# QMC and the CASINO program: QMC exam answers

*These are sample answers for the QMC exam.*
*The criteria for awarding points is given at the end of each question in italics.*

**1.**

(i) We can expand $\Phi$ at imaginary time $t$ as

$$\Phi(\mathbf{R},t) = \sum_{n=0}^{\infty} c_n(t)\Phi_n(\mathbf{R}) \,, \tag{1}$$

where $\Phi_n$ is the $n$th eigenfunction of $\hat{H}$ and $c_n(t)$ are unknown functions. Substituting this expression into the imaginary-time-dependent Schrödinger equation gives

$$\sum_{n=0}^{\infty} \left[ (E_n - E_T)\, c_n(t) + \frac{\partial c_n(t)}{\partial t} \right] \Phi_n(\mathbf{R}) = 0 \,, \tag{2}$$

where $E_n$ is the eigenvalue of the $n$th eigenstate of $\hat{H}$. Since eigenfunctions of the Hamiltonian are orthogonal to each other, we can equate the contents of the square brackets to zero, and the solution of the resulting differential equation is

$$c_n(t) = \exp\left[-\left(E_n - E_T\right) t\right] c_n(0) \,, \tag{3}$$

and therefore we can write

$$\Phi(\mathbf{R},t) = \sum_{n=0}^{\infty} \exp\left[-\left(E_n - E_T\right) t\right] c_n(0)\Phi_n(\mathbf{R}) \,. \tag{4}$$

The exponential $\exp\left[-\left(E_n - E_T\right) t\right]$ will die away exponentially if $(E_n - E_T) > 0$, i.e., if $E_n > E_T$. If $E_T$ is set to $E_0$, the $n > 0$ components of $\Phi$ will die away exponentially, leaving only the ground state.

*1 point given for writing Eq. 4*

*2 points given for indicating behaviour of exponentials*

(ii.1) Substituting $f = \Phi\Psi$ into

$$-\frac{1}{2}\nabla^2 f + \nabla\cdot\left(\Psi^{-1}\nabla\Psi f\right) + \left(\Psi^{-1}\hat{H}\Psi - E_T\right) f = -\frac{\partial f}{\partial t} \tag{5}$$

gives

$$-\frac{1}{2}\nabla^2\left(\Phi\Psi\right) + \nabla\cdot\left(\Phi\nabla\Psi\right) + \left(\hat{H}\Psi - E_T\Psi\right)\Phi = -\Psi\frac{\partial\Phi}{\partial t} \,, \tag{6}$$

where we have made use of the fact that $\Psi$ is independent of time. Noting that $\nabla^2(\Phi\Psi) = \Psi\nabla^2\Phi + 2\nabla\Phi\cdot\nabla\Psi + \Phi\nabla^2\Psi$, that $\nabla\cdot(\Phi\nabla\Psi) = \Phi\nabla^2\Psi + \nabla\Phi\cdot\nabla\Psi$, and that $\hat{H} = -\frac{1}{2}\nabla^2 + V$, and simplifying the expression we arrive at

$$-\frac{1}{2}\Psi\nabla^2\Phi + V\Phi\Psi - E_T\Phi\Psi = -\Psi\frac{\partial\Phi}{\partial t} \,. \tag{7}$$

$\Psi$ can be factored out of the equation and eliminated, giving the imaginary-time Schrödinger equation,

$$\left(\hat{H} - E_T\right)\Phi = -\frac{\partial\Phi}{\partial t} \,, \tag{8}$$

as required.

*1 point given for attempting the derivation*

*2 points given for correctly substituting $f = \Phi\Psi$ and manipulating the resulting expression*

*1 point given for correctly simplifying the equation and completing the derivation*

(ii.2)

* Introduces the fixed-node approximation.
* More adequately distributes DMC configurations according to $\Phi\Psi$ rather than $|\Phi|$
* Reduces the fluctuations of the exponent of the branching factor because the local energy is better-behaved than the potential.

*2 points given for "introduces the fixed-node approximation"*

*1 point for any other valid point*

**2.**

(i) Because the quality of DMC results is only dependent on the quality of the nodes of the wave function.

*3 points*

(ii)

* Backflow transformations
* Orbital optimization
* Multi-determinant expansions

*2 points for any valid answer*

*1 point for any other valid answer*

(iii) (c)

*2 points*

(iv) Equating $\Psi_{\rm HF} = 0$ we find that either

$$\phi_1(r_1)\phi_2(r_2) - \phi_1(r_2)\phi_2(r_1) = 0 \,, \tag{9}$$

or

$$\phi_1(r_3)\phi_2(r_4) - \phi_1(r_4)\phi_2(r_3) = 0 \tag{10}$$

must hold. Given that we are asked for the nodes as a function of $r_1$, we discard the second possibility, and are left with

$$\phi_1(r_1)\phi_2(r_2) = \phi_1(r_2)\phi_2(r_1) \,, \tag{11}$$

which is satisfied if $r_1 = r_2$. That is, the shape of the node seen by particle 1 is a sphere of radius $r_2$ around the nucleus.

A multi-determinant expansion can be expressed as

$$\Psi_{\rm MD}(\mathbf{R}) = \Psi_{\rm HF}(\mathbf{R}) + F(\mathbf{R}) \,, \tag{12}$$

whose nodes are given by $\Psi_{\rm HF}(\mathbf{R}) + F(\mathbf{R}) = 0$. Knowledge of the region where $\Psi_{\rm HF}(\mathbf{R}) = 0$ does not help in finding the region where $\Psi_{\rm MD}(\mathbf{R}) = 0$; the nodes of the HF and MD wave functions can be entirely different.

A backflow transformation represents a smooth deformation of configuration space, and, by construction, the nodes of $\Psi_{\rm BF}$ are close to those of $\Psi_{\rm HF}$.

*1 point for stating that the nodes are given by $r_1 = r_2$, or that they are spherical*

*1 point for stating that backflow produces small deformations of the nodes while additional determinants change the nodes in a deeper manner*

**3.**

(i) Cusp conditions are conditions imposed on the wave function such that the kinetic energy cancels the divergences in the potential energy when two electrons or an electron and a nucleus coincide in space.

The Kato cusp condition is

$$\left(\frac{1}{\Psi}\frac{\partial\Psi}{\partial r_{ij}}\right)_{r_{ij}=0} = \frac{2q_iq_j\mu_{ij}}{2 \pm d} \ , \tag{13}$$

where $q_i$ and $q_j$ are the charges of particles $i$ and $j$, $\mu_{ij} = (1/m_i + 1/m_j)^{-1}$, $d$ is the dimensionality of the system, and the positive sign is used for indistinguishable particles (e.g., parallel-spin electrons), while the negative sign is for distinguishable particles (e.g., anti-parallel-spin electrons, an electron and a positron, an electron and a nucleus, etc.).

It is important to impose the cusp conditions because the statistical fluctuations of the local energy are much reduced as a result. If cusp conditions are applied, the local energy will diverge only at the nodes of the trial wave function, which are rarely sampled. If they are not applied, divergences may occur in regions where the value of the wave-function is finite.

*3 points for describing the efect of cusp conditions and/or giving the form of the Kato cusp condition*

*2 points for mentioning divergences of the local energy*

(ii)

(ii.1) Because the Slater determinant does not explicitly depend on the electron-electron distances.

*3 points*

(ii.2) So that the location of the electron-nucleus cusps enforced by the orbitals is not shifted by the backflow transformation. Such shift would make the kinetic and potential energy diverge at different points and would cause statistical problems.

*2 points for mentioning shift in e-n cusp*

**4.**

(i) The local energy of a configuration $\mathbf{R}$ is

$$E_L(\mathbf{R}) = \frac{\hat{H}(\mathbf{R})\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \ . \tag{14}$$

The Schrödinger equation states that for the $n$th eigenstate of the Hamiltonian,

$$\hat{H}(\mathbf{R})\Phi_n(\mathbf{R}) = E_n\Phi_n(\mathbf{R}) \ . \tag{15}$$

Combining these equations we can see that for the $n$th eigenstate of the Hamiltonian the local energy is independent of $\mathbf{R}$, $E_L(\mathbf{R}) = E_n \ \ \forall \ \mathbf{R}$, and the variance of any sample of constant local energies will be zero.

The variance of the local energy is zero for the ground state of the Hamiltonian, and therefore it is reasonable to assume that minimizing the variance of the local energies associated with a trial wave function $\Psi$ will bring the wave function closer to the ground state.

*4 points for mentioning zero variance of ground state*

(ii) The local energy of a configuration diverges at the nodes of the wave function, since the denominator of $\frac{\hat{H}\Psi}{\Psi}$ goes to zero at the nodes, while the numerator remains finite in general. Optimization within VMC is done using correlated sampling, which means that a set of configurations is generated for an initial set of parameters $\boldsymbol{\alpha}_0$, and they are kept fixed as the parameters are changed.

When parameters are changed, the wave function value at one of these fixed configurations may change from $\Psi(\mathbf{R};\boldsymbol{\alpha}_0) \neq 0$ to $\Psi(\mathbf{R};\boldsymbol{\alpha}) = 0$ (or close to zero). This is usually called "moving a node through a configuration", which causes the local energy to diverge, and the unreweighted variance of the local energy to diverge as well. This results in artificial barriers in the energy landscape, which prevent sampling parameter space correctly, cause numerical difficulties and make the optimization procedure unlikely to succeed.

*2 points for describing the problem of moving a node through a configuration, or mentioning the expression "move a node through a configuration"*

(iii)
- Mean absolute deviation from the median energy minimization
- Reweighted variance minimization
- [Linear least-squares] energy minimization

*2 points for any valid answer*

*2 points for any other valid answer*

**5.**

(i)

(i.1) Ignoring serial correlation,

$$\sigma_{\bar{E}}^2 = \frac{\sigma_{E_L}^2}{M} \ , \tag{16}$$

where $M$ is the number of moves in the calculation.

With serial correlation such that the average correlation time of the sample is $\tau$ steps,

$$\sigma_{\bar{E}}^2 = \tau \frac{\sigma_{E_L}^2}{M} \ . \tag{17}$$

*3 points for either of the above equations*

(i.2) The variance of the local energy $\sigma_{E_L}^2$ can be reduced by improving the trial wave function.

*2 points*

(i.3) The standard error $\sigma_{\bar{E}}$ is proportional to $1/\sqrt{M}$. Therefore to obtain half the standard error obtained with $M_0$ steps we would need to run a total of $4M_0$ steps. Therefore the continuation run needs to be $3M_0$ steps long.

*2 points for reaching the conclusion that one needs $3M_0$ more steps, or that the total run should be $4M_0$ steps long*

(ii) Serial correlation is the similarity of consecutive local energies (or other expectation values) due to the spatial proximity of consecutive configurations, which cause the usual expression for the standard error to underestimate the actual standard error by a factor of $\sqrt{\tau}$.

The reblocking algorithm removes serial correlation by replacing groups of $n$ contiguous data by their average, and analyzing the behaviour of the standard error of the average of the blocked series of data as the block size $n$ changes. When the computed standard error reaches a plateau, the height of that plateau is the correlation-corrected error bar.

*1 point for mentioning grouping in blocks*

*2 points for drawing / mentioning a plateau*