurations remains roughly constant.

Results of DMC simulation

sI methane clathrate (178 atoms per cell)



DMC cohesive energies

Method	Si	Ge	C	BN	NiO
LSDA	5.28	4.59	8.61	15.07	10.96
VMC	4.48±0.01	3.80±0.02	7.36±0.01	12.85±0.03	8.57±0.01
DMC	4.63±0.02	3.85±0.02	7.35±0.01		9.44±0.01
Exp.	4.62	3.85	7.37	12.9	9.45

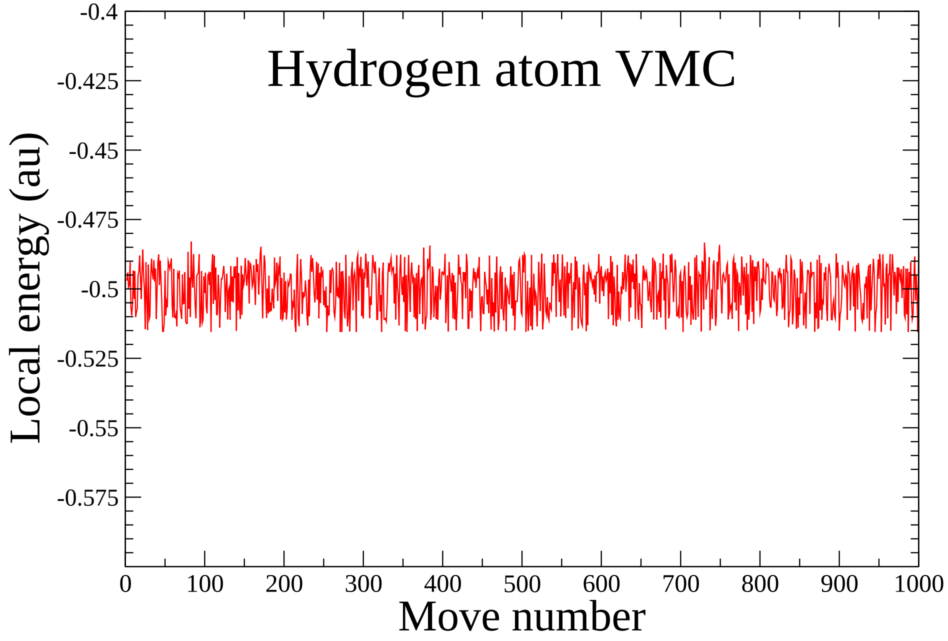
Units: eV per atom Si/Ge/C and eV per 2 atoms BN/NiO

DMC is a big improvement!

VMC : review

$$E_{\mathrm{VMC}} = \frac{\int \Psi \hat{H} \Psi \, d\mathbf{R}}{\int \Psi^2 \, d\mathbf{R}} = \frac{\int \Psi^2 \left(\frac{\hat{H}\Psi}{\Psi} \right) \, d\mathbf{R}}{\int \Psi^2 \, d\mathbf{R}}$$

•Generate points distributed according to Ψ^2 using a random walk and the Metropolis algorithm. Average local energies $\frac{\hat{H}\Psi}{\Psi}$over the walk $\longrightarrow$ E_{VMC}.



NOTE: VMC is not particularly useful as a technique in its own right.
Its main purpose is to provide high-quality trial wave functions for DMC.

DMC : review

- Propagate Ψ in imaginary time to enhance ground state component.

$$\Psi(\mathbf{R},\tau+\delta\tau)=\int K(\mathbf{R},\mathbf{R}',\delta\tau)\Psi(\mathbf{R}',\tau)\,d\mathbf{R}'$$

$$\begin{aligned}K(\mathbf{R},\mathbf{R}',\delta\tau) &= (2\pi\delta\tau)^{-\frac{3N}{2}}\exp\left(-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\delta\tau}\right)\\ &\quad\times\exp\left[-\delta\tau\left(\frac{V(\mathbf{R})+V(\mathbf{R}')}{2}-E_{\mathrm{T}}\right)\right]\end{aligned}$$

DMC : review

- Propagate Ψ in imaginary time to enhance ground state component.

$$\Psi(\mathbf{R},\tau+\delta\tau)=\int K(\mathbf{R},\mathbf{R}',\delta\tau)\Psi(\mathbf{R}',\tau)\,d\mathbf{R}'$$

$$\begin{aligned}K(\mathbf{R},\mathbf{R}',\delta\tau) &= (2\pi\delta\tau)^{-\frac{3N}{2}}\exp\left(-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\delta\tau}\right)\\ &\quad\times\exp\left[-\delta\tau\left(\frac{V(\mathbf{R})+V(\mathbf{R}')}{2}-E_{\mathrm{T}}\right)\right]\end{aligned}$$

- Impractical for fermions (ground state bosonic; huge fluctuations).

DMC : review

•Propagate Ψ in imaginary time to enhance ground state component.

$$\Psi(\mathbf{R},\tau+\delta\tau) = \int K(\mathbf{R},\mathbf{R}',\delta\tau)\Psi(\mathbf{R}',\tau)\, d\mathbf{R}'$$

$$\begin{aligned} K(\mathbf{R},\mathbf{R}',\delta\tau) &= (2\pi\delta\tau)^{-\frac{3N}{2}} \exp\left(-\frac{|\mathbf{R}-\mathbf{R}'|^2}{2\delta\tau}\right) \\ &\times \exp\left[-\delta\tau\left(\frac{V(\mathbf{R})+V(\mathbf{R}')}{2} - E_{\mathrm{T}}\right)\right] \end{aligned}$$

•Impractical for fermions (ground state bosonic; huge fluctuations).
•Solution : use $f(\mathbf{R},\tau) = \Psi(\mathbf{R},\tau)\Phi_T(\mathbf{R})$ instead of $\Psi(\mathbf{R},\tau)$.

$$\begin{aligned} K(\mathbf{R}',\mathbf{R},\delta\tau) &= (2\pi\delta\tau)^{-\frac{3N}{2}} \exp\left[-\left(\frac{\mathbf{R}'-\mathbf{R}-\delta\tau\mathbf{F}(\mathbf{R})}{2\delta\tau}\right)^2\right] \\ &\times \exp\left[-\delta\tau\left(\frac{E_{\mathrm{L}}(\mathbf{R})+E_{\mathrm{L}}(\mathbf{R}')}{2} - E_{\mathrm{T}}\right)\right] \end{aligned}$$

Practical Issues

Issues with the wave function

- Wave function generation and transfer. Simulation cells.
- Wave function evaluation. Representation of orbitals. Basis sets.
- Cusp conditions
- Pseudopotentials
- Spin
- Excited states
- Fixed-node approximation and DMC minutiae
- Jastrow factor and optimization strategies

Issues with calculating the total energy

- Electron-electron Coulomb interactions (Ewald, MPC etc. methods)
- Finite size effects
- Electron-ion interactions (non-local integration)

Computational, performance and scaling issues

- Where does the calculation spend its time?
- Scaling with system size
- Sources of error
- Problems remaining to be solved

More about many-electron wave functions

Slater-Jastrow function

$$\Psi(\mathbf{X}) = e^{J(\mathbf{X})} \sum_n c_n D_n(\mathbf{X})$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$, D is a determinant of orbitals, and $e^{J(\mathbf{X})}$ is a 'Jastrow factor' explicitly dependent on the mutual separation of pairs of electrons and their position with respect to the nuclear framework.

Points to note

- Sometimes useful to use more than one determinant.
- Nodal surface given entirely by determinant part (important for DMC).
- Best way to 'guess' appropriate orbitals for the determinant part is to use results from e.g. molecular orbital theory/band theory calculations.
- Calculating the orbitals is often the most expensive part of the calculation. Need to pay attention to representing them *efficiently*.
- Parameters in Jastrow factor obtained by optimization (usually variance minimization). For linear parameters an extra-fast method is available.
- Can optimize parameters that affect nodal surface (orbital parameters; determinant coefficients; backflow) but this introduces additional problems. See later.

How to generate a trial wave function

Is the system finite or does it have periodic boundary conditions?

Usually use a 'molecular orbital theory' or 'band theory' method to generate suitable orbitals. A QMC code thus requires links to standard programs.

This is often not necessary in model systems where orbitals are already known (e.g. homogeneous electron gas) or where they can be simply parametrized.

What method to use?

Hartree-Fock theory - best possible set of orbitals for a single determinant wave function in the case where the only correlation between particles is due to antisymmetry (parallel spins only).

Density functional theory - best possible set of orbitals for a single determinant (Kohn-Sham) wave function in a fictitious 'non-interacting' system with the same density as the true one. All the complicated many-body physics is transferred to the exchange-correlation functional which gives the 'interacting bit' of the total energy from this density.

Quantum chemistry correlated wave function methods - various multideterminant methods with acronyms (e.g. MCSCF, CASSCF, CIS, CISD, CCSD(T)). Remember the advantage of QMC is its scaling; try not to use an N^7 method to create the trial wave function, otherwise what's the point?

What do we do with the wave function?

(1) Move the electrons

To implement the Metropolis algorithm, propose random electron moves accepted with probability :

$$a(\mathbf{r} \longrightarrow \mathbf{r}') = \min\left[1, \frac{\rho(\mathbf{r}')}{\rho(\mathbf{r})}\right] = \min\left[1, \frac{\Psi(\mathbf{X}')^2}{\Psi(\mathbf{X})^2}\right]$$

Slater matrix contains value of every orbital at current location of every electron. Determinant of this matrix times the Jastrow factor gives the value of the many-electron wave function at that point in configuration space. The ratio of the squares of the old and new wave functions after an electron hop appears in the Metropolis acceptance probability above.

What do we do with the wave function?

(1) Move the electrons

To implement the Metropolis algorithm, propose random electron moves accepted with probability :

$$a(\mathbf{r} \longrightarrow \mathbf{r}') = \min\left[1, \frac{\rho(\mathbf{r}')}{\rho(\mathbf{r})}\right] = \min\left[1, \frac{\Psi(\mathbf{X}')^2}{\Psi(\mathbf{X})^2}\right]$$

Slater matrix contains value of every orbital at current location of every electron. Determinant of this matrix times the Jastrow factor gives the value of the many-electron wave function at that point in configuration space. The ratio of the squares of the old and new wave functions after an electron hop appears in the Metropolis acceptance probability above.

NOTE: Not necessary to reevaluate entire determinant (scales as N^3) every time we move an electron. In fact we:

- Propose move of an electron from point A to point B.
- Evaluate the value of *every* orbital at point B (i.e. we recalculate one column of the Slater matrix).
- New value of *ratio* given by a dot product of column of transpose of inverse Slater matrix with the regenerated column (scales as N). Update of inverse Slater matrix then scales as N^2.

What do we do with the wave function?

(2) Need $\{x, y, z\}$ Cartesian first derivatives and the Laplacian of the orbitals at the new electron position to calculate energy and drift, e.g.

Kinetic energy

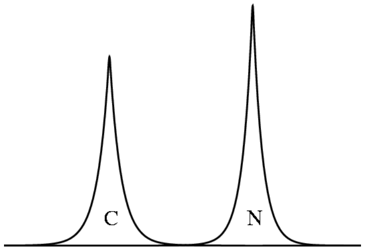
$$
\begin{aligned}
K = \sum_{i=1}^{N} K_i &= \sum_{i=1}^{N} -\tfrac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} \\
\mathbf{F}_i &= \frac{1}{\sqrt{2}}\nabla_i\left(\ln|\Psi|\right) = \frac{1}{\sqrt{2}}\frac{\nabla_i\Psi}{\Psi} \\
T_i &= -\frac{1}{4}\nabla_i^2\left(\ln|\Psi|\right) = -\frac{1}{4}\frac{\nabla_i^2\Psi}{\Psi} + \frac{1}{4}\left(\frac{\nabla_i\Psi}{\Psi}\right)^2
\end{aligned}
$$

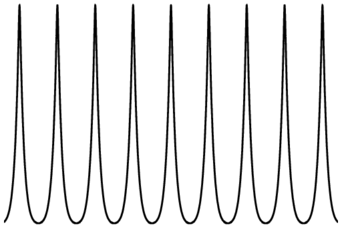
K_i can then be expressed in terms of $\mathbf{F}_i$ and T_i as $K_i = 2T_i - |\mathbf{F}_i|^2$. Integrating shows that $\langle K_i\rangle = \langle|\mathbf{F}_i|^2\rangle = \langle T_i\rangle$ - useful!

Drift vector

$$
G^{\text{diffusion}}(\mathbf{R}', \mathbf{R}, \delta\tau) = (2\pi\delta\tau)^{-\frac{3N}{2}} \exp\left[-\left(\frac{\mathbf{R}' - \mathbf{R} - \delta\tau\frac{\nabla_{\mathbf{R}}\Psi}{\Psi}}{2\delta\tau}\right)^2\right]
$$

Representation of orbitals





Can represent orbitals and derivatives on a grid, or we can use an expansion in a basis set (which may consist of localized or delocalized functions).

Orbitals in periodic potential can be made to obey BLOCH'S THEOREM:

$$\Psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \text{ or } \Psi_{n\mathbf{k}}(\mathbf{r}+\mathbf{t}) = \Psi_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{t}}$$

Delocalized basis (plane waves) : $\Psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_n^{\mathbf{G}} \, e^{i\mathbf{G}\cdot\mathbf{r}}$.

Localized basis (e.g. atom-centred Gaussians $\{\chi\}$): form *Bloch sum* by combining χ and periodic images modulated by a phase factor: $\Phi_{m\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{t}} \chi_m^{\mathbf{t}}(\mathbf{r} - \mathbf{r}_a - \mathbf{t}) e^{i\mathbf{k}\cdot\mathbf{t}}$ where $\chi_m^{\mathbf{t}}(\mathbf{r} - \mathbf{r}_a - \mathbf{t})$ is an atomic function (located at $\mathbf{r_a}$ in the zero cell) translated into cell $\mathbf{t}$. Then : $\Psi_{n\mathbf{k}}(\mathbf{r}) = \sum_m c_n^m \Phi_{m\mathbf{k}}(\mathbf{r})$

In QMC, it is hugely preferable to use localized functions in the basis, since only a subset of them contribute to each orbital at a point. Gain a factor of N over plane waves in scaling with system size.

Math(s)

Given three *not-necessarily orthogonal* basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, the component of an arbitrary vector $\mathbf{d}$ along each basis vector is $\mathbf{d} \cdot \mathbf{b}_i$ where $\mathbf{b}_i$ is a *reciprocal vector* to $\mathbf{a}_i$:

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \qquad \mathbf{d} = \sum_i (\mathbf{d} \cdot \mathbf{b}_i)\mathbf{a}_i$$

Math(s)

Given three *not-necessarily orthogonal* basis vectors $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$, the component of an arbitrary vector $\mathbf{d}$ along each basis vector is $\mathbf{d} \cdot \mathbf{b}_i$ where $\mathbf{b}_i$ is a *reciprocal vector* to $\mathbf{a}_i$:

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3} \qquad \mathbf{d} = \sum_i (\mathbf{d} \cdot \mathbf{b}_i)\mathbf{a}_i$$

Solid-state physics

If $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ are the primitive translation vectors of a real-space crystalline lattice, then the *reciprocal lattice* is then mapped out by the reciprocal lattice vectors $\mathbf{G}_\mathbf{n}$ defined by $\mathbf{G}_\mathbf{n} = 2\pi(n_1\mathbf{b_1} + n_2\mathbf{b_2} + n_3\mathbf{b_3})$ where $\mathbf{n}$ represents an arbitrary triplet of integers n_1, n_2, n_3.

The **reciprocal lattice** is a *Fourier space for arbitrary functions that have the lattice periodicity*

$$f(\mathbf{r}) = \sum_\mathbf{n} c_\mathbf{n} \exp(i\mathbf{G}_\mathbf{n} \cdot \mathbf{r})$$

This is because all functions $\exp(i\mathbf{G}_\mathbf{n} \cdot (\mathbf{r} + \mathbf{R}_\mathbf{m}))$ are invariant under all possible lattice translations $\mathbf{R}_\mathbf{m}$ since $\exp(i\mathbf{G}_\mathbf{n} \cdot \mathbf{R}_\mathbf{m}) = 1$.

Arbitrary functions of position within the crystal

Consider a 1D reciprocal space. Make two sets of points :(1) the reciprocal lattice vectors $\mathbf{G_n}$, and (2) a set of $\mathbf{k}$ points (all the points between $-\frac{1}{2}\mathbf{G_1} = -\frac{\pi}{a}$ and $\frac{1}{2}\mathbf{G_1} = \frac{\pi}{a}$ where a is real space primitive lattice constant).

- Waves $\exp(i\mathbf{G_n} \cdot \mathbf{r})$ are either constant ($\mathbf{G_0} = 0$) or have a wavelength less than or equal to a ($\mathbf{G_n} \neq 0$)). All are periodic in the primitive lattice.
- Waves $\exp(i\mathbf{k_n} \cdot \mathbf{r})$ are *not* periodic in the primitive lattice, and have a wavelength longer than a. Any $\mathbf{k}$ not in the *first Brillouin zone* (i.e. not in red cell) can be reduced into it since it is the sum of some $\mathbf{G}$ and some $\mathbf{k}$ between $-\frac{1}{2}\mathbf{G_1}$ and $\frac{1}{2}\mathbf{G_1}$.

Can therefore Fourier expand an arbitrary function of position within the solid which is not necessarily periodic in the real space lattice but obeys the boundary conditions at the surface :

$$g(\mathbf{r}) = \sum_{\mathbf{m}} \sum_{\mathbf{n}} c_{\mathbf{m,n}} e^{i(\mathbf{k_n}+\mathbf{G_m})\cdot\mathbf{r}} = \sum_{\mathbf{n}} \left(\sum_{\mathbf{m}} c_{\mathbf{m,n}} e^{i\mathbf{G_m}\cdot\mathbf{r}}\right) e^{i\mathbf{k_n}\cdot\mathbf{r}} = \sum_{\mathbf{n}} u_{\mathbf{n}}(\mathbf{r}) e^{i\mathbf{k_n}\cdot\mathbf{r}}$$

where the red parts of the formula are cell-periodic functions.

Bloch functions

- The plane wave basis states $\exp(i(\mathbf{k_n} + \mathbf{G_m})$ are not generally eigenfunctions ϕ of the Hamiltonian unless the potential is independent of position. We have seen that some linear combination of them must be:

$$\phi(r) = \sum_{\mathbf{m}} \sum_{\mathbf{n}} c_{\mathbf{m},\mathbf{n}} e^{i(\mathbf{k_n}+\mathbf{G_m})\cdot\mathbf{r}}$$

- Energy eigenvalues are found to depend on a specific vector $\mathbf{k_n}$. Can energy eigenfunctions be constructed from individual subsets of the set of terms appearing in this general expansion, each subset corresponding to a single value of $\mathbf{k_n}$? Yes :

$$\phi_{\mathbf{n}}(\mathbf{r}) = \sum_{\mathbf{m}} \sum_{\mathbf{n}'} \delta_{\mathbf{n},\mathbf{n}'} c_{\mathbf{m},\mathbf{n}'} e^{i(\mathbf{k}'_{\mathbf{n}}+\mathbf{G_m})\cdot\mathbf{r}} = e^{i\mathbf{k_n}\cdot\mathbf{r}} \sum_{\mathbf{m}} c_{\mathbf{n},\mathbf{m}} e^{i\mathbf{G_m}\cdot\mathbf{r}}$$

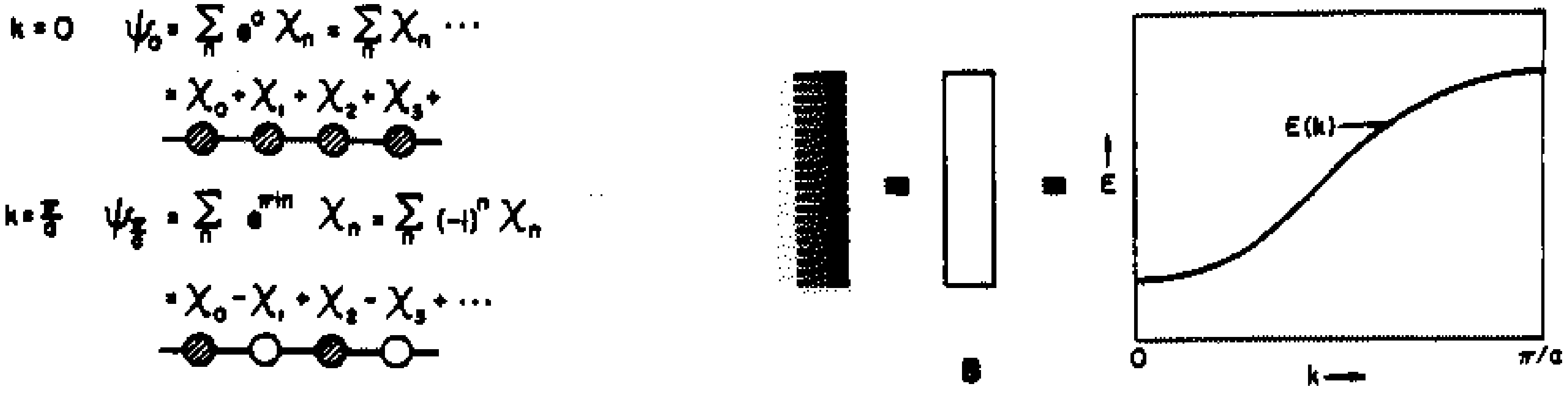
- Energy eigenfunctions for a periodic potential may be written in the form of **Bloch functions** :

$$\phi_n(\mathbf{r}) = u_n(\mathbf{r}) e^{i\mathbf{k_n}\cdot\mathbf{r}}$$

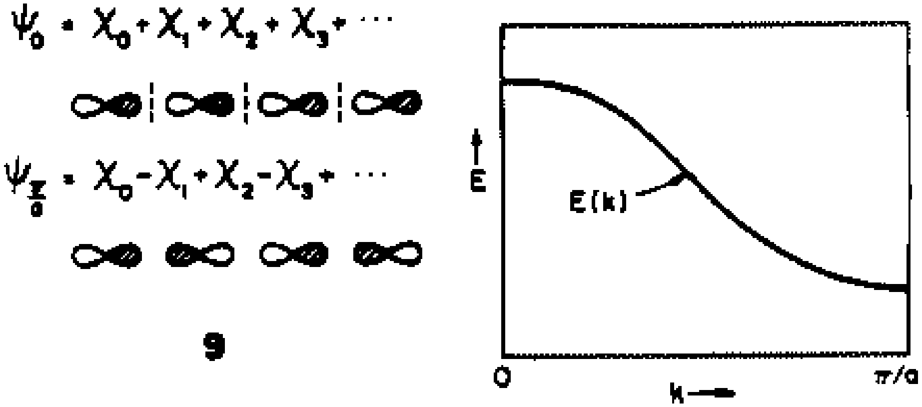
Any arbitrary function of position may be written as a linear combination of Bloch functions from different **k**, and we will do so later in constructing such things as *maximally localized Wannier orbitals* for ‘improved scaling’ QMC.

Band structure

s orbitals ($\mathbf{k} = \frac{\pi}{a}$ most ‘antibonding’)



p orbitals ($\mathbf{k} = 0$ most ‘antibonding’)



The topology of orbital interactions determines which way bands ‘run’.

k point sampling

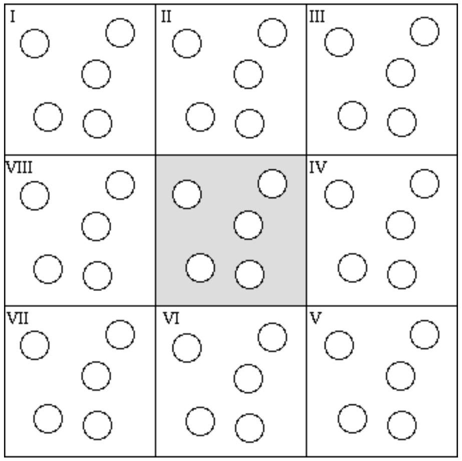
- Electronic states allowed only at set of **k** points determined by boundary conditions on the bulk solid. Density of allowed **k** points is proportional to the volume of the solid.

- Infinite number of electrons in infinite solid accounted for by an infinite number of **k** points in the 1st Brillouin zone, and only a finite number of electronic states are occupied at each **k** point.

- Don't need to calculate electronic states at infinite number of **k** points, since the wave functions at **k** points that are very close together will be almost identical. Represent *region* of **k** space by single **k** point. Then only finite number of objects need to be calculated in order to calculate the total energy of the solid.

- Calculation of properties requires *Brillouin zone integration* - done by straightforward sums over states using a special point scheme and modest number of **k** points. Usual to use Monkhorst-Pack mesh (uniformally spaced **k**-points with a possible offset of the origin).

Many-body wave functions and periodic boundary conditions



- In band theory can reduce a problem into the real space primitive cell using **k** points. In general with a many-body wave function we cannot.
- A primitive cell band theory calculation done on a $2 \times 2 \times 2$ **k**-point grid defines a many-body wave function for a $2 \times 2 \times 2$ real-space supercell.
- How big does a cell need to be to represent the wave function properly? Needs to accomodate the exchange-correlation hole around each electron. Forces on particles in zero cell need to be approximately what they would be if the same cell was placed in genuine (aperiodic) bulk solid.
- For silicon, $2 \times 2 \times 2$ supercell might be sufficient (larger cells required for metals). Concept of 'plucking' useful to reduce errors in Brillouin zone integration in band theory calculation (subset of **k** grid used in orbitals for supercell).

Many-body Bloch theorem

Invariance of $\hat{H}$ under translation of *any* electron by simulation cell lattice vector :

$$\Psi_{\mathbf{k}_s}(\{\mathbf{r}_i\}) = u_{\mathbf{k}_s}(\{\mathbf{r}_i\}) \exp\left(i\mathbf{k}_s \cdot \sum_{i=1}^{N} \mathbf{r}_i\right)$$

Invariance of $\hat{H}$ under translation of *all* electrons by primitive lattice vector.

$$\Psi_{\mathbf{k}_p}(\{\mathbf{r}_i\}) = w_{\mathbf{k}_s}(\{\mathbf{r}_i\}) \exp\left(i\mathbf{k}_p \cdot \frac{1}{N}\sum_{i=1}^{N} \mathbf{r}_i\right)$$

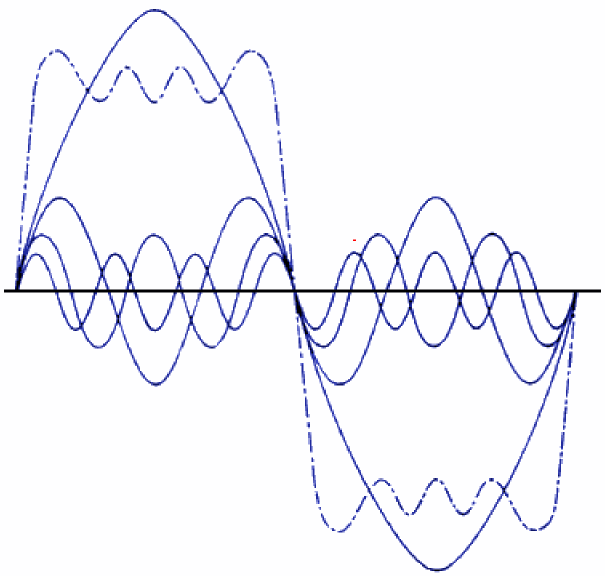
If simulation cell contains more than one primitive cell, the two wave vectors are distinct labels of the many-body wave function and both are required to specify the translational symmetry.

Other QMC considerations

- Bloch functions are normally complex, and often prefer to make them real (with appropriate linear combinations) for use in FN-DMC.
- Minimize 'independent particle finite size effects' by choosing k_s appropriately (or even 'twist averaging' over different values). Still leaves 'Coulomb finite size effects' caused by the artificial periodicity in the many-electron Hamiltonian. Later.

Basis sets for orbitals : plane waves

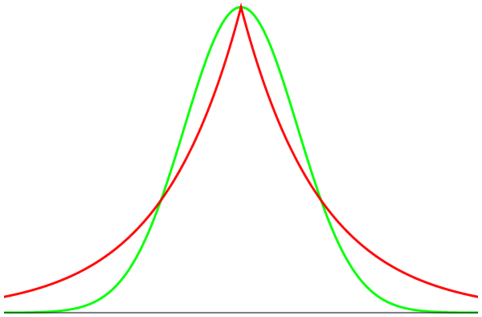
Use a Fourier expansion in plane waves, where the expansion coefficients are assumed to be zero for wave vectors $\mathbf{G}$ whose kinetic energy $\frac{|\mathbf{G}|^2}{2}$ exceeds a cutoff.



GOOD : Orthonormal complete set. Universal. Systematic improvement by changing single cutoff parameter. DFT analytic gradients/forces accurate and cheap. Codes widely available.

BAD : Restricted detail in real space - need to use pseudopotentials. Must use periodic boundary conditions, even for molecules/surfaces etc. Non-local exchange difficult. FFTs not good for massively parallel computation. Delocalization $\longrightarrow$ extra factor of N in scaling with system size in QMC.

Basis sets for orbitals : Gaussian functions

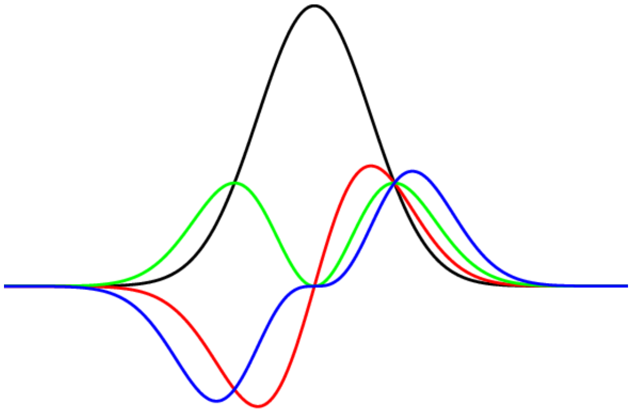


In quantum chemistry, long range behaviour and nuclear cusp originally suggested use of Slater-type orbitals:

$$\chi^{STO} = r^{n-1} exp\left(-\zeta r\right) Y_{lm}\left(\theta, \phi\right)$$

Not suitable for fast calc of multi-centre integrals so Cartesian Gaussian functions usually used instead:

$$\chi^{GTF} = exp\left(-\alpha r^2\right) x^l y^m z^n$$



SEPARABLE : $\chi^{GTF} = \chi_x \chi_y \chi_z$ with $\chi_x = exp\left(-\alpha(x-x_a)^2\right)(x-x_a)^l$ etc.. Categorize into 'shells' with $L = l + m + n$:

s : 1 $(L = 0)$

p : x, y, z $(L = 1)$

d : x^2, xy, xz, y^2, yz, z^2 (L=2)

GOOD : Localized. Don't require pseudopotentials. Don't require periodic boundary conditions. Easier non-local exchange in one-electron codes (HF/hybrid DFT.. $\longrightarrow$ good for e.g. systems containing transition elements). Scales better with system size than PW.

BAD : Not universal. Disobey electron-nuclear cusp conditions. Basis set superposition error. Physicists generally refuse to use them, chemists will use nothing else.

Tedious Acronyms for Gaussian basis sets

STO-2G, STO-3G, STO-6G, STO-3G*, 3-21G, 3-21++G, 3-21G*, 3-21GSP, 4-31G, 4-22GSP, 6-31G, 6-31G-Blaudeau, 6-31++G, 6-31G*, 6-31G**, 6-31G*-Blaudeau, 6-31+G*, 6-31++G**, 6-31G(3df,3pd), 6-311G, 6-311G*, 6-311G**, 6-311+G*, 6-311++G**, 6-311++G(2d,2p), 6-311G(2df,2pd), 6-311++G(3df,3pd), MINI (Huzinaga), MINI (Scaled), MIDI (Huzinaga), MIDI!, SV (Dunning-Hay), SVP+Diffuse (Dunning-Hay), DZ (Dunning), DZP (Dunning), DZP+Diffuse (Dunning), TZ (Dunning), Chipman DZP+Diffuse, cc-pVDZ, cc-PVTZ, cc-pVQZ, cc-pV5Z, cc-pV6Z, pV6Z, pV7Z, cc-pVDZ(seg-opt), cc-pVTZ(seg-opt), cc-PVQZ(seg-opt), cc-pCVDZ, cc-pCVTZ, cc-pCVQZ, cc-pCV5Z, aug-cc-pVDZ, aug-cc-pVTZ, aug-cc-pVQZ, aug-cc-pV5Z, aug-cc-pV6Z, aug-pV7Z, aug-cc-pCVDZ, aug-cc-pCVTZ, aug-cc-pCVQZ, aug-cc-pCV5Z, d-aug-cc-pVDZ, d-aug-cc-pVTZ, d-aug-cc-pVQZ, d-aug-cc-pV5Z, d-aug-cc-pV6Z, Feller Misc. CVDZ, Feller Misc cVTZ, Feller Misc. CVQZ, NASA Ames ANO, Roos Augmented Double Zeta ANO, Roos Augmented Triple Zeta ANO, WTBS, GAMESS VTZ, GAMESS PVTZ, Partridge Uncontr. 1, Partridge Uncontr. 2, Partridge Uncontr. 3, Ahlrichs VDZ, Ahlrichs, pVDZ, Ahlrichs VTZ, Ahlrichs TZV, Binning/Curtiss SV, Binning/Curtiss VTZ, Binning/Curtiss SVP, Binning-Curtiss VTZP, Mclean/Chandler VTZ, SV+Rydberg (Dunning-Hay), SVP+Rudberg (Dunning-Hay), SVP+Diffuse+Rydberg, DZ+Rydberg (Dunning), DZP+Rydberg (Dunning), DZ+Double Rydberg (Dunning-Hay), SV+Double Rydberg (Dunning-Hay), Wachters+f, Bauschlicher ANO, Sadlej pVTZ, Hay-Wadt MB(n+1)ECP, Hay-Wadt VDZ(n+1)ECP, LANL2DZ ECP, SBKJC VDZ ECP, CRENBL ECP, CRENBS ECP, Stuttgart RLC ECP, Stuttgart RSC ECP, DZVP (DFT Orbital), DZVP2 (DFT Orbital), TZP (DFT Orbital), DeMon Coulomb Fitting, DGauss A1 DFT Coulomb Fitting, DGauss A1 DFT Exchange Fitting, DGauss A2 DFT Coulomb Fitting, DGauss A2 DFT Exchange Fitting, Ahlrichs Coulomb Fitting, cc-pVDZ-fit2-1, cc-pVTZ-fit2-1, cc-pVDZ_DK, cc-pVTZ_DK, cc-pVQZ_DK, cc-pV5Z_DK, cc-pVDZ(pt/sf/fw), cc-PVTZ(pt/sf/fw), cc-pVQZ(pt/sf/fw), cc-pV5Z(pt/sf/fw), cc-pVDZ(fi/sf/fw), cc-pVTZ(fi/sf/fw), cc-pVQZ(fi/sf/fw), cc-pV5Z(fi/sf/fw), cc-pVDZ(pt/sf/sc), cc-pVDZ(pt/sf/lc), cc-pVTZ(pt/sf/sc), cc-PVTZ(pt/sf/lc), ccp-PVQZ(pt/sf/sc), cc-pVQZ(pt/sf/lc), cc-PV5Z(pt/sf/sc), cc-PV5Z(pt/sf/lc), cc-pVDZ(fi/sf/sc), cc-PVDZ(fi/sf/lc), cc-PVTZ(fi/sf/sc), cc-PVTZ(fi/sf/lc), cc-PVQZ(fi/sf/sc), cc-PVQZ(fi/sf/lc), cc-PV5Z(fi/sf/sc), cc-pV5Z(fi/sf/lc), Pople-Style Diffuse, STO-3G* Polarization, 3-21G* Polarization, 6-31G* Polarization, 6-31G** Polarization, 6-311G* Polarization, 6-311G** Polarization, Pople (2d/2p) Polarization, Pople (3df,3pd) Polarization, HONDO7 Polarization, Huzinaga Polarization, Dunning-Hay Diffuse, aug-cc-pVDZ Diffuse, aug-cc-pVTZ Diffuse, aug-cc-pVQZ Diffuse, aug-cc-pV5Z Diffuse, aug-cc-pV6Z Diffuse, aug-pV7Z Diffuse, d-aug-cc-pVDZ Diffuse, d-aug-cc-pVTZ Diffuse, d-aug-cc-pVQZ Diffuse, d-aug-cc-pV5Z Diffuse, d-aug-cc-pV6Z Diffuse, DHMS Polarization, Dunning-Hay Rydberg, Dunning-Hay Double Rydberg, Binning-Curtiss (1d Polarization), Binning-Curtiss (df) Polarization, Ahlrichs Polarization, Glendenning Polarization, Blaudeau Polarization, Core/val. Functions (cc-pCVDZ), Core/val. Functions (cc-pCVTZ), Core/val. Functions (cc-pCVQZ), Core/val. Functions (cc-pCV5Z).

Contraction schemes

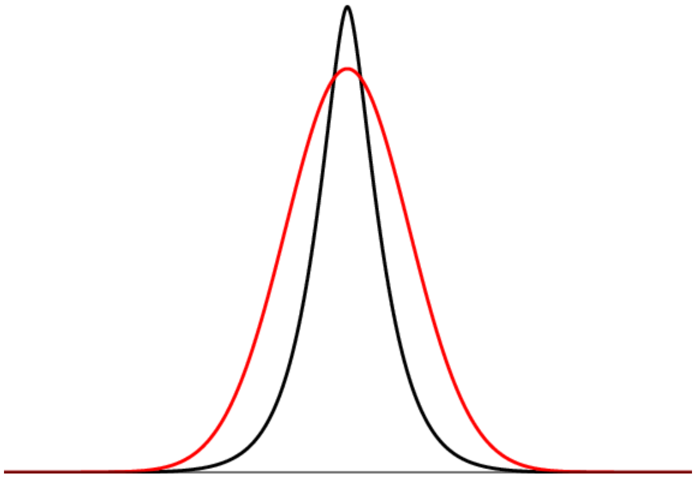
contraction n. A linear combination of Gaussian primitives to be used as a basis function in electronic structure calculations.

$$\chi_i^{CGTF}(r) = \sum_{j=1}^{L} c_j \chi_j^{GTF}(r)$$

$$\chi_j^{GTF}(r) = N_{lm}(\alpha)\, r^l Y_{lm}(\theta,\phi) \exp\left(-\alpha_j r^2\right)$$

In e.g. HF calculations, best results would be obtained if all coefficients in the Gaussian expansion were allowed to vary, but contractions give large increase in efficiency in HF codes by reducing no. of orbital coefficients in the wave function.

α	c_j
3450660.8	0.000034
467601.94	0.000322
92314.514	0.0021
21992.520	0.0112
6082.9917	0.0475
1915.2715	0.157
676.43927	0.3524
263.00267	0.4238
106.89395	0.1593



Generic Gaussian basis set file

```
8 5
S 8 1.0
 8020.0 0.00108
 1338.0 0.00804
 255.4 0.05324
 69.22 0.1681
 23.90 0.3581
 9.264 0.3855
 3.851 0.1468
 1.212 0.0728
SP 4 1.0
 49.43 -0.00883 0.00958
 10.47 -0.0915 0.0696
 3.235 -0.0402 0.2065
 1.217 0.379 0.347
SP 1 1.0
 0.4764 1.0 1.0
SP 1 1.0
 0.1802 1.0 1.0
D 1 1.0 - polarization function
 0.31 1.0
```

Oxygen basis with five shells.

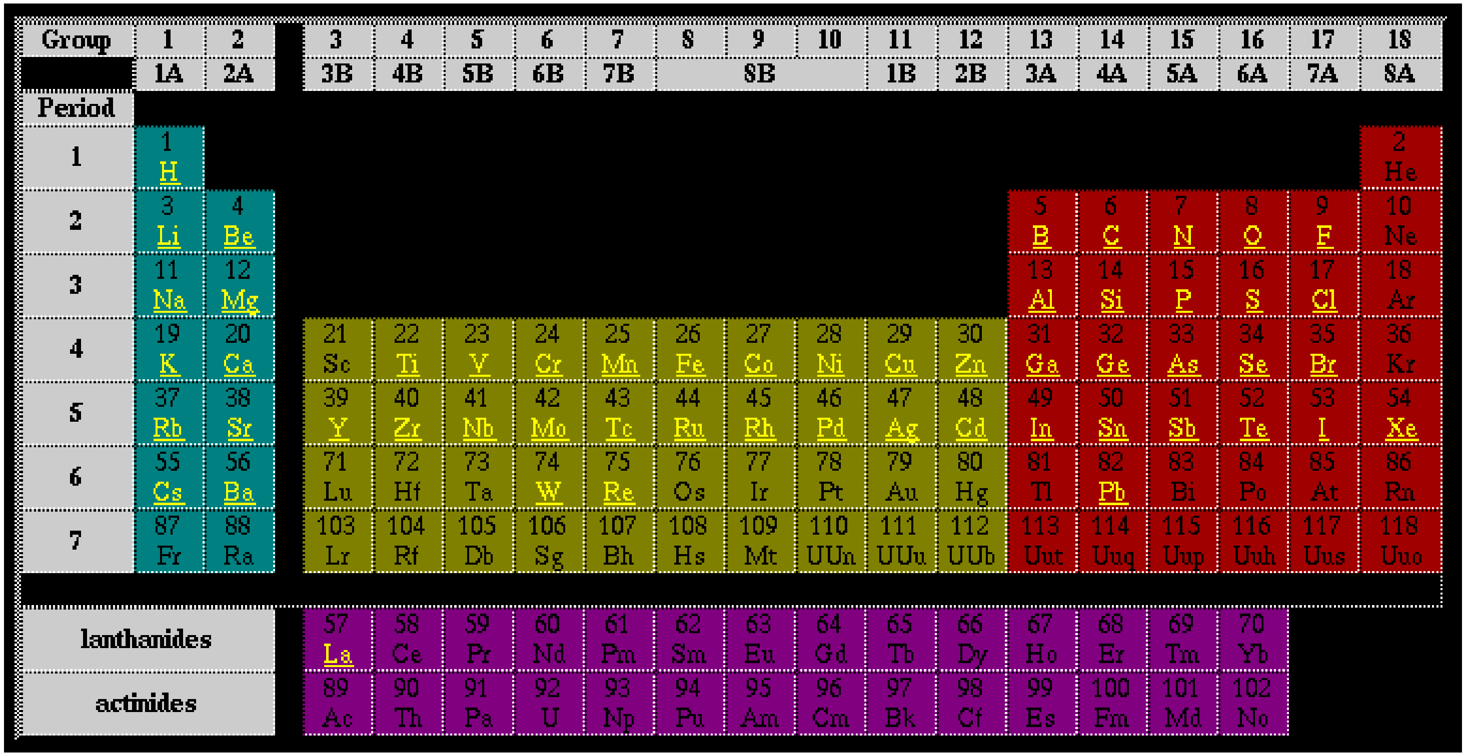
Web libraries

EMSL Molecular basis set library

bse.pnl.gov or www.cse.clrc.ac.uk/qcg/basis/

My basis set library (largely for solid-state calcs)

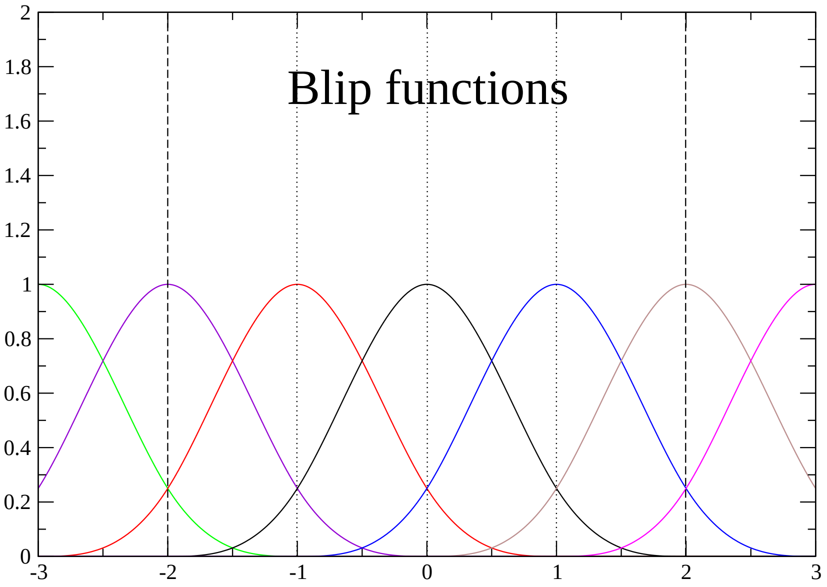
www.tcm.phy.cam.ac.uk/∼mdt26/crystal.html



Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	1A	2A	3B	4B	5B	6B	7B	8B			1B	2B	3A	4A	5A	6A	7A	8A
Period																		
1	1 H																	2 He
2	3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
3	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
4	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
5	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
6	55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
7	87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 UUn	111 UUu	112 UUb	113 Uut	114 Uuq	115 Uup	116 Uuh	117 Uus	118 Uuo
lanthanides			57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb		
actinides			89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No		

Basis sets for orbitals : blip functions (splines)

Expansion in localized spline functions on a uniform grid



Grid spacing unity

$$\begin{aligned}\phi(x) &= 1-\frac{3}{2}x^2+\frac{3}{4}|x|^3 \quad 0\le |x| \le 1\\ &= \frac{1}{4}(2-|x|)^3 \quad 1<|x|\le 2\end{aligned}$$

Grid spacing a

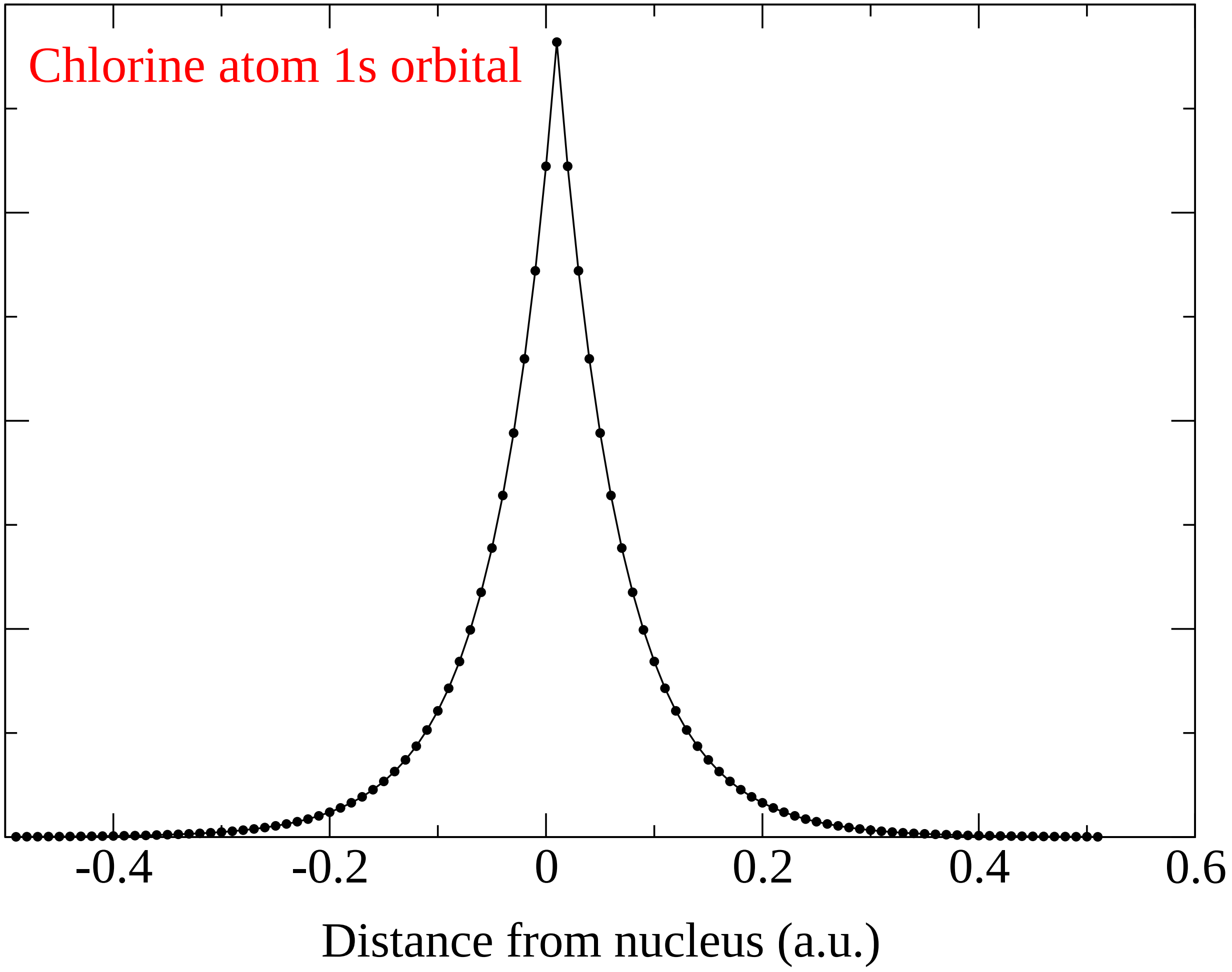
$$\begin{aligned}X_i &= ia\\ \Theta(x-X_i) &= \phi\left(\frac{(x-X_i)}{a}\right)\end{aligned}$$

In three dimensions there are only 64 non-zero blips for each position $\mathbf{r}$. With plane waves the number of functions in e.g. silicon is around 100 per atom.

GOOD : Achieved from transformation of wave function expanded in plane waves with accompanying huge efficiency increase. Localized. Universal.

BAD : Somewhat greedy with memory and disk. Extra step required (blip transformation of plane wave data file).

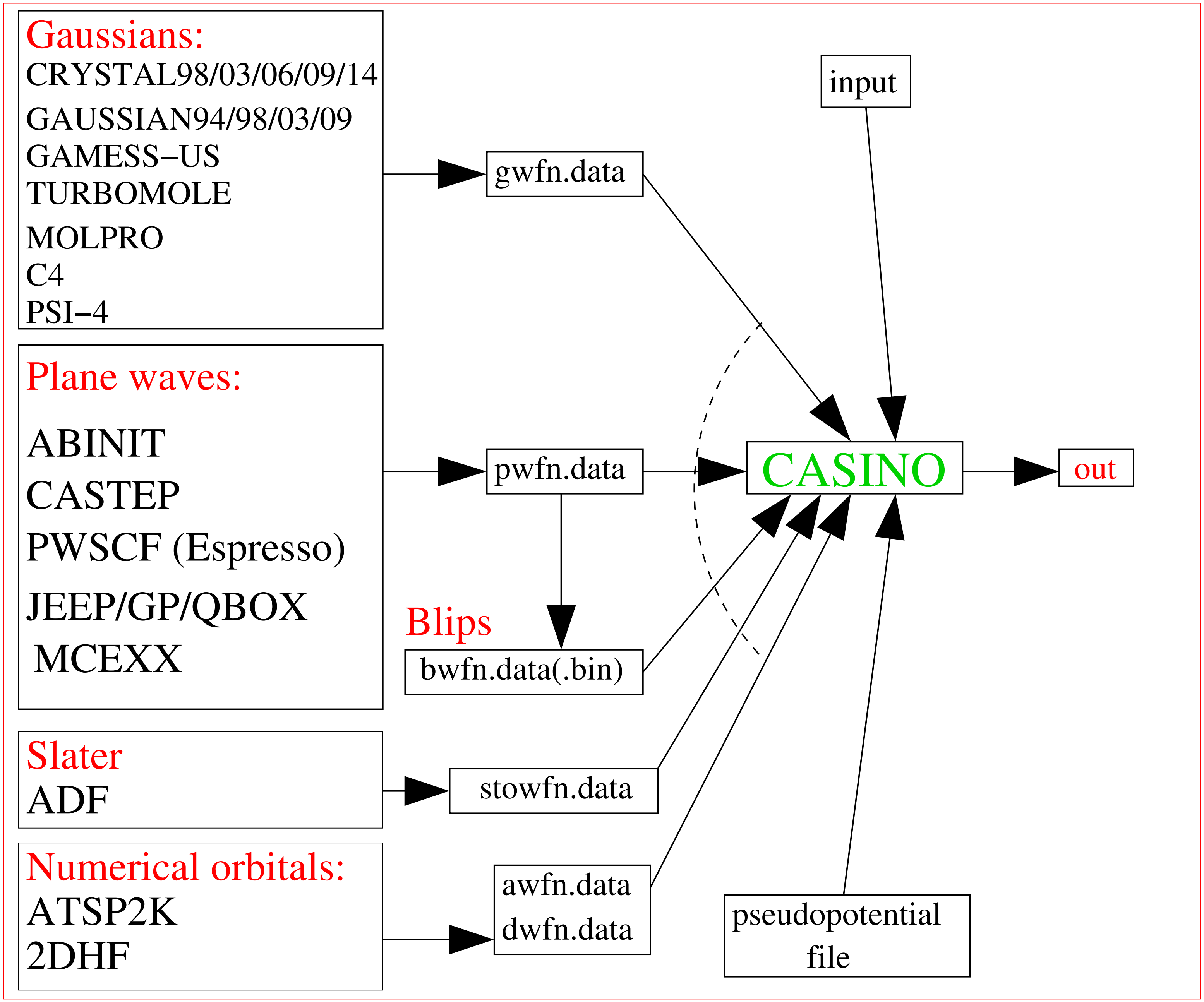
No basis set : use a grid instead



Orbitals and their derivatives are represented numerically on a radial grid. Interpolate to get values at arbitrary position in space.

- Very expensive and inefficient for larger systems.
- Implemented in CASINO for atoms and dimers only.

Interfaces to other programs



Gaussians:
CRYSTAL98/03/06/09/14
GAUSSIAN94/98/03/09
GAMESS–US
TURBOMOLE
MOLPRO
C4
PSI–4

Plane waves:
ABINIT
CASTEP
PWSCF (Espresso)
JEEP/GP/QBOX
MCEXX

Slater
ADF

Numerical orbitals:
ATSP2K
2DHF

gwfn.data

pwfn.data

Blips
bwfn.data(.bin)

stowfn.data

awfn.data
dwfn.data

input

CASINO

out

pseudopotential
file

Spin

Antisymmetric wave function for an N-electron system $(N = N_\uparrow + N_\downarrow)$ in an $S_z = (N_\uparrow - N_\downarrow)/2$ state (non-relativistic; no external magnetic field) can be decomposed in terms of spin components:

$$\Psi(\mathbf{r}_1 s_1, \ldots, \mathbf{r}_N s_N) = \sum_{i=1}^{K} F_i(\mathbf{r}_1, \ldots, \mathbf{r}_N)\chi_i(s_1, \ldots, s_N)$$

Permutation symmetry implies expectation value of spin-dependent operator with (Slater-Jastrow) wave function can be evaluated with a spin-assigned wave function:

$$\Psi(\mathbf{R}) = e^{J(\mathbf{R})} \sum_i D_i^\uparrow(\mathbf{r}_1, \ldots, \mathbf{r}_{N_\uparrow}) D_i^\downarrow(\mathbf{r}_{N_\uparrow+1}, \ldots, \mathbf{r}_N)$$

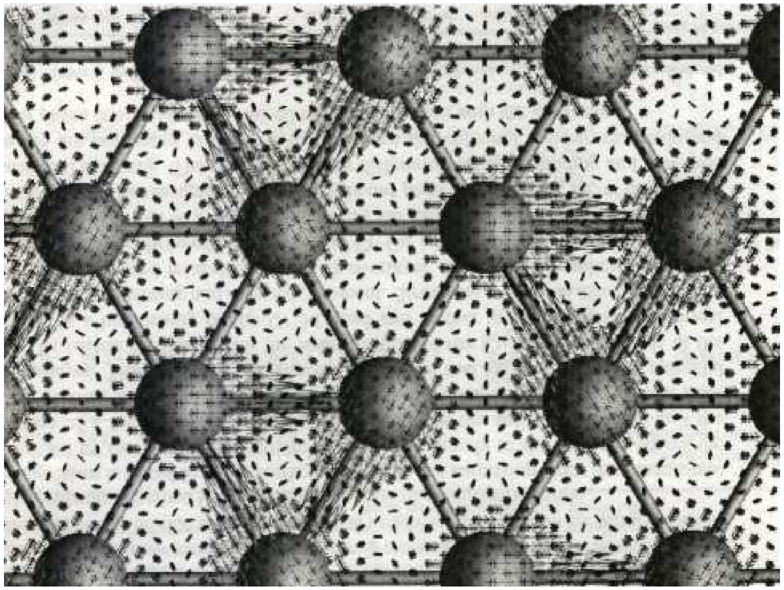
Is it an eigenstate of $\hat{S}^2$ as well as $\hat{S}_z$? Determinant part can be constructed to be so (as in quantum chemistry). Jastrow not necessarily (not invariant under exchange of two antiparallel electrons). Options?

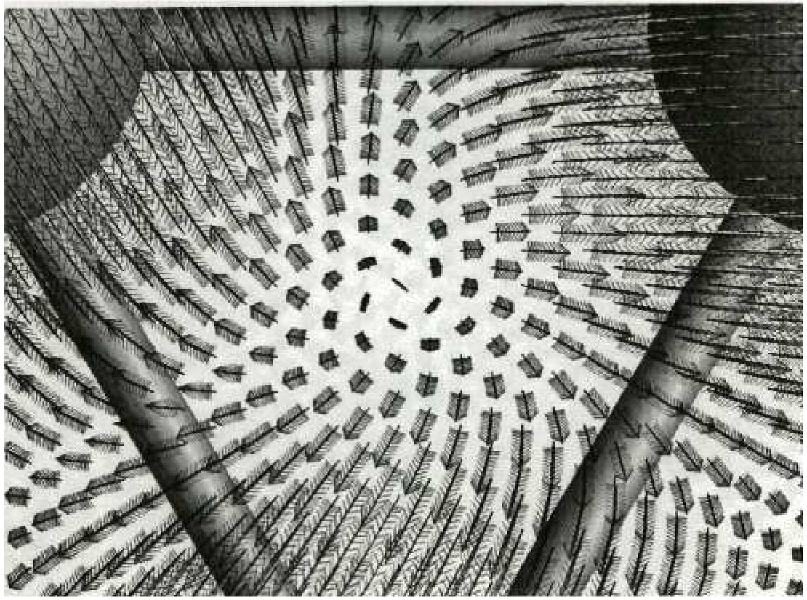
- Use totally symmetric Jastrow factor (increases energy and variance!)
- Wave function with asymmetric Jastrow satisfying cusp conditions (not generally eigenstate of $\hat{S}^2$, but optimization reduces spin contamination).
- Can construct wave function that satisfies cusp conditions and is eigenstate of $\hat{S}^2$ (but increases scaling by N, since need to do explicit sum over spin components unless Jastrow is spin-independent).

Single determinants of one-electron spin orbitals

- Restricted form All spin orbitals are pure space-spin products of the form $\phi_n\alpha$ or $\phi_n\beta$ and are occupied singly or in pairs with a common orbital factor ϕ_n.
- Unrestricted form Spin orbitals no longer occupied in pairs but still pure space-spin products $\phi_n\alpha$ or $\bar{\phi}_n\beta$. However, now have different spatial factors ϕ_n and $\bar{\phi}_n$ for different spins.

What to do for non-collinear spin states?





- General unrestricted form No longer restrict to simple product form. Each spin orbital now a 2-component complex spinor orbital: $\Psi_1 = \phi_1^\alpha\alpha + \phi_1^\beta\beta$ and $\Psi_2 = \phi_2^\alpha\alpha + \phi_2^\beta\beta$. Interesting QMC project!

Total energy

Total energy = KE + electron-ion + electron-electron + ion-ion

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i\alpha} v_\alpha(\mathbf{r}_i, \mathbf{r}_\alpha) + \frac{1}{2}\sum_i \sum_{j\neq i} v(\mathbf{r}_i, \mathbf{r}_j) + \frac{1}{2}\sum_\alpha \sum_{\beta\neq\alpha} v(\mathbf{r}_\alpha, \mathbf{r}_\beta)$$

Electron-electron Coulomb interactions in periodic systems

Solve Poisson's equations subject to periodic boundary conditions:

$$v_{\mathrm{E}}(\mathbf{r}, \mathbf{r}_j) = \sum_{\mathbf{R}} \frac{\mathrm{erfc}\left(\gamma^{\frac{1}{2}}|\mathbf{r} - (\mathbf{r}_j + \mathbf{R})|\right)}{|\mathbf{r} - (\mathbf{r}_j + \mathbf{R})|} - \frac{\pi}{\Omega\gamma}$$

3D Ewald formula

$$+ \frac{4\pi}{\Omega} \sum_{\mathbf{G}\neq\mathbf{0}} \frac{\exp\left(-G^2/4\gamma\right)}{G^2} \exp(i\mathbf{G}\cdot(\mathbf{r} - \mathbf{r}_j))$$

MANY-BODY BLOCH THEOREM

Ewald energy remains unchanged if single electron translated by a supercell lattice vector, or *all* electrons translated by a primitive cell lattice vector.

Coulomb finite size effects

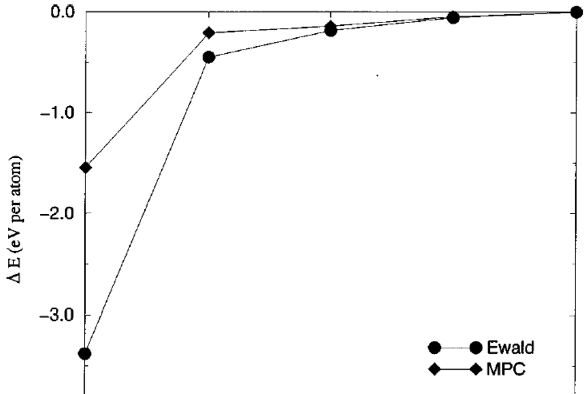
How big does cell have to be before Ewald energy of zero cell in field of the rest of the crystal closely matches infinite supercell size limit?

Sources of error :
(1) ‘Squeezing of XC hole’ (minor effect, usually).
(2) Interaction with periodic array of XC holes, OR EQUIVALENTLY Ewald interaction contains effective ‘depolarization field’ to cancel field due to surface charges. All supercells contain *same net dipole* due to random arrangement of electrons with respect to nuclei. Dipole and field interact.

One solution (there are other more modern ones) is to change many-body Hamiltonian so that interaction with XC hole is exactly $1/r$, without altering Hartree energy. Thus:

$$\hat{H}^{\mathrm{exact}}_{e-e} = \sum_{i>j} f(\mathbf{r}_i - \mathbf{r}_j) + \sum_i \int_{\mathrm{WS}} \rho(\mathbf{r}) \left[v_{\mathrm{E}}(\mathbf{r}_i - \mathbf{r}) - f(\mathbf{r}_i - \mathbf{r}) \right] d\mathbf{r}$$

MODIFIED PERIODIC
COULOMB (MPC)
INTERACTION

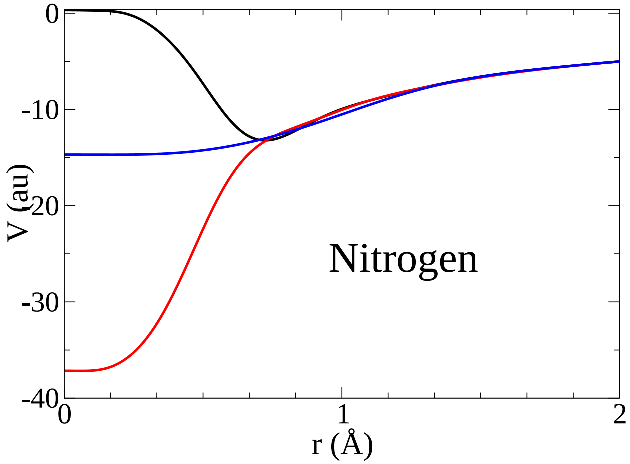


Electron-ion interactions with pseudopotentials

$$\hat{H}_{e-i} = \sum_{i=1}^{N} V_i^{\mathrm{long}}(\mathbf{R}) + \sum_{i=1}^{N} V_i^{\mathrm{short}}(\mathbf{R}) + \sum_{i=1}^{N} \frac{\hat{V}_{\mathrm{nl},i}^{\mathrm{short}} \Psi(\mathbf{R})}{\Psi(\mathbf{R})}$$

long-range local + short-range local + short-range non-local

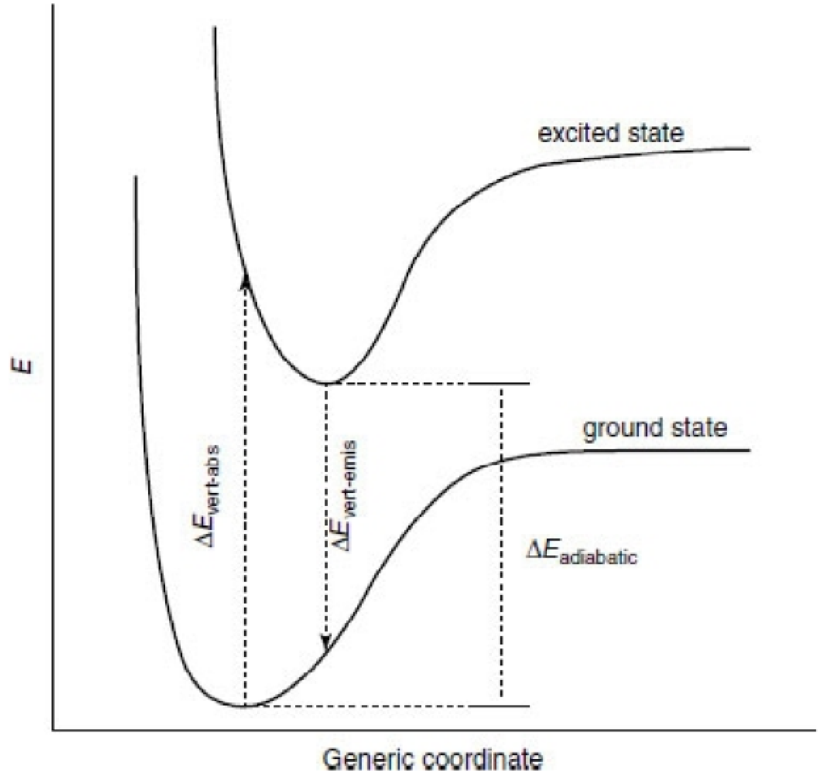
$$V_{\mathrm{nl},i} = \sum_{l} V_{\mathrm{nl},l}^{\mathrm{ps}}(r_i) \frac{2l+1}{4\pi} \int P_l\left[\cos(\theta_i')\right] \times \frac{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i', \mathbf{r}_{i+1}, \ldots, \mathbf{r}_N)}{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_N)} \, d\Omega_{\mathbf{r}_i'}$$



- Horrible non-local angular integration - takes a lot of time.
- Done *approximately* in DMC. Error should be small but not tested extensively (‘LOCALITY APPROXIMATION’).

Excited states

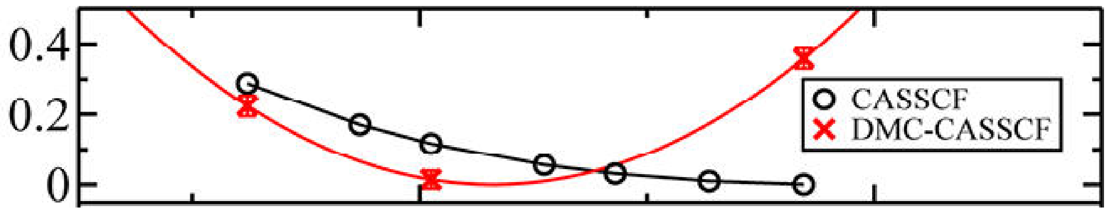
Obviously DMC suffers from same variational collapse issue as other methods: any attempt to optimize the orbitals with respect to the energy (or here, do the DMC imaginary-time projection) will necessarily return Ψ to lowest energy state compatible with the symmetry, usually the ground state unless trial function is orthogonal to it. If that's what you want, then fine; DMC calculations of such states are straightforward. Otherwise we would seem to have a problem.



- BUT fixed-node approximation turns out to be a blessing in disguise. ‘Fixing the nodes’ is actually a very strong constraint, and this essentially ensures convergence to the *lowest energy state compatible with the nodal surface of the trial wave function*.
- DMC therefore gives the exact energy of any state if the nodal surface is exact, and it gives an approximate energy with an approximate nodal surface.
- Important difference from GS: existence of variational principle for excited states cannot be guaranteed in general. Depends on symmetry of Ψ_T. If excited state of interest is lowest state of a given spin symmetry belonging to a 1D irreducible representation (i.e. it's not degenerate or has only accidental degeneracies), the DMC energy is variational. In all other cases, DMC is no longer variational and quality of trial function becomes increasingly important.

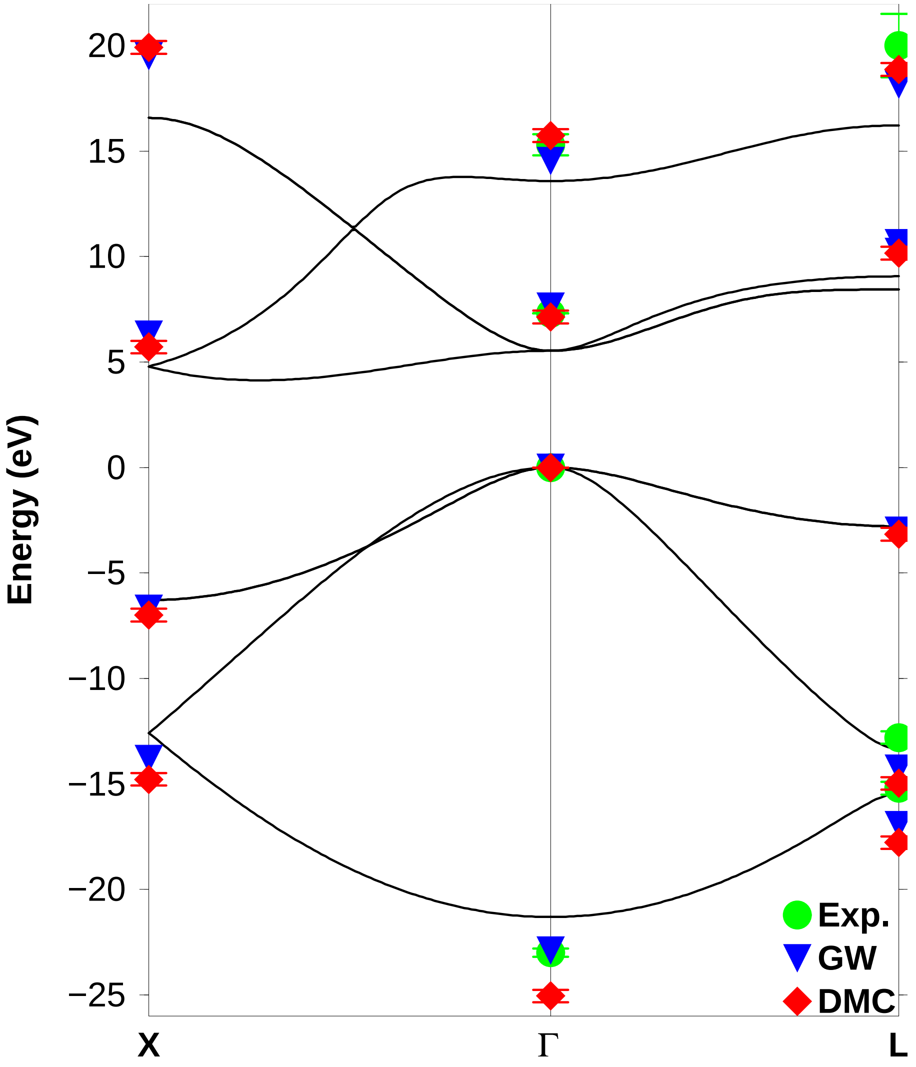
Excited states: where does DMC fit in?

- Closed-shell singlets etc. readily described by single determinants. For general treatment of excited-state potential energy surfaces chemists use CASSCF.
- In typical applications, CASSCF treats only the *static (nondynamical) correlation*. Computing missing dynamical part with e.g. multireference CI only for v. small molecules. Community seems to have settled on using multireference perturbation theory - usually CASPT2 (better but still unfavourable scaling).
- CASSCF/CASPT2 model pretty accurate and robust: model of choice for study of photochemistry and excited-state reactivity.
- Newly developed models based on TDDFT also useful for excited states, but accuracy compromised compared to higher-level correlated quantum chemistry methods, e.g., description of excitations with charge transfer, multireference, or Rydberg character problematic in TD-DFT.
- Insufficient accuracy of all these methods severely limits predictive power; even CASPT2 where short-range correlations treated only perturbatively - in contrast to the essentially exact treatment in QMC. Good-quality DMC calculations potentially very useful. Restoration of missing CASSCF dynamic correlation through DMC [Dubecky *et al.*, vanadium-benzene half-sandwich]. The energy of the VBz^+ system (in eV) is shown as a function of the V^+-Bz distance. We see that the CASSCF calculation gives a very shallow minimum at a substantially larger distance – more than 0.2Å– compared to the more accurate DMC calculation.



Many interesting questions and open problems! See later lecture.

Diamond band structure with single determinant trial function



LDA band structure overlaid on DMC quasiparticle energies from electron-hole pair excitations (and GW results). DMC energy of state at top of valence band is zero by definition, but rest of energies are meaningful.

Low-lying quasiparticle energies very accurate, but calculated energies of hole lying deeper in the valence band significantly overestimated. Probably because fixed-node approximation recovers larger fraction of the correlation energy of ground state than highly excited state.

Main sources of error in DMC

- Incompetence
- Statistical error
- Born-Oppenheimer approximation. Using geometries from band theory. (forces have been difficult to calculate directly in QMC, though the situation is improving).
- Fixed-node error
- Time step error
- Pseudopotentials in general
- Localization approximation (non-local integration of pseudopotentials)
- Possible ‘spin contamination’

Conclusions

- VMC using Slater-Jastrow wave functions with $\sim$30 variational parameters can recover between 70 and 85% of the valence correlation energy, and DMC calculations can recover roughly 95% plus. The remaining 5% is largely due to fixed-node error.
- In solids, QMC is the *only* practical method based on many-body correlated wave functions, the variational principle, and the many-electron Schrödinger equation. It is now the method of choice for tackling large quantum many-body problems.
- Efficient implentations of VMC and DMC for finite and periodic systems have been made in our computer program CASINO. Much remains to be done to make QMC as flexible and easy to use as traditional methods.
- With its emphasis on many-electron wave functions and probabilities, QMC has shown that it is possible to study interacting electrons in real solids (as complicated as doped lanthanum cuprate superconductors in recent years) using very direct computational techniques; no need to resort to perturbation theory or mean-field approximations. Medium to large computers generally required however..

Download CASINO : `vallico.net/casinoqmc`

Coffee break