Optimisation of Jastrow Factors by Variance Minimisation

Neil D. Drummond
and Richard J. Needs

October 18, 2005

Slater–Jastrow Wave Functions

- *Slater–Jastrow* wave functions are used in QMC:

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

- Free parameters in the Jastrow factor are:
  1. polynomial expansion coefficients,
  2. plane-wave expansion coefficients, and
  3. cutoff lengths for isotropic terms.

- The plane-wave and polynomial expansion coefficients occur **linearly** in the Jastrow exponent.

Why Optimise the Jastrow Factor?

- The DMC energy is independent of the Jastrow factor (in the limit of zero time step and infinite population).

- Optimising the Jastrow factor reduces time-step and population-control biases.

- Better wave functions give smaller statistical error bars on estimated quantities.

- Extrapolated estimation (e.g. of charge density) requires a good wave function.

Reweighted Variance Minimisation

- Generate a set of configurations distributed according to Ψ_0^2 using VMC.

- Associate a weight of Ψ^2/Ψ_0^2 with each config. (Jastrow factor doesn't affect nodal surface, so weights don't diverge.)

- Reweighted mean *local energy* $\Psi^{-1}\hat{H}\Psi$ estimates $\langle\hat{H}\rangle_\Psi$ and **reweighted variance** σ_w^2 of $\Psi^{-1}\hat{H}\Psi$ estimates $\sigma^2 \equiv \langle\hat{H}^2\rangle_\Psi - \langle\hat{H}\rangle_\Psi^2$.

- Can therefore estimate the energy variance for any given parameter set using a fixed sampling of configuration space.

- Variance is positive, but is zero if Ψ is an eigenstate of $\hat{H}$. So, *minimise the variance to optimise the wave function*.

Unreweighted Variance Minimisation

- If Ψ is an eigenstate of $\hat{H}$, then $\Psi^{-1}\hat{H}\Psi$ is constant in configuration space.

- Can therefore optimise Ψ by minimising the **unreweighted variance** σ_u^2 of the local energies of *any* fixed set of configurations.

- In limit of perfect sampling, σ_w^2 is independent of Ψ_0. This is not true of σ_u^2.

- Can iterate unreweighted varmin to **self-consistency**, however.

Example: 1D Quartic Potential

- Suppose a particle of unit mass moves in a 1D quartic potential. The Hamiltonian is

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + x^4.$$

- Let the trial wave function be

$$\Psi(x) = \exp\left(-\alpha x^2\right),$$

- The energy expectation value is

$$E = \frac{8\alpha^3 + 3}{16\alpha^2}.$$

  Minimum is $E = 0.6814$, when $\alpha = 0.9086$.

- The variance of the energy is

$$\sigma^2 = \frac{4\alpha^6 - 6\alpha^3 + 3}{8\alpha^4}.$$

The minimum of σ^2 is at $\alpha = 0.9676$, for which $E = 0.6841$.

- Suppose configurations are distributed according to $\Psi_0^2 \equiv \exp(-2\beta x^2)$. Then

$$\sigma_u^2 = \frac{4\beta^2\alpha^4 + 3 - 6\beta\alpha^2}{8\beta^4}.$$

The minimum of σ_u^2 for a given β is at

$$\alpha = \frac{1}{2}\sqrt{\frac{3}{\beta}}.$$

- Repeated unreweighted varmins generate a sequence of values of α, starting with $\alpha_0 = \beta$, and satisfying $\alpha_n = (1/2)\sqrt{3/\alpha_{n-1}}$.

- The stationary point of this process, irrespective of the initial β, is $\alpha_\infty = 0.9086$. This is equal to the result of *energy minimisation*, not variance minimisation!

Studies of Model Systems

- Model systems exist for which energy minimisation, reweighted varmin and SC unreweighted varmin give different results.

- In all cases studied, SC unreweighted varmin leads to a lower energy than reweighted varmin.

- Results of SC unreweighted varmin are often exactly the same as the results of energy minimisation.

Linear Jastrow Parameters

- Suppose the Jastrow exponent is linear in P parameters $\{\alpha\}$:

$$J(\mathbf{R}) = \sum_{i=1}^{P} f_i(\mathbf{R})\alpha_i + J_0(\mathbf{R}).$$

- Then the local energy $\Psi^{-1}\hat{H}\Psi$ is a quadratic function of the parameters.

- Suppose N_C configurations are distributed according to Ψ_0^2.

- The unreweighted average local energy is also quadratic in $\{\alpha\}$. It has a global maximum at $\{\alpha\}$ corresponding to Ψ_0.

- The unreweighted variance is a **quartic** function of the parameters.

- Sum over configs not required, so σ_u^2 can be evaluated *extremely* rapidly.

- The coefficients of the quartic variance can be accumulated in VMC.

- Along any line in parameter space, σ_u^2 is a quartic polynomial of a single parameter.

- The global minimum of a quartic polynomial can be determined analytically.

- Hence σ_u^2 can be minimised rapidly, exactly and globally along lines in parameter space.

- After using BFGS to find a (local) minimum of σ_u^2, one can perform billions of analytic line minimisations along random directions to look for lower minima.

Scaling

- Standard varmin scales as

$$\mathcal{O}\left(N^2 P^2 N_C\right).$$

- The new varmin method scales as

$$\mathcal{O}\left(P^4\right).$$

- Evaluating the "basis functions" $f_i(\mathbf{R})$ in VMC scales as

$$\mathcal{O}\left(P N^2 N_C\right).$$

- Evaluating the quartic coefficients in VMC scales as

$$\mathcal{O}\left(P^4 N_C\right).$$

Nature of the Variance in the Space of Linear Jastrow Parameters

- Nonglobal minima of σ_u^2 are *only* found for a very poor sampling of config space.

- σ_w^2 is more likely to contain nonglobal minima than σ_u^2.

- For good samplings of config space, unreweighted and reweighted varmin give very similar results; for poor samplings, unreweighted varmin is more stable.

- For poor Jastrow factors and large samplings of config space, the minima of the energy, σ_w^2 and σ_u^2 clearly differ.

- For high-quality Jastrow factors and good samplings of config space, the minima of the energy, σ_w^2 and σ_u^2 are very close.

Sampling of Configuration Space

- Wave-function quality increases rapidly with no. of configs used in optimisation.

- For small molecules, 10,000 configurations are usually sufficient to optimise linear Jastrow parameters.

- Variance is a smooth function of linear parameters; it is not a smooth function of the cutoff lengths, however.

- Improving the sampling of config space makes the variance a smoother function of the cutoff lengths.

- Can try to improve the quality of the sampling of config space by removing configs whose local energies are far from the mean.

Wave-Function Flexibility

- Wave-function quality generally increases rapidly as the number of free parameters is increased, before saturating.

- The results obtained with a very large number of parameters are often poorer than those obtained with smaller numbers.

Timing Results

- The optimisation phase in the new method is almost instantaneous, irrespective of the system size.

- The time taken to evaluate the basis functions is generally negligible.

- The time taken to compute the quartic coefficients is significant for small systems, but not for large systems (it is independent of system size).

- Overall, new method is 3–10 times faster than old method for a range of systems.

Conclusions

Self-consistent unreweighted variance minimisation is better than reweighted variance minimisation because:

1. In the limit of infinite sampling, the resulting energy is generally lower;

2. For small samplings of configuration space, it is more stable;

3. The unreweighted variance is a quartic function of linear Jastrow parameters, enabling rapid evaluation;

4. For any reasonable sampling of configuration space, the unreweighted variance does not have nonglobal minima in the space of linear Jastrow parameters.