

# Overview of Continuum Quantum Monte Carlo Methods for One-Dimensional Systems



**Neil D. Drummond**

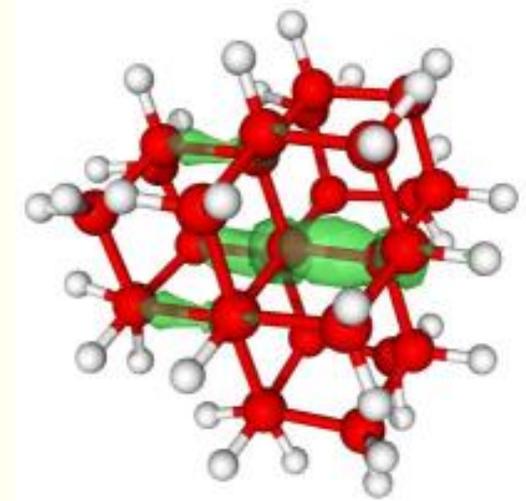
**Department of Physics, University of Lancaster, UK**

*Correlations and transport in one-dimensional structures, TU Dresden, Germany*

Friday 5th July, 2019

# Electronic Structure Calculation: the Challenge

- Suppose you want to calculate the atomic structure that a material adopts or the equation of state or the phonon modes or the electronic band structure or . . .
  - Work within the [Born–Oppenheimer](#) approximation.
  - Left with the nontrivial problem of determining the electronic ground-state (or an excited-state) energy for a given set of nuclear positions.



- [Nonrelativistic many-electron Hamiltonian](#) (first quantisation):

$$\hat{H} = \sum_i -\frac{1}{2}\nabla_i^2 - \sum_i \sum_I \frac{Z_I}{r_{iI}} + \sum_{i>j} \frac{1}{r_{ij}} + \sum_{I>J} \frac{Z_I Z_J}{r_{IJ}}.$$

- Repulsion between electrons couples their behaviour. Makes life difficult/interesting.
- Similar Hamiltonians arise for [homogeneous](#) and [inhomogeneous](#) models of [interacting charge carriers in semiconductors](#), moving in 1D, 2D or 3D. Or [cold atomic gases](#). Or [nuclear matter](#). Etc. *Interactions need not be of Coulomb form.*

# Variational Principle

- Suppose we have a trial many-electron spatial wave function  $\Psi(\mathbf{R})$ .
- **Variational principle:** Let  $E_0$  be the ground state energy. Then

$$\frac{\int \Psi^* \hat{H} \Psi d\mathbf{R}}{\int |\Psi|^2 d\mathbf{R}} \geq E_0.$$

- Evaluating the integrals is difficult in general, because they are  $3N$ -dimensional in a 3D system ( $N$  dimensional in a 1D system).
- Unless you happen to be interested in things like the hydrogen atom, you must either
  - (i) restrict the form of the wave function so that the  $3N$ -dimensional integration can be broken into a series of 3D or 6D integrals, or
  - (ii) use **Monte Carlo integration**.

# Variational Monte Carlo

- Expectation value of Hamiltonian:

$$\frac{\int \Psi^* \hat{H} \Psi d\mathbf{R}}{\int |\Psi|^2 d\mathbf{R}} = \frac{\int |\Psi|^2 \frac{\hat{H}\Psi}{\Psi} d\mathbf{R}}{\int |\Psi|^2 d\mathbf{R}} = \langle E_L \rangle,$$

where

$$E_L = \frac{\hat{H}\Psi}{\Psi}$$

is called the **local energy**.

- **VMC**: use the Metropolis algorithm to generate sets of electron coordinates distributed as  $|\Psi|^2$  and average the local energies to obtain an estimate of the energy.
  - **Zero-variance principle**: variance of the local energy goes to zero as trial wave function improves, becoming zero when  $\Psi$  is an eigenfunction of  $\hat{H}$ .
- Evaluate expectation values of other operators  $\hat{A}$  in a similar fashion (simultaneously) by averaging  $\Psi^{-1} \hat{A} \Psi$ .

# Scaling

- Variance of local energies increases as  $N$ : think about adding the energy contributions of distant, largely uncorrelated molecules.
- So number of samples required to achieve a given error bar increases as  $N$ .
- The time taken to evaluate the wave function (see later) increases as  $N^2$ .
- So the cost of achieving a given accuracy increases as  $N^3$ .
- Same as scaling of standard implementations of rival methods, but with much larger prefactor (typically 100–10,000 times larger than DFT).



# Metropolis–Hastings Algorithm (I)

- If you are familiar with using the Metropolis algorithm to sample the Boltzmann distribution, the algorithm is very similar in quantum Monte Carlo.
- At each step, propose a **trial move** of the particles drawn from the proposal probability density  $T(\mathbf{R} \rightarrow \mathbf{R}')$ .
- **Accept** the move with probability

$$A(\mathbf{R} \rightarrow \mathbf{R}') = \min \left\{ 1, \frac{T(\mathbf{R}' \rightarrow \mathbf{R})|\Psi(\mathbf{R}')|^2}{T(\mathbf{R} \rightarrow \mathbf{R}')|\Psi(\mathbf{R})|^2} \right\}.$$

- Moves to regions of lower  $|\Psi|^2$  are like “uphill moves” when sampling the Boltzmann distribution and moves to regions of larger  $|\Psi|^2$  are like “downhill moves”.

## Metropolis–Hastings Algorithm (II)

- Let  $P_s(\mathbf{R})$  be the distribution of configurations at step  $s$ .

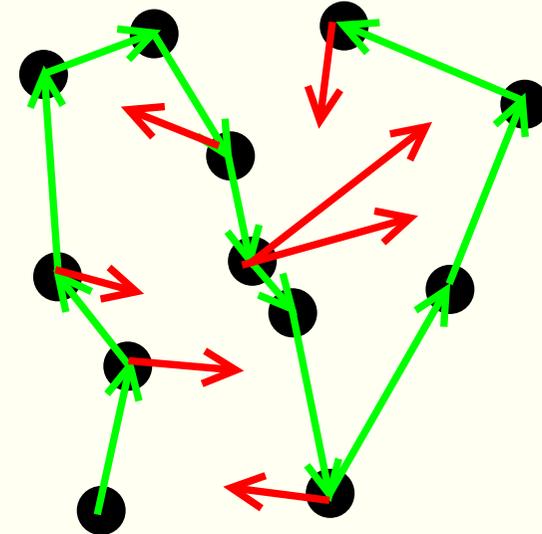
- It is easy to verify that the **detailed balance** condition is satisfied:

$$|\Psi(\mathbf{R})|^2 T(\mathbf{R} \rightarrow \mathbf{R}') A(\mathbf{R} \rightarrow \mathbf{R}') = |\Psi(\mathbf{R}')|^2 T(\mathbf{R}' \rightarrow \mathbf{R}) A(\mathbf{R}' \rightarrow \mathbf{R}).$$

- Hence  $P_s(\mathbf{R}) = |\Psi(\mathbf{R})|^2$  is a stationary distribution of the process, meaning that  $P_{s+1}(\mathbf{R}') = \int P_s(\mathbf{R}) T(\mathbf{R} \rightarrow \mathbf{R}') A(\mathbf{R} \rightarrow \mathbf{R}') d\mathbf{R} = P_s(\mathbf{R}')$ .
- The Metropolis algorithm converges to this stationary distribution provided the transition probability density is **ergodic** (any configuration can be reached from any other with finite probability in a finite number of moves).
- We normally use a **Gaussian** probability density  $T(\mathbf{R} \rightarrow \mathbf{R}')$ .
- We use the electron-by-electron algorithm, in which electrons are moved one at a time, in turn.

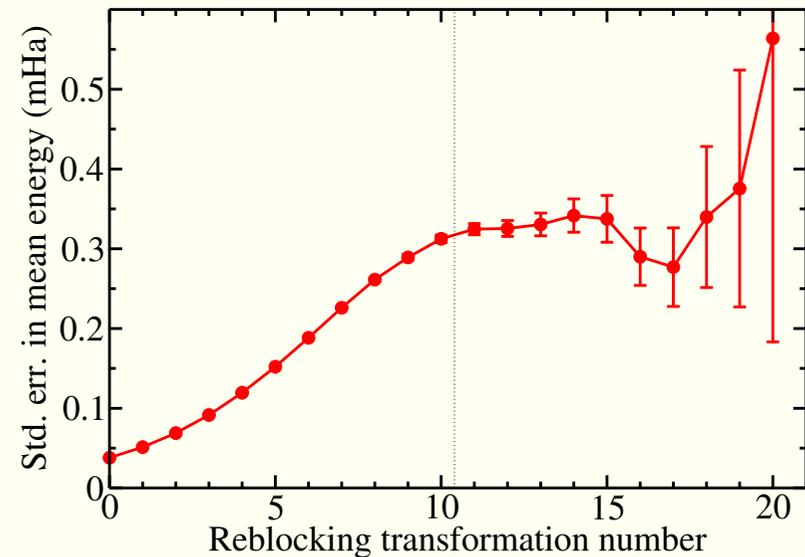
## Serial Correlation

- Although the configurations generated along the random walk are distributed as  $|\Psi|^2$ , they are **serially correlated**: (i) proposed moves are of finite range and (ii) some proposed moves are rejected, resulting in successive configurations that are identical.
- The resulting energy data are serially correlated.
- A naïve calculation of the standard error in the mean energy as the standard deviation divided by the square root of the number of data points sampled **underestimates** the error bar.
- In effect, serial correlation implies that the number of independent data points is less than the number of actual data points sampled.



# Reblocking Analysis

- A simple and robust way of addressing this problem is **reblocking analysis**:
  - *Successive data points are averaged in pairs, producing a new data set with half as many data points as the original, but with a lower variance and weaker serial correlation (and an unchanged mean).*
  - *The naïve expression for the standard error in the mean is computed using the new data set.*
  - *The process is then repeated.*
- Once the block length exceeds the longest correlation period, the reblocked data points are independent and the naïve expression gives the correct standard error.
- The signature of this having happened is that the standard error does not change significantly on further reblocking transformations.



## Wave-Function Forms

- First stage of any VMC or DMC calculation: *generate a trial wave function*.
- Basic approximation comes from a density-functional theory (DFT) or Hartree–Fock (HF) code (e.g. CASTEP or Gaussian) or is known on theoretical grounds (e.g. plane-wave orbitals for the fluid phase of an electron gas).
  - “Typically” one might retrieve about 98% of the total energy using HF theory.
  - The remaining 2% (*correlation energy*) is very significant for chemistry, however.
- Parameters in explicitly correlated wave function are optimised within QMC.
  - *Most difficult technical aspect of QMC.*
  - But. . . **Good wave functions are vital for successful QMC work.**
- Usually we are interested in **spin-independent Hamiltonians**, so we are looking for a solution to the Schrödinger equation that is an eigenfunction of spin operator  $\hat{S}_z$ .
  - Hence we can work with a **spatial wave function**  $\Psi(\mathbf{R})$  for  $N_\uparrow$  spin-up electrons and  $N_\downarrow$  spin-down electrons that is only antisymmetric under exchange of the spatial coordinates of the same-spin electrons.

# Slater–Jastrow Wave Functions

- Most QMC calculations use *Slater–Jastrow* trial wave functions:

$$\Psi(\mathbf{R}) = \exp[J(\mathbf{R})] \sum_n c_n D_n^\uparrow(\mathbf{R}) D_n^\downarrow(\mathbf{R}),$$

where  $D^\uparrow$  and  $D^\downarrow$  are **Slater determinants** for spin-up and down electrons, and  $\exp(J)$  is a **Jastrow factor**.

- **Time-reversal symmetry**: Wave function can always be chosen to be real for a real Hamiltonian with appropriate boundary conditions.
  - Either real or complex wave functions can be used in QMC calculations. Former are faster, but latter are necessary if “twisted” boundary conditions are used.
- Wave function must be as cheap as possible to evaluate, and should not contain an exponentially scaling number of parameters.
- **Compact description of correlation**: number of free parameters to retrieve a given fraction of the correlation energy increases as a (low-degree) polynomial of  $N$ .

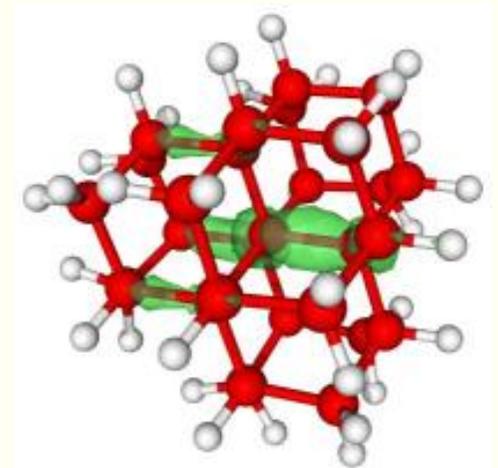
# Slater Wave Function

- Slater wave function is our first approximation to the many-electron wave function.
- Each Slater determinant is of the form

$$D^\uparrow(\mathbf{R}) = \begin{vmatrix} \psi_1^\uparrow(\mathbf{r}_1) & \cdots & \psi_{N_\uparrow}^\uparrow(\mathbf{r}_1) \\ \vdots & & \vdots \\ \psi_1^\uparrow(\mathbf{r}_{N_\uparrow}) & \cdots & \psi_{N_\uparrow}^\uparrow(\mathbf{r}_{N_\uparrow}) \end{vmatrix},$$

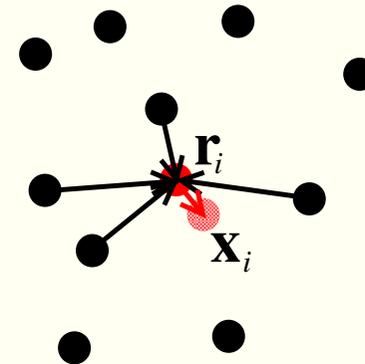
and similarly for spin-down determinants.

- Orbitals  $\{\psi_i^\sigma\}$  are usually generated in either DFT or HF calculations.
  - It is sometimes possible to optimise orbitals within QMC.
  - Orbitals may be represented in different ways (plane waves, Gaussians, “blips”, . . . ).



# Multideterminant and Backflow Wave Functions

- Multideterminant expansion coefficients  $\{c_n\}$  can be taken from post-HF calculations, or can be optimised within QMC.
- Most QMC calculations use only one term in the expansion: the HF wave function.
- Instead of evaluating Slater wave function at actual electron coordinates  $\mathbf{R}$ , one can evaluate it at *quasiparticle coordinates*  $\mathbf{X}(\mathbf{R})$ .
- Quasiparticle coordinates are related to the original coordinates by a *backflow transformation*, which depends on all coordinates.
  - Backflow function contains free parameters to be optimised.
- *Backflow remains useful in 1D systems: efficient param. of multibody correlations.*



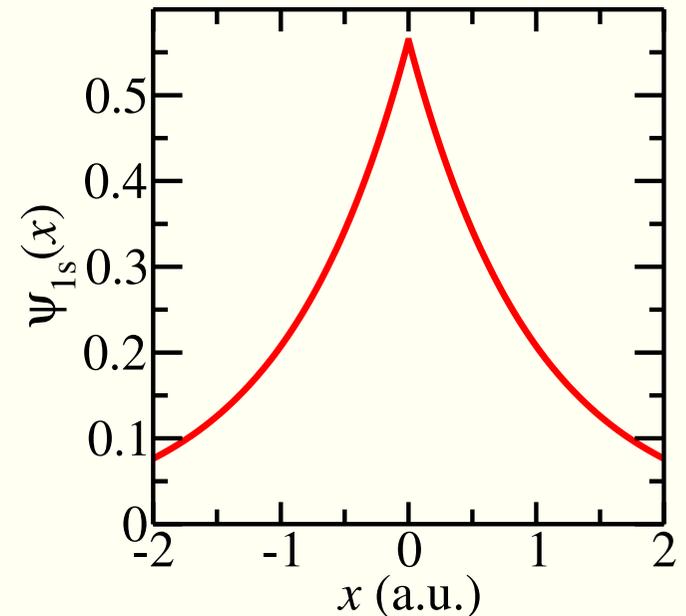
Two-body backflow fn:  $\mathbf{x}_i = \mathbf{r}_i + \sum_{j \neq i} \eta(r_{ij}) \mathbf{r}_{ij}$ , where  $\eta$  is a smoothly truncated polynomial.

# Jastrow Factor

- Jastrow factor  $\exp(J)$  is an explicit function of interparticle distance.
  - Hence Jastrow factor can be compactly parametrised, *unlike e.g. multideterminant calculations.*
  - *Number of parameters in the Jastrow factor required to retrieve a given fraction of correlation energy is essentially independent of the system size.*
- Slater wave function has the required antisymmetry, so  $J$  must be symmetric under the exchange of same-spin electrons.
  - We choose  $J$  to transform as the trivial representation of the symmetry group of the Hamiltonian.
  - All wave function symmetry is determined by the Slater part of the wave function.
  - The Slater wave (in particular the *topology of the nodal surface*) determines the state of the system described by the trial wave function.
  - At long range  $J$  describes the *envelope wave function* for quasiparticles.
- $J$  contains free parameters, to be determined by an optimisation method.

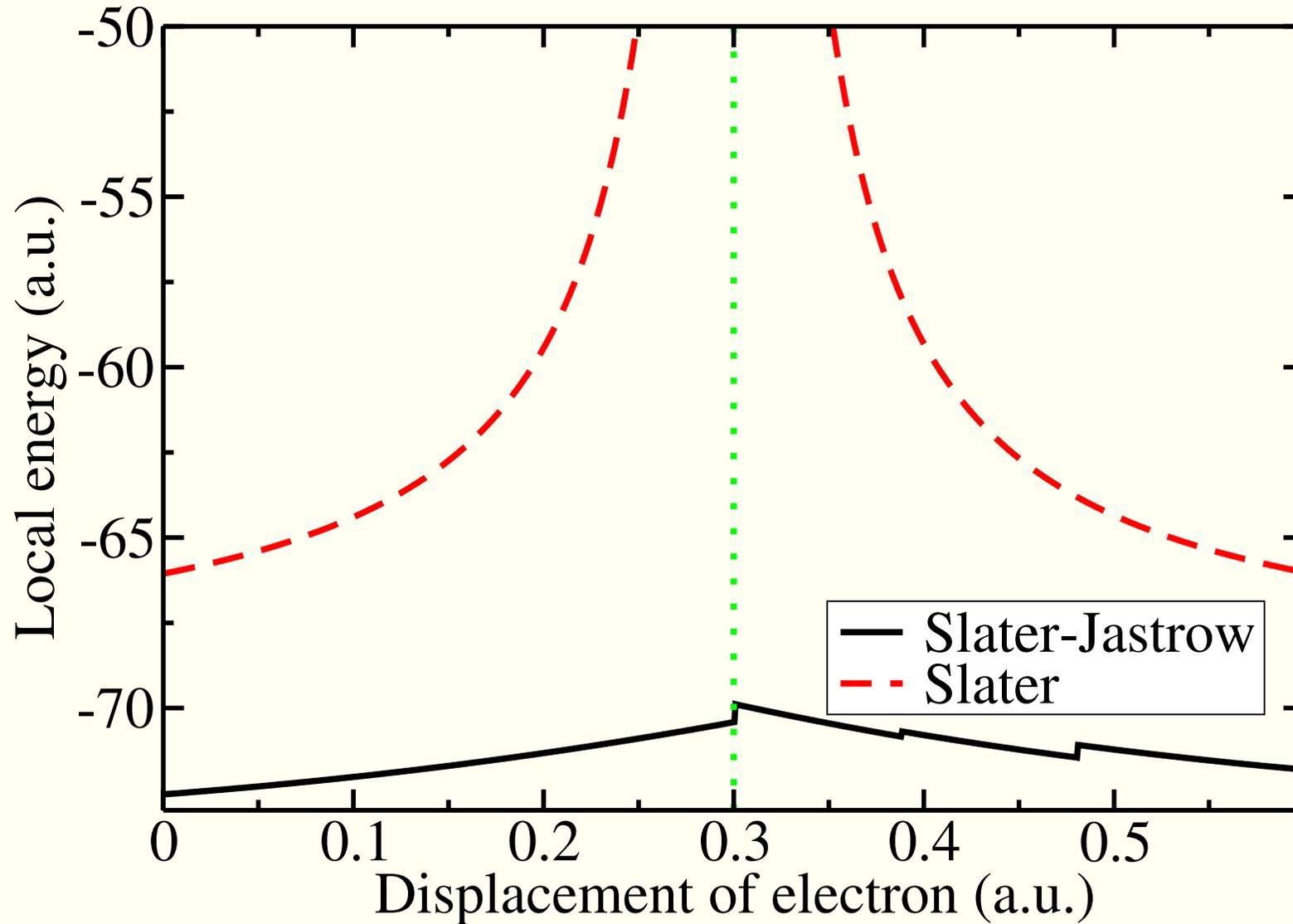
## Kato Cusp Conditions (I)

- If two electrons approach each other then the potential energy diverges as  $1/r_{ij}$ .
- **The local energy is constant for an eigenstate:** so this divergence must be cancelled by an equal and opposite divergence in the local kinetic energy  $-(\nabla^2\Psi)/(2\Psi)$ .
- To get a divergence in the local kinetic energy, the wave function must be cusp-like at electron–electron coalescence points.
- These **cusp conditions** can be imposed by the Jastrow factor.
- Likewise there is an electron–nucleus cusp condition.
- DMC calculations can be unstable if the cusp conditions are not satisfied.



*Hydrogenic 1s orbital*  
 $\psi = (1/\sqrt{\pi}) \exp(-r)$  clearly satisfies Kato cusp condition.

## Kato Cusp Conditions (II)



*Local energy as one electron is moved straight through another of opposite spin in  $H_2O$ . Jastrow factor imposes Kato cusp condition.*

# Coalescence Conditions and Wave Functions in 1D

- Consider a strictly 1D system interacting via the Coulomb interaction  $v(x) = 1/|x|$ .
  - In 1D there is no divergent term in the Laplacian at short range, so the local kinetic energy cannot cancel the divergent potential energy at divergences.
  - Hence the wave function must go to zero at all coalescence points, for both parallel and antiparallel spins.
  - Impose this via the Slater wave function: choose the Slater wave function to be antisymmetric under all particle exchanges.
- Now suppose the interaction is nondivergent at coalescent points (modelling finite-width effects), e.g.,  $v(x) = (x^2 + d^2)^{-1/2}$  where  $d$  is small.
  - Then the wave function has “pseudonodes” at antiparallel-spin coalescence points.
  - Antiparallel-spin electrons occasionally swap positions during a simulation.
  - Leads to ergodicity problems (very long equilibration and correlation times).

↑↑↑↑↑↓↓↓↓↓

(long time)

→

↑↑↑↑↓↑↓↓↓↓

## Optimisation: Variance Minimisation

- Suppose the trial wave function happens to be an eigenstate of the Hamiltonian with eigenvalue  $E_i$ . Then the local energy is

$$E_L(\mathbf{R}) = \frac{\hat{H}\Psi}{\Psi} = E_i,$$

which is **constant** throughout configuration space.

- So, start with an initial wave function (e.g., the Hartree-Fock wave function) and:
  - Generate a set of configurations  $\{\mathbf{R}\}$  distributed as  $|\Psi|^2$ .
  - Adjust the parameters in the wave function so as to minimise the variance of the local energies of the configuration set  $\{E_L(\mathbf{R})\}$ .
  - Repeat until the process has converged.
- *If wave function form can describe an eigenfunction then this has zero variance and will be found by minimisation of the variance. If not, we get something which is close to an eigenfunction.*

# Imaginary-Time Schrödinger Equation (I)

- *Imaginary-time Schrödinger equation (ITSE):*

$$\left[ \hat{H} - E_T \right] \Phi = -\frac{1}{2} \nabla^2 \Phi + U \Phi - E_T \Phi = -\frac{\partial \Phi}{\partial t},$$

where  $\Phi(\mathbf{R}, t)$  is a function of configuration  $\mathbf{R}$  and imaginary time  $t$ ,  $U(\mathbf{R})$  is potential energy and  $E_T$  is a *reference energy*.

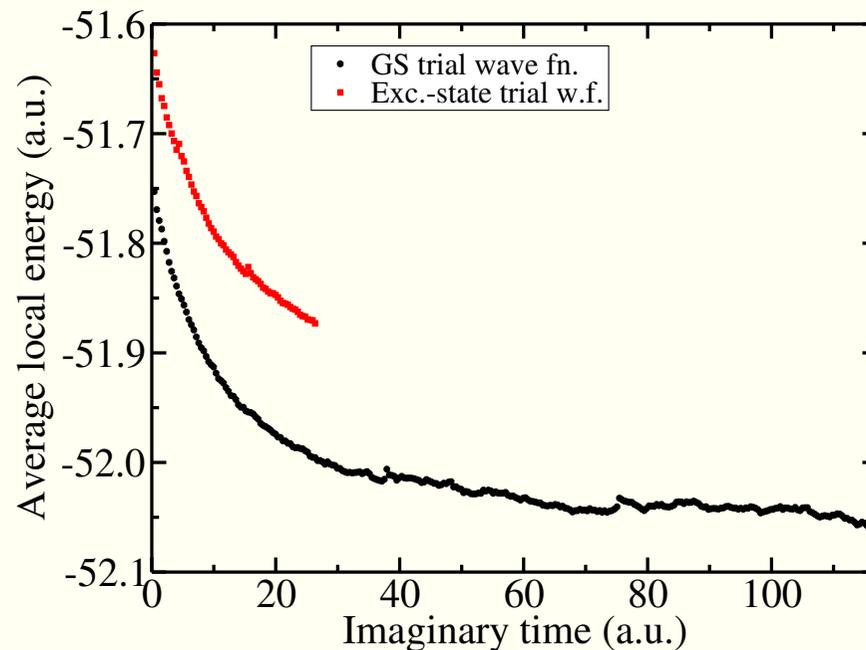
- Time-dependence of ITSE is separable. Can write

$$\Phi = \sum_{n=0}^{\infty} c_n \phi_n \exp[-(E_n - E_T)t],$$

where  $E_n$  and  $\phi_n(\mathbf{R})$  are  $n$ th eigenvalue and eigenfunction of Hamiltonian  $\hat{H}$ .

## Imaginary-Time Schrödinger Equation (II)

- Excited states die away exponentially compared with ground state.
- If  $E_T = E_0$  and initial conditions have  $c_0 \neq 0$  then, in limit that  $t \rightarrow \infty$ ,  $\Phi$  is proportional to  $\phi_0$ .
- Ground-state component of  $\Phi$  is “projected out”.
- This is true for any reasonable boundary conditions on  $\Phi$ : see later.



# Importance-Sampling Transformation

- Suppose we have a **trial wave function**  $\Psi(\mathbf{R})$ . Let  $f = \Phi\Psi$ . Then

$$-\frac{1}{2}\nabla^2 f + \nabla \cdot [\mathbf{V}f] + [E_L - E_T] f = -\frac{\partial f}{\partial t},$$

where

- $f(\mathbf{R}, t) = \Phi\Psi$  is the *importance-sampled* or *mixed wave function*,
  - $\mathbf{V}(\mathbf{R}) = \Psi^{-1}\nabla\Psi$  is the *drift velocity*,
  - $E_L(\mathbf{R}) = \Psi^{-1}\hat{H}\Psi$  is the *local energy*.
- Proof: substitute  $\Phi = \Psi^{-1}f$  into ITSE.
  - Consequences of importance sampling:
    1. Term in ITSE involving potential  $U(\mathbf{R})$  is replaced by term involving local energy  $E_L(\mathbf{R})$ , which is relatively uniform. *Makes branching DMC algorithm stable.*
    2. Configurations are distributed according to  $f = \Phi\Psi$  rather than  $|\Phi|$ .
    3. *Fixed-node approximation* is introduced (see later).

# Importance-Sampled Imaginary Time Schrödinger Equation

- Without the last term on the LHS, ISITSE is a **Fokker–Planck** equation:

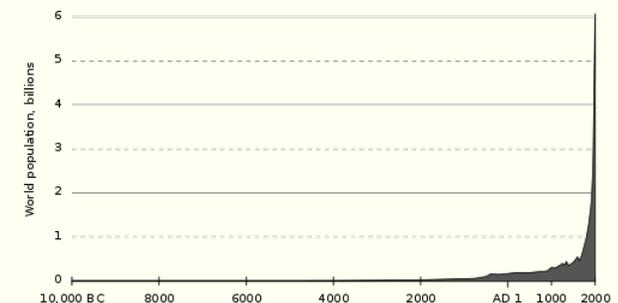
$$-\frac{1}{2}\nabla^2 f + \nabla \cdot [\mathbf{V} f] = -\frac{\partial f}{\partial t}.$$

- This equation describes the time-evolution of the density of a set of “particles” undergoing random diffusion in a  $3N$ -dimensional fluid of velocity field  $\mathbf{V}$ .

- Without the first two terms on the LHS, ISITSE is a **rate** equation:

$$[E_L - E_T] f = -\frac{\partial f}{\partial t}.$$

- Gives an exponential increase or decrease in the density of “particles” at each point in the  $3N$ -dimensional configuration space.



## Green's Function (I)

- Green's function  $G(\mathbf{R} \leftarrow \mathbf{R}', t)$  is solution to differential equation satisfying initial condition  $G(\mathbf{R} \leftarrow \mathbf{R}', 0) = \delta(\mathbf{R} - \mathbf{R}')$ .
- Integral form of ISITSE:

$$f(\mathbf{R}, t + \tau) = \int f(\mathbf{R}', t) G(\mathbf{R} \leftarrow \mathbf{R}', \tau) d\mathbf{R}'.$$

- Formally,  $G(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \langle \mathbf{R} | e^{-\tau(\hat{F} + E_L - E_T)} | \mathbf{R}' \rangle$ , where  $\hat{F}$  is the Fokker–Planck operator (operator on LHS of Fokker–Planck equation).
- Use the Trotter–Suzuki formula:

$$\begin{aligned} G(\mathbf{R} \leftarrow \mathbf{R}', \tau) &= \langle \mathbf{R} | e^{-\tau(E_L - E_T)/2} e^{-\tau\hat{F}} e^{-\tau(E_L - E_T)/2} | \mathbf{R}' \rangle + \mathcal{O}(\tau^3) \\ &= e^{-\tau[E_L(\mathbf{R}) + E_L(\mathbf{R}') - 2E_T]/2} \langle \mathbf{R} | e^{-\tau\hat{F}} | \mathbf{R}' \rangle + \mathcal{O}(\tau^3) \\ &\equiv G_B(\mathbf{R} \leftarrow \mathbf{R}', \tau) G_D(\mathbf{R} \leftarrow \mathbf{R}', \tau) + \mathcal{O}(\tau^3). \end{aligned}$$

## Green's Function (II)

- Green's function for Fokker–Planck equation (Langevin Green's function):

$$G_D(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \frac{1}{(2\pi\tau)^{3N/2}} \exp\left(-\frac{|\mathbf{R} - \mathbf{R}' - \tau\mathbf{V}(\mathbf{R}')|^2}{2\tau}\right) + \mathcal{O}(\tau^2).$$

- Hence Green's function for ISITSE:

$$\begin{aligned} G(\mathbf{R} \leftarrow \mathbf{R}', \tau) &= G_D(\mathbf{R} \leftarrow \mathbf{R}', \tau)G_B(\mathbf{R} \leftarrow \mathbf{R}', \tau) + \mathcal{O}(\tau^2) \\ &= G_{\text{DMC}}(\mathbf{R} \leftarrow \mathbf{R}', \tau) + \mathcal{O}(\tau^2). \end{aligned}$$

- DMC Green's function is exact in limit of small time steps  $\tau$ .

## Propagation of Configuration Population (I)

- At any given moment in a DMC simulation,  $f$  is represented by a population of “configurations” or “walkers” in configuration space:

$$f(\mathbf{R}, t) = \left\langle \sum_{\alpha} w_{\alpha} \delta(\mathbf{R} - \mathbf{R}_{\alpha}) \right\rangle,$$

where  $\mathbf{R}_{\alpha}$  is the position of configuration  $\alpha$  and  $w_{\alpha}$  is its weight and the angled brackets are an ensemble average.

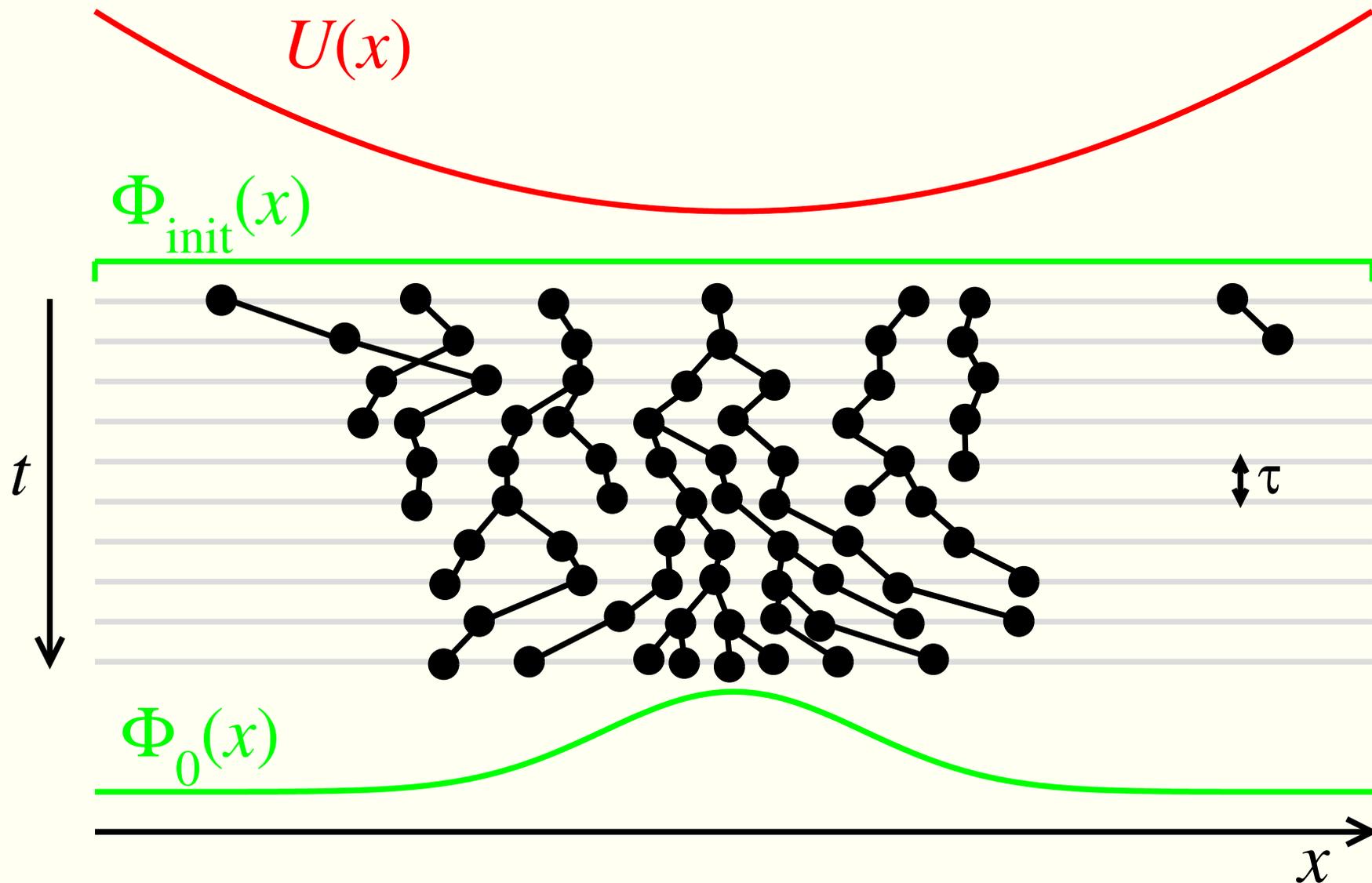
- Ensemble averaging commutes with differentiation, etc.
- Substitute above expression for  $f$  into integral form of ISITSE to find distribution of configurations one time step  $\tau$  later:

$$f(\mathbf{R}, t + \tau) = \left\langle \sum_{\alpha} w_{\alpha} G_{\text{B}}(\mathbf{R} \leftarrow \mathbf{R}_{\alpha}, \tau) G_{\text{D}}(\mathbf{R} \leftarrow \mathbf{R}_{\alpha}, \tau) \right\rangle.$$

## Propagation of Configuration Population (II)

- Green's functions can be treated as **transition-probability densities**; ensemble average has correct behaviour.
  - Configurations drift by  $\tau \mathbf{V}(\mathbf{R}_\alpha)$  and diffuse (are displaced by a random vector, Gaussian-distributed with variance  $\tau$ ).
  - Branching factor is absorbed into a new weight.
- First allow excited-state components of  $\Phi$  to die away: *equilibration phase*.
- Then continue to propagate configurations, but gather energy data: *statistics-accumulation phase*.
- Finite time-step error is introduced into  $f$  at rate  $\mathcal{O}(\tau)$ , but is removed at a rate proportional to the error in  $f$ .
  - Hence in equilibrium the finite time-step error in  $f$  is linear in the time step  $\tau$ .

## Propagation of Configuration Population (III)



*Drift, diffusion and branching for a single particle moving in 1D.*

# Wave-Function Antisymmetry

- Want to find **fermionic** (antisymmetric) ground state.
- Lowest-energy wave function is **bosonic** (symmetric) ground state.
- Therefore have to constrain DMC to preserve antisymmetry.
- Constraint is “automatic” in importance-sampled DMC algorithm:
  - If  $\Phi$  and  $\Psi$  have different nodal surfaces, there must exist regions where  $f$  is negative.
  - Our algorithm is based on interpreting  $f$  as a *probability density*.
  - Can never have a negative  $f$  in our algorithm.
  - So we cannot describe a change in the nodal surface of  $\Phi$ .
- *By importance sampling and not permitting weights to become negative, we have introduced the fixed-node approximation.*

## Fixed-Node Approximation (I)

- Nodes of  $\Psi$  divide configuration space into *nodal pockets*.
- Within each nodal pocket  $\lambda$  we in effect solve Schrödinger equation subject to boundary condition that (asymmetric) wave function  $\Phi_\lambda$  is zero outside pocket.
- So  $\hat{H}\Phi_\lambda = E_0^\lambda\Phi_\lambda + \delta_\lambda$ , where  $\delta_\lambda$  are  $\delta$  functions at pocket boundary arising from discontinuity of derivative of  $\Phi_\lambda$  and  $E_0^\lambda$  is *pocket energy*.
- Consider antisymmetric wave function  $\tilde{\Phi}_\lambda(\mathbf{R}) = \hat{\mathcal{A}}\Phi_\lambda(\mathbf{R}) \equiv \sum_{\hat{\mathcal{P}}} (-1)^p \hat{\mathcal{P}}\Phi_\lambda(\mathbf{R})$ , where  $\{\hat{\mathcal{P}}\}$  are operators that permute like-spin coordinates and  $\{p\}$  are corresponding parities.
- **Variational principle:**

$$E_0^{\text{F}} \leq \frac{\langle \tilde{\Phi}_\lambda | \hat{H} | \tilde{\Phi}_\lambda \rangle}{\langle \tilde{\Phi}_\lambda | \tilde{\Phi}_\lambda \rangle} = \frac{\langle \Phi_\lambda | \hat{\mathcal{A}} \hat{H} \hat{\mathcal{A}} | \Phi_\lambda \rangle}{\langle \Phi_\lambda | \hat{\mathcal{A}}^2 | \Phi_\lambda \rangle} = E_0^\lambda,$$

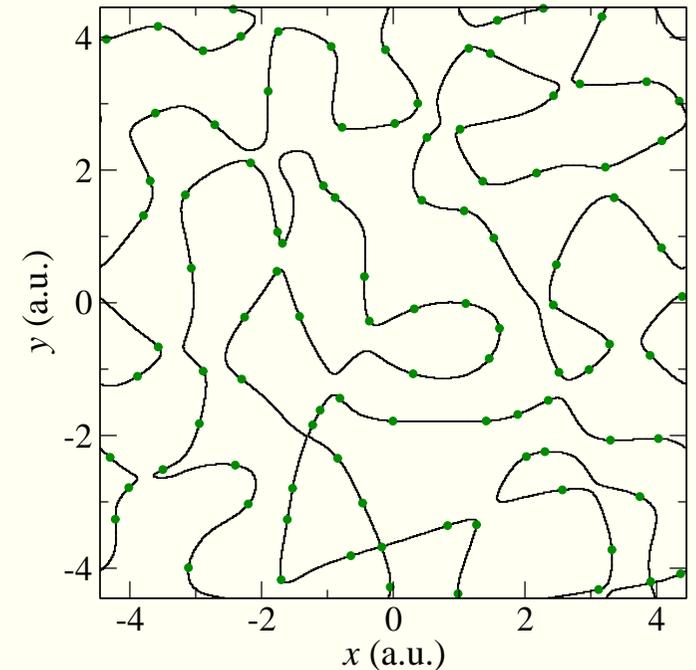
so each pocket energy is greater than the fermion ground-state energy  $E_0^{\text{F}}$ .

## Fixed-Node Approximation (II)

- Have used fact that  $\hat{A}$  is Hermitian and that it commutes with  $\hat{H}$ , and that contribution due to  $\delta_\lambda$  vanishes because  $\tilde{\Phi}_\lambda = 0$  at nodes.
- Within each nodal pocket  $\lambda$ , mixed estimator gives energy  $E_0^\lambda$ .
- Configuration populations in high-energy pockets die out, so DMC energy is  $\min\{E_0^\lambda\} \geq E_0^F$ .
- **Fixed-node DMC energy exceeds fermion ground-state energy, becoming equal in limit that nodal surface is exact.**
- *FNA is only fundamental approximation in DMC.* Error in DMC energy is second order in error in nodal surface.
- Drift velocity diverges at nodal surface, carrying away configurations.
- At finite time steps, configurations can drift/diffuse across surface. Source of time-step bias.

## Fixed-Node Approximation (III)

- FNA with antisymmetric trial wave function enables us to approximate lowest-energy antisymmetric eigenstate.
  - Likewise, FNA with trial wave function of any given symmetry enables us to approximate lowest-energy state with that symmetry.
- Always the case that DMC gives the energy of an excited state exactly if the nodal surface is exact for that state.
  - Hence we can use fixed-node DMC to calculate excited-state energies by using an appropriate trial wave function.
  - Variational principle does not hold for excited-state energies in general.
- **Ground state of a 1D system:** Nodal “surface” is the set of coalescence points; nodal surface of trial wave function is exact; no fixed-node error!



*2D slice intersecting with the 107-dimensional nodal hypersurface of a 54-electron 2D electron gas.*

## Mixed Estimator of the Energy

- Fixed-node ground state satisfies  $\hat{H}\phi_0 = E_0\phi_0$ . Hence

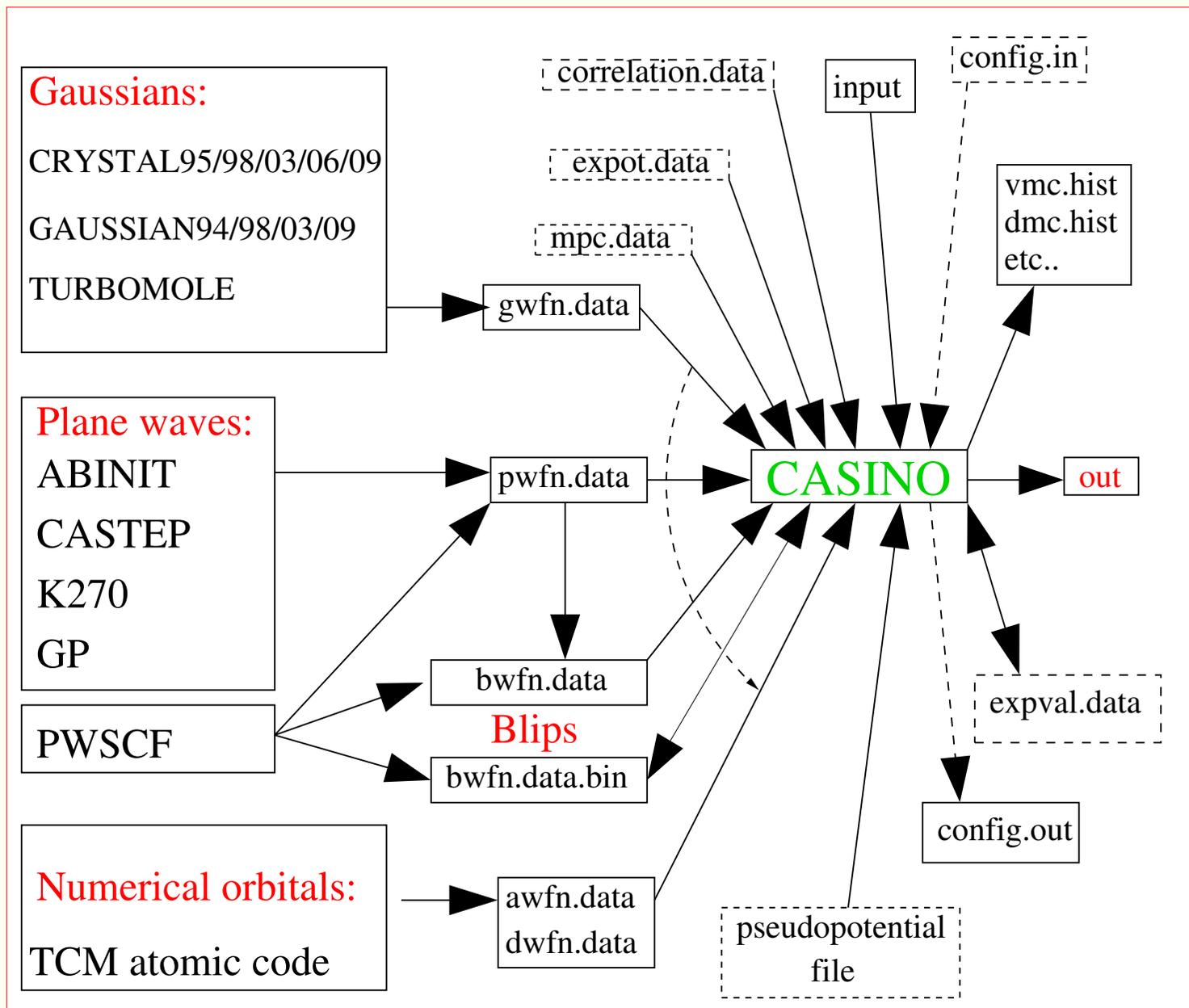
$$E_0 = \frac{\langle \phi_0 | \hat{H} | \Psi \rangle}{\langle \phi_0 | \Psi \rangle} = \frac{\int \phi_0 \Psi E_L d\mathbf{R}}{\int \phi_0 \Psi d\mathbf{R}} = \langle E_L \rangle_{\phi_0 \Psi}.$$

- After equilibration, DMC configuration population has the **mixed** distribution  $\phi_0 \Psi$ .
- So, average the local energies (with weights) of the configurations generated in the statistics accumulation phase to obtain an estimate of  $E_0$ .
- In unweighted DMC, branching factor is used in place of weights.
- At each iteration, local energies are averaged over configuration population (weighted by branching factors).
- $\mathcal{O}(\tau)$  error in mixed distribution gives  $\mathcal{O}(\tau)$  time-step bias in mixed estimator.

# CASINO QMC Software

- VMC and DMC for 1D, 2D and 3D systems, with periodicity in 0, 1, 2 or 3 dimensions.
  - Wave-function optimisation by variance or energy minimisation.
  - Model systems: electron[–hole] gases, excitonic molecules, cold atomic gases, . . .
- Slater[–Jastrow[–backflow]] trial wave functions. The Slater part may consist of (spin-polarised) multiple determinants. Pairing (geminal) wave functions.
  - Basis functions: plane waves, blips, atom-centred Gaussians [with cusp corrections] and Slater functions. Numerical orbitals for atoms and molecular dimers.
  - Excited states from promotion, addition or subtraction of particles.
- Periodic interactions with Ewald or model periodic Coulomb interactions.
- Various expectation values: energy components, charge & spin density, pair correlation function, structure factor, density matrices, dipole moment, momentum density, . . .
- Portability: requires just Fortran and C compilers, MPI library and optionally BLAS and LAPACK. MPI parallelism over walkers; OpenMP parallelism over electrons.

# Interfaces to Other Codes



# Computational Expense

- “Heroic” QMC calculations have more than 1,000 electrons.
- Methods are “embarrassingly parallel”: i.e., require little interprocessor communication on parallel computers.
  - VMC involves no interprocessor communication apart from setup and averaging.
  - DMC requires a small amount of communication: setting of reference energy, and load-balancing transfers of configuration population.
  - CASINO scales to tens of thousands of cores.
- Methods are not as black-box-like as DFT, especially wave-function optimisation.
- **Common question: “how long is a QMC calculation?”**
  - Depends on **system size** and **target precision** (error bar  $\sim 1/\sqrt{\text{expense}}$ ).
  - Handwaving, approximate answer: “Computational effort is 100–10,000 times greater than the corresponding DFT calculation.”



## Some QMC References

- Useful (but old) book: B. L. Hammond, W. A. Lester, Jr. and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific (1994).
- General overview: W. M. C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- More recent general overview: R. J. Needs, M. D. Towler, N. D. Drummond and P. López Ríos, *J. Phys.: Condens. Matter* **22**, 023201 (2010).
- CASINO website:

<https://vallico.net/casinoqmc/>