Correlated sampling without reweighting, computing properties with size-independent variances

Roland Assaraf

Laboratoire de Chimie Théorique, CNRS-UMR 7616, Université Pierre et Marie Curie Paris VI, Case 137, 4, place Jussieu 75252 PARIS Cedex 05, France

BIRS 2012
Some perspective on Quantum Monte Carlo (QMC)

Many problem in Quantum physics at zero temperature

The Schroedinger equation,

\[ H\Phi = (-\sum_{i=1}^{N} \Delta_i + V(r_1, r_2 \ldots r_N))\Phi = E\Phi \]  

- \( N \) number of particles.
- \( r_i \), 3 spatial coordinates of particle \( i \).
- \( E \) lowest eigenvalue, the groundstate energy.
- \( \Phi(r_1 \ldots r_N) \) the lowest eigen vector, the groundstate
- \( \Phi \) antisymmetric for electrons (fermions).
Stochastic technics in principle adapted for solving the Schrödinger equation:

Solving the many problem in Quantum Physics

\[ \leftarrow \rightarrow \]

Computing integrals in large dimensions.
Example: variational energy

Variational energy

\[ E_V \equiv \langle \Psi | \hat{H} | \Psi \rangle \]

Average on a probability distribution

\[
\langle \Psi | \hat{H} | \Psi \rangle = \int dR \Psi^2(R) \frac{H\Psi}{\Psi}(R)
\]

\[
= \left\langle \frac{H\Psi}{\Psi}(R) \right\rangle \Psi^2 = \left\langle e(R) \right\rangle \Psi^2
\]

\( R \): 3N coordinates of the \( N \) interacting particles

\[
E_v = \frac{1}{N} \sum_{k=1}^{N} e(R_k)
\]
Example: variational energy

Variational energy

$$E_V \equiv \langle \Psi | \hat{H} | \Psi \rangle$$

Average on a probability distribution

$$\langle \Psi | \hat{H} | \Psi \rangle = \int d\mathbf{R} \Psi^2(\mathbf{R}) \frac{H\Psi}{\Psi}(\mathbf{R})$$

$$= \left\langle \frac{H\Psi}{\Psi}(\mathbf{R}) \right\rangle \Psi^2 = \left\langle e(\mathbf{R}) \right\rangle \Psi^2$$

\(\mathbf{R}\): 3N coordinates of the \(N\) interacting particles

$$E_v = \frac{1}{N} \sum_{k=1}^{N} e(\mathbf{R}_k)$$
Example: variational energy

Variational energy

\[ E_V \equiv \langle \Psi | \hat{H} | \Psi \rangle \]

Average on a probability distribution

\[
\langle \Psi | \hat{H} | \Psi \rangle = \int dR \Psi^2(R) \frac{H \Psi}{\Psi}(R)
\]

\[
= \left\langle \frac{H \Psi}{\Psi}(R) \right\rangle \Psi^2 = \left\langle e(R) \right\rangle \Psi^2
\]

\( R \): 3N coordinates of the \( N \) interacting particles

\[ E_v = \frac{1}{N} \sum_{k=1}^{N} e(R_k) \]

Roland Assaraf

Correlated sampling without reweighting
More generally

$$E_{\text{QMC}} = \left\langle e(R) \right\rangle_{\Pi}$$

Depending on the QMC method, the nature of $R$ might change:

- $3N$ particle coordinates (VMC, DMC..).
- Trajectories in the space of $3N$ particle coordinates (PDMC, PIMC, reptation...)
### Accurate energies

<table>
<thead>
<tr>
<th>No analytical integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Flexibility (choice of $\psi$ in VMC).</td>
</tr>
<tr>
<td>- Weak limitation in system sizes.</td>
</tr>
<tr>
<td>- Possibility to improve “arbitrarily” the accuracy (“zero-variance zero-bias principle”, choice of $\psi$ in VMC..).</td>
</tr>
</tbody>
</table>

In practice, reference methods for total energies on large systems (large $N$)
They are energy differences.

**Exploiting accurate total energies**

More tricky in QMC than in a deterministic method.

- Energy differences are usually very small.
- Statistical uncertainties.

Small statistical uncertainty on a total energy might be huge on a difference if energies are computed independently.
Why energy differences are usually small?

Examples

- Binding energies, transition state energies. One, two particle gaps (electron affinities, ionization energies) …

- First order derivatives of the energy: Any observable (force, dipole, moment, densities...).

- Higher order derivatives: spectroscopic constants …

They are groundstate energies of similar systems
Paradigm: Calculation of an observable $O$

$H_\lambda = H + \lambda O \Rightarrow \bar{O} = \frac{dE_\lambda}{d\lambda} \simeq \frac{E_\lambda - E_0}{\lambda} = \frac{\Delta_\lambda}{\lambda}$ \hspace{1cm} (2)

$\Delta_\lambda = E_\lambda - E_0 \propto \lambda \text{ small}$

Behavior as a function of the system size

$$\lim_{N \to \infty} \Delta_\lambda(N) = K \text{ finite.}$$

The perturbation $\lambda O$ depends usually on a few degrees of freedom.

$\Delta_\lambda$ has a locality property
In summary

**Small $\lambda$ and large $N$**

$$\Delta_\lambda(N) \propto \lambda$$

**Accuracy on $\Delta_\lambda$ in an independent energy calculation**

$$\frac{\delta \Delta_\lambda}{\Delta_\lambda} \propto \frac{\delta E_0}{\lambda} \propto \frac{\sqrt{N}}{\lambda}$$

No locality property for the statistical uncertainty.

**Comparison to total energy**

$$\frac{\delta \Delta_\lambda}{\Delta_\lambda} \propto \frac{N^{\frac{3}{2}} \delta E}{\lambda E}$$
Overview

1. Introduction

2. Correlated sampling with reweighting
   - The method
   - Statistical uncertainties
   - Numerical illustration

3. Correlated sampling with no reweighting
   - The method
   - Numerical illustration

4. Conclusion and perspectives
correlated sampling with reweighting

We have to compute the difference

\[ E_\lambda - E_0 = \langle e_\lambda(R) \rangle_{\pi_\lambda} - \langle e(R) \rangle_{\pi} \]

Sampling the same distribution for the two energies

\[ E_\lambda - E_0 = \frac{\langle e_\lambda \frac{\pi_\lambda}{\pi} \rangle_{\pi}}{\langle \frac{\pi_\lambda}{\pi} \rangle_{\pi}} - \langle e \rangle_{\pi}. \]  

(3)

weight \( w_\lambda \)

Different contexts

- Variational Monte Carlo \( e_\lambda(R) = \frac{H_\lambda \psi_\lambda}{\psi_\lambda}(R), \ w_\lambda(R) = \frac{\psi_\lambda^2}{\psi_2^2}(R) \)
- Forward walking method inc on text of DMC algorithms.
- ...
**Introduction**

**Correlated sampling with reweighting**

1. **Correlated sampling with no reweighting**

**Conclusion and perspectives**

---

**The method**

**Statistical uncertainties**

**Numerical illustration**

---

**Correlated sampling with reweighting**

We have to compute the difference

\[ E_\lambda - E_0 = \langle e_\lambda (R) \rangle_{\pi_\lambda} - \langle e (R) \rangle_{\pi} \]

Sampling the same distribution for the two energies

\[ E_\lambda - E_0 = \frac{\langle e_\lambda \frac{\pi_\lambda}{\pi} \rangle_{\pi}}{\langle \frac{\pi_\lambda}{\pi} \rangle_{\pi}} - \langle e \rangle_{\pi}. \tag{3} \]

---

**Different contexts**

- Variational Monte Carlo: \( e_\lambda (R) = \frac{H_\lambda \psi_\lambda}{\psi_\lambda} (R), \ w_\lambda (R) = \frac{\psi_\lambda^2}{\psi^2} (R) \)
- Forward walking method in context of DMC algorithms.
- ...

---

**Roland Assaraf**

**Correlated sampling without reweighting**
Introduction

Correlated sampling with reweighting
Correlated sampling with no reweighting
Conclusion and perspectives

The method
Statistical uncertainties
Numerical illustration

We have to compute the difference

\[ E_\lambda - E_0 = \langle e_\lambda(R) \rangle_{\pi_\lambda} - \langle e(R) \rangle_{\pi} \]

Sampling the same distribution for the two energies

\[ E_\lambda - E_0 = \frac{\langle e_\lambda \pi_{\lambda \pi} \rangle_{\pi}}{\langle \pi_{\lambda \pi} \rangle_{\pi}} - \langle e \rangle_{\pi}. \] (3)

Different contexts

- Variational Monte Carlo: \[ e_\lambda(R) = \frac{H_\lambda \psi_\lambda}{\psi_\lambda}(R), \quad w_\lambda(R) = \frac{\psi_\lambda^2}{\psi^2}(R) \]
- Forward walking method in context of DMC algorithms.
- ...

Roland Assaraf
Correlated sampling without reweighting
General expression

Compact expression

\[ \Delta_\lambda = E_\lambda - E_0 = \left\langle e_\lambda - e \right\rangle_\pi + \frac{\text{cov}(e_\lambda, w_\lambda)}{\left\langle w_\lambda \right\rangle_\pi} \] (4)

\[ E_\lambda - E_0 = \lambda \frac{\partial E_\lambda}{\partial \lambda} \bigg|_{\lambda=0} + o(\lambda). \]

\[ E'_\lambda = \left\langle e'_\lambda \right\rangle_\pi + \text{cov}(e_\lambda, w'_\lambda) \]

Zero-Variance (ZV) estimator

Pulay correction

Finite statistical uncertainty on \( E'_\lambda \) \( \implies \frac{\delta \Delta_\lambda}{\Delta_\lambda} = K + o(\lambda) \)
Pair correlation function


\[ O_u = \sum_{i<j} \delta(r_{ij} - u) \]

Probability density to find a pair of electrons at distance \( u \)

ZV term : \( \frac{dE}{d\lambda} = \langle O_u + \frac{(H-e)\psi'}{\psi_0} \rangle \psi_0^2 = \langle O_u \rangle \psi_0^2 \)
N-dependence

R. Assaraf, D. Domin, W. Lester.

Model of two separated (non interacting) subsystems

Particles coordinates $\mathbf{R}^l$ and $\mathbf{R}^u$. $H_\lambda = H_\lambda^l + H_\lambda^u$

Variational Monte Carlo

- $\mathbf{R} = (\mathbf{R}^l, \mathbf{R}^u)$
- $\Psi_\lambda(\mathbf{R}) = \Psi_\lambda(\mathbf{R}^l, \mathbf{R}^u) = \Psi_\lambda^l(\mathbf{R}^l)\Psi_\lambda^u(\mathbf{R}^u)$
- Local energy $e_\lambda(\mathbf{R}) = e_\lambda^l(\mathbf{R}^l) + e_\lambda^u(\mathbf{R}^u)$

\[
E_\lambda - E = \langle e_\lambda - e_0^l \rangle + \frac{\text{COV}(e_\lambda, w^l)}{\langle w^l \rangle} \tag{5}
\]
First term (ZV)

\[ \langle e^I_\lambda - e^I_0 \rangle \text{ depends only on } R^I \]

\[ \rightarrow \text{ Locality property of its variance} \]
The Pulay term

\[
\frac{\text{COV}(e^\lambda, w^l)}{\langle w^l \rangle} = \frac{\text{COV}(e^l, w^l)}{\langle w^l \rangle} + \frac{\text{COV}(e^u, w^l)}{\langle w^l \rangle} \tag{6}
\]

- **Local**
- **Non local**

- The non local contribution is 0 \((e^u \text{ and } w^l \text{ independent})\)!
- Its variance on a finite sample \((e^u(R^u_i), w^l(R^l_i))_{i \in [1..M]} : \)
  \[
  \propto V(e^u) \propto N
  \]

\[
\implies \delta \Delta^\lambda(N) \propto \sqrt{N} \text{ for large } N.
\]

Non locality property of the Pulay term.
Conclusion

\[ \frac{\delta \Delta \lambda}{\Delta \lambda} \propto \sqrt{N} \quad (7) \]

Correlated sampling with reweighting solves the small \( \lambda \) difficulty but not the large \( N \) one
Is the analysis for non interacting subsystems holds for interacting systems?

- Hydrogen chains, metallic and insulating
- Calculation of the force on the first nucleus: derivative of the energy with respect to the position of the first nucleus
- Variational calculation
- $\psi$ is a single determinant (Restricted Hartree Fock)
Metallic hydrogen chains

**Figure:** Energy derivative, different estimators

Roland Assaraf

Correlated sampling without reweighting
Insulating hydrogen chains

**Fig.**: Energy derivative, different estimators
Histogram of the ZV term, metallic chain

**Fig.**: Histogram of the energy derivative in the $H_n$ chain

ZV contribution has the local property!!
The local energy has not the local property
Statistical uncertainties

**Fig.:** Statistical uncertainties in the insulating $H_n$ chain
**Statistical uncertainties**

**Figure:** Statistical uncertainties in the metallic $H_n$ chain

Roland Assaraf

Correlated sampling without reweighting
The method
Assaraf, Caffarel, Kollias 2011

Basic idea

\[ \langle e_\lambda(R) \rangle_{\pi_\lambda} - \langle e(R) \rangle_\pi = \langle e_\lambda(R_\lambda) - e(R) \rangle_{\Pi(R, R_\lambda)} \]

- Marginal distributions of \( \Pi(R, R_\lambda) \) must be \( \pi(R), \pi_\lambda(R_\lambda) \).
- Differences of the order of \( \lambda \), \( \langle (R_\lambda - R)^2 \rangle = K\lambda^2 \)
How to build such a process

- Choosing close stochastic processes, $L, L_\lambda$ having $\pi$ and $\pi_\lambda$ as stationary states.
- Stability versus chaos. Two trajectories with the different initial conditions and same pseudo random numbers meet exponentially fast.
- Insures that close processes will produce close trajectories.
For example, with the overdamped Langevin process one would have

\[
\begin{align*}
R(t + dt) &= R(t) + \mathbf{b} \, [R(t)] \, dt + dW \\
R_\lambda(t + dt) &= R_\lambda(t) + \lambda \, \mathbf{b}_\lambda \, [R_\lambda(t)] \, dt + dW
\end{align*}
\]

(8)  
(9)
Stability of the process versus chaos

- Chain of 120 Hydrogens (120 electrons).
- Same process but different initial conditions.
- Perturbed system one atom displaced of $\lambda = 10^{-4} \text{a.u}$ (finite difference derivative).

![Graph showing synchronization of trajectories in H120 molecule]

$(R_{\lambda=0} - R)^2(t) \approx \langle (R_{\lambda=0} - R)^2(t) \rangle$ (average on 100 walkers at time $t$)
Independance of the uncertainties on $\lambda$

**Figure:** Quadratic distances between the two processes
Locality of the algorithm

**Fig.**: Square average of the inter electron distance at a given distance from the first atom
Fig.: Energy derivative with the correlated sampling with no reweighting
**Fig.:** Uncertainty as a function of $N$, metallic chains
Figure: Histogram of the correlated difference metallic chain
**Introduction**

Correlated sampling with reweighting

Correlated sampling with no reweighting

**Conclusion and perspectives**

**The method**

**Numerical illustration**

---

**Fig.:** Histogram of the correlated difference metallic chain
Reweighting introduces statistical fluctuations difficult to control

- Solves the small perturbation problem ($\lambda$ small).
- Sometimes large prefactors in the variance.
- Same large $N$ behavior as independent energy calculations.

Correlated sampling with no reweighting

- Solves the small $\lambda$ and large $N$ undesirable behavior.
- Perspective to obtain small energy differences with comparable accuracy to the energy.
- Relies on some particular dynamics (stability with respect to the chaos).
Possible to build such stable dynamics

- At the core of perfect sampling (criteria of time convergence, see Fahy, Krauth...).
- Building such dynamics for general molecules is underway.
- Vast subject (numerically, mathematically). Collaborations are welcome...