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

Marseille, GDR Correl 2015

Some perspective on Quantum Monte Carlo (QMC)

Many problem in Quantum physics at zero temperature

The Schroedinger equation,

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

- N number of particles.
- $\mathbf{r}_i$, 3 spatial coordinates of particle i.
- E lowest eigenvalue, the groundstate energy.
- $\Phi(\mathbf{r}_1 \ldots \mathbf{r}_N)$ the lowest eigen vector, the groundstate
- Φ antisymetric for electrons (fermions).

Stochastic technics in principle adapted for solving the Schroedinger equation :

Solving the many problem in Quantum Physics

<=>

Computing integrals in large dimensions.

Example : variational energy

Variational energy

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

Average on a probability distribution

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

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

$\mathbf{R}$: 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

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

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

$\mathbf{R}$: 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

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

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

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

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

In general

More generally

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

Depending on the QMC method, the nature of $\mathbf{R}$ and π might be different. e.g. for $\mathbf{R}$:

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

Precision on energies

$$E \propto N \quad \text{(homogenous) system of size } N \tag{2}$$

Statistical uncertainty

$$\delta E \propto \sqrt{\frac{V(e)}{M}} \propto \sqrt{\frac{N}{M}} \quad \text{sample size } M \tag{3}$$

$$\frac{\delta E}{E} \propto \frac{1}{\sqrt{MN}} \tag{4}$$

- Weak limitation in system size N (cpu and memory $\propto N ... N^3$)
- $V(e)$ can be lowered (choice of ψ) : zero-variance principle .

Quantities of physical interest

Energy differences

$$E_\lambda - E_0 = \langle e_\lambda \rangle_{\pi_\lambda} - \langle e_0 \rangle_{\pi_0} \tag{5}$$

"includes energy derivatives"

$$\lim_{\lambda \to 0} \frac{E_\lambda - E_0}{\lambda} \tag{6}$$

Quantities useful :

- Physical or chemical properties.
- Observables.
- Optimization of the energy.

Paradigm : Calculation of an observable O

$$H_\lambda = H + \lambda O => \bar{O} = \frac{dE_\lambda}{d\lambda} \simeq \frac{E_\lambda - E_0}{\lambda} = \frac{\Delta_\lambda}{\lambda} \tag{7}$$

Energy difference Δ_λ small

- λ small.
- λO often localized (forces...)

$$\Longrightarrow \Delta_\lambda \textbf{ has a locality property :}$$

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

$$\frac{\Delta_\lambda}{E_0} \propto \frac{\lambda}{N} \tag{8}$$

Paradigm : Calculation of an observable O

$$H_\lambda = H + \lambda O => \bar{O} = \frac{dE_\lambda}{d\lambda} \simeq \frac{E_\lambda - E_0}{\lambda} = \frac{\Delta_\lambda}{\lambda} \tag{7}$$

Energy difference Δ_λ small

- λ small.
- λO often localized (forces...)

$$\Longrightarrow \Delta_\lambda \textbf{ has a locality property :}$$

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

$$\frac{\Delta_\lambda}{E_0} \propto \frac{\lambda}{N} \tag{8}$$

First method : Independent calculation

Δ_λ 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 the total energy

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

Basic idea

We have to compute the difference

$$E_\lambda - E_0 = \langle e_\lambda(\mathbf{R}) \rangle_{\pi_\lambda} - \langle e_0(\mathbf{R}) \rangle_{\pi}$$

Sampling the same distribution for the two energies

$$E_\lambda - E_0 = \frac{\left\langle e_\lambda \frac{\pi_\lambda}{\pi_0} \right\rangle_{\pi}}{\left\langle \frac{\pi_\lambda}{\pi_0} \right\rangle_{\pi_0}} - \langle e_0 \rangle_{\pi_0}. \quad (9)$$

weight w_λ

Basic idea

We have to compute the difference

$$E_\lambda - E_0 = \langle e_\lambda(\mathbf{R})\rangle_{\pi_\lambda} - \langle e_0(\mathbf{R})\rangle_{\pi}$$

Sampling the same distribution for the two energies

$$E_\lambda - E_0 = \frac{\left\langle e_\lambda \frac{\pi_\lambda}{\pi_0}\right\rangle_{\pi}}{\left\langle \frac{\pi_\lambda}{\pi_0}\right\rangle_{\pi_0}} - \langle e_0\rangle_{\pi_0}. \qquad (9)$$

weight w_λ

Useful expression

$$\Delta_\lambda = E_\lambda - E_0 = \langle e_\lambda - e_0 \rangle_{\pi_0} + \frac{\mathrm{cov}(e_\lambda, w_\lambda)}{\langle w_\lambda \rangle_{\pi_0}} \tag{10}$$

λ dependence

$$E_\lambda - E_0 = \lambda \frac{\partial E_\lambda}{d\lambda}|_{\lambda=0} + o(\lambda).$$

$$E'_\lambda = \langle e'_\lambda [\psi'_\lambda] \rangle_{\pi_0} + \mathrm{cov}(e_\lambda, w'_\lambda)$$

Zero-Variance (ZV) estimator

Pulay correction

Finite statistical uncertainty on $E'_\lambda \Longrightarrow \frac{\delta \Delta_\lambda}{\Delta_\lambda} = K + o(\lambda)$

N-dependence

R. Assaraf, D. Domin PRE 2014

Model of two separated (non interacting) subsystems

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

Variational Monte Carlo

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

$$E_\lambda - E = \langle e^l_\lambda - e^l_0 \rangle + \frac{\mathrm{cov}(e_\lambda, w^l)}{\langle w^l \rangle} \qquad (11)$$

First term (ZV)

$\langle e_\lambda^l - e_0^l \rangle$ depends only on $\mathbf{R}^l$

$\Longrightarrow$ Locality property of its variance

Second term ("Pulay" term)

$$\frac{\mathrm{cov}(e_\lambda, w^l)}{\langle w^l \rangle} = \frac{\mathrm{cov}(e^l_\lambda, w^l)}{\langle w^l \rangle} + \frac{\mathrm{cov}(e^u, w^l)}{\langle w^l \rangle} \tag{12}$$

Local — Non local

- The non local contribution is 0 (e^u and w^l independant)!
- But the variance on a finite sample $(e^u(\mathbf{R}^u_i), w^l(\mathbf{R}^l_i))_{i\in[1..M]}$:

$$\propto V(e^u) \propto N$$

.

$\Longrightarrow \delta\Delta_\lambda(N) \propto \sqrt{N}$ for large N.

Non locality property for the statistical uncertainty for the Pulay term.

Conclusion

$$\frac{\delta\Delta_\lambda}{\Delta_\lambda} \propto \sqrt{N} \tag{13}$$

Correlated sampling with reweighting solves the small λ difficulty but not the large N one

Numerical illustration

Calculation of the force on a nucleus

Histograms, metallic chain

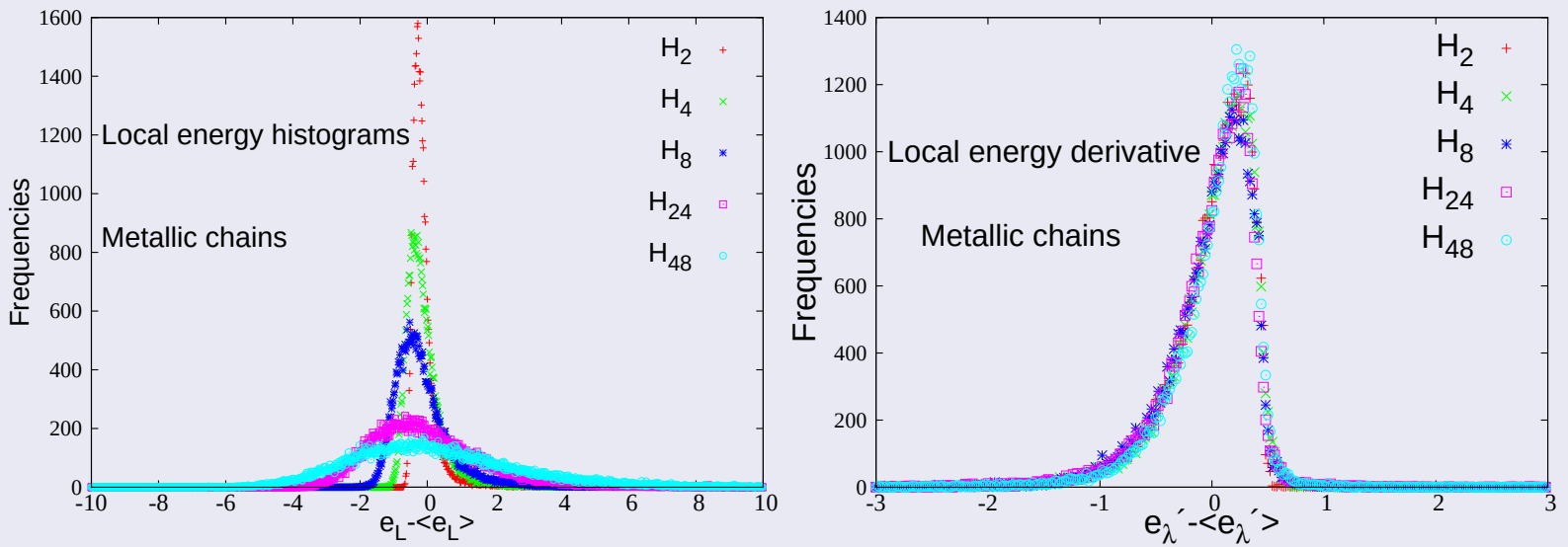

FIG.: Histogram of the ZV term and the local energy in the H_n chain

The ZV contribution has the locality property

Statistical uncertainties

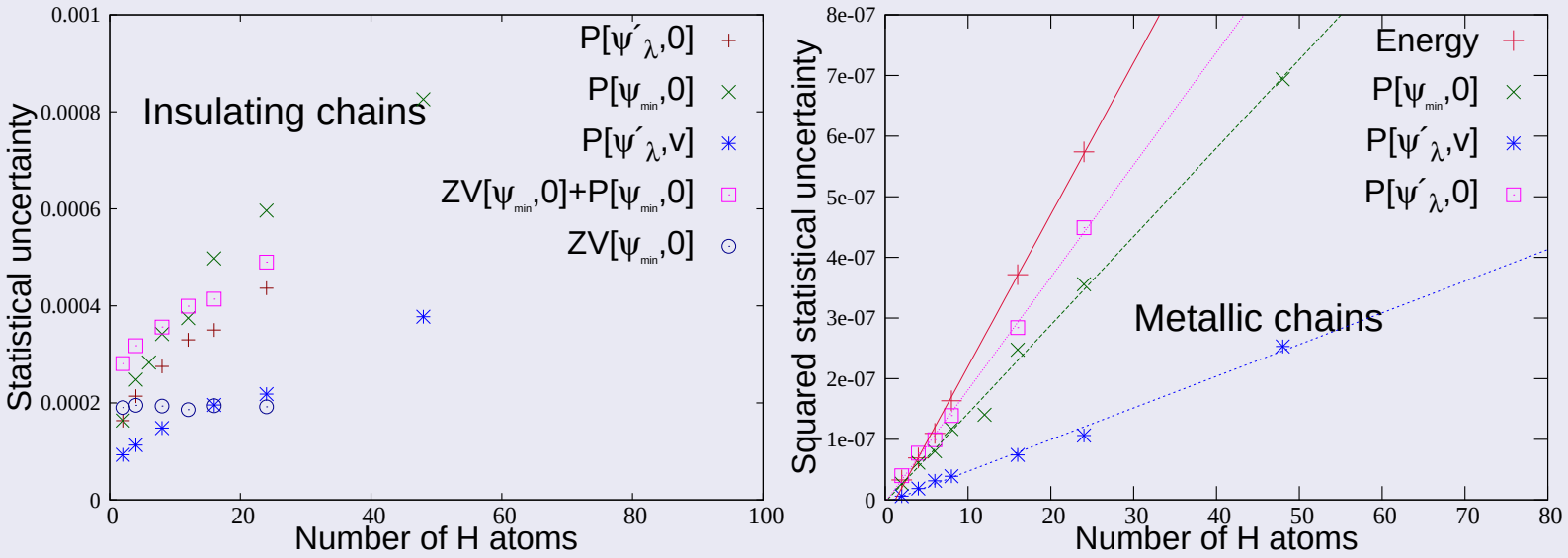

FIG.: Statistical uncertainties in the metallic H_n chain

The fluctuations of the weigths (Pulay correction) dominate for large N

The method

Assaraf, Caffarel, Kollias 2011

Basic idea

$$\langle e_\lambda(\mathbf{R})\rangle_{\pi_\lambda} - \langle e(\mathbf{R})\rangle_{\pi} = \langle e_\lambda(\mathbf{R}_\lambda) - e(\mathbf{R})\rangle_{\Pi(\mathbf{R},\mathbf{R}_\lambda)}$$

- Marginal distributions of $\Pi(\mathbf{R},\mathbf{R}_\lambda)$ must be $\pi(\mathbf{R}), \pi_\lambda(\mathbf{R}_\lambda)$.
- differences of the order of λ, $\langle(\mathbf{R}_\lambda) - \mathbf{R})^2\rangle = K\lambda^2$

How to build such a process

- Choosing close stochastic processes, L_0, L_λ having π_0 and π_λ 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

$$\mathbf{R}_0(t+dt) = \mathbf{R}_0(t) + \mathbf{b}_0\left[\mathbf{R}(t)\right]dt + \mathbf{dW} \quad (14)$$

$$\mathbf{R}_\lambda(t+dt) = \mathbf{R}_\lambda(t) + \mathbf{b}_\lambda\left[\mathbf{R}_\lambda(t)\right]dt + \mathbf{dW} \quad (15)$$

Stability of the process versus chaos

- Chain of 120 Hydrogens (120 electrons).
- Same process but different initial conditions.

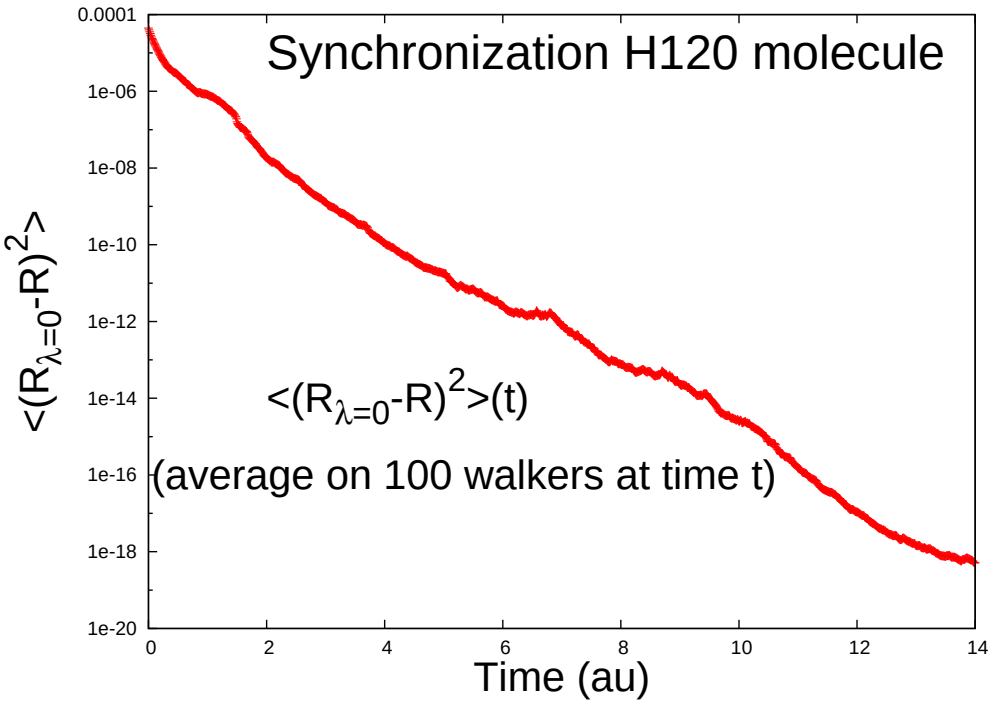

FIG.: Synchronization of trajectories in H_{120}

Back to the fully separated model

Same dynamic for $\mathbf{R}_\lambda$ and $\mathbf{R}_0$ in the region unaffected by the perturbation (region u) :

$$\mathbf{R}_\lambda = (\mathbf{R}_\lambda^l, \mathbf{R}^u)$$

The separability of the Hamiltonian and the trial wavefunction is transfered to the local energy :

$$e_\lambda(\mathbf{R}_\lambda) = e_\lambda^l(\mathbf{R}_\lambda^l) + e_0^u(\mathbf{R}_0^l)$$

$$\Longrightarrow e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0) = e_\lambda^l(\mathbf{R}_\lambda^l) - e_0^l(\mathbf{R}_0^l)$$

independent of $\mathbf{R}^u \Longrightarrow$ locality property for the variance

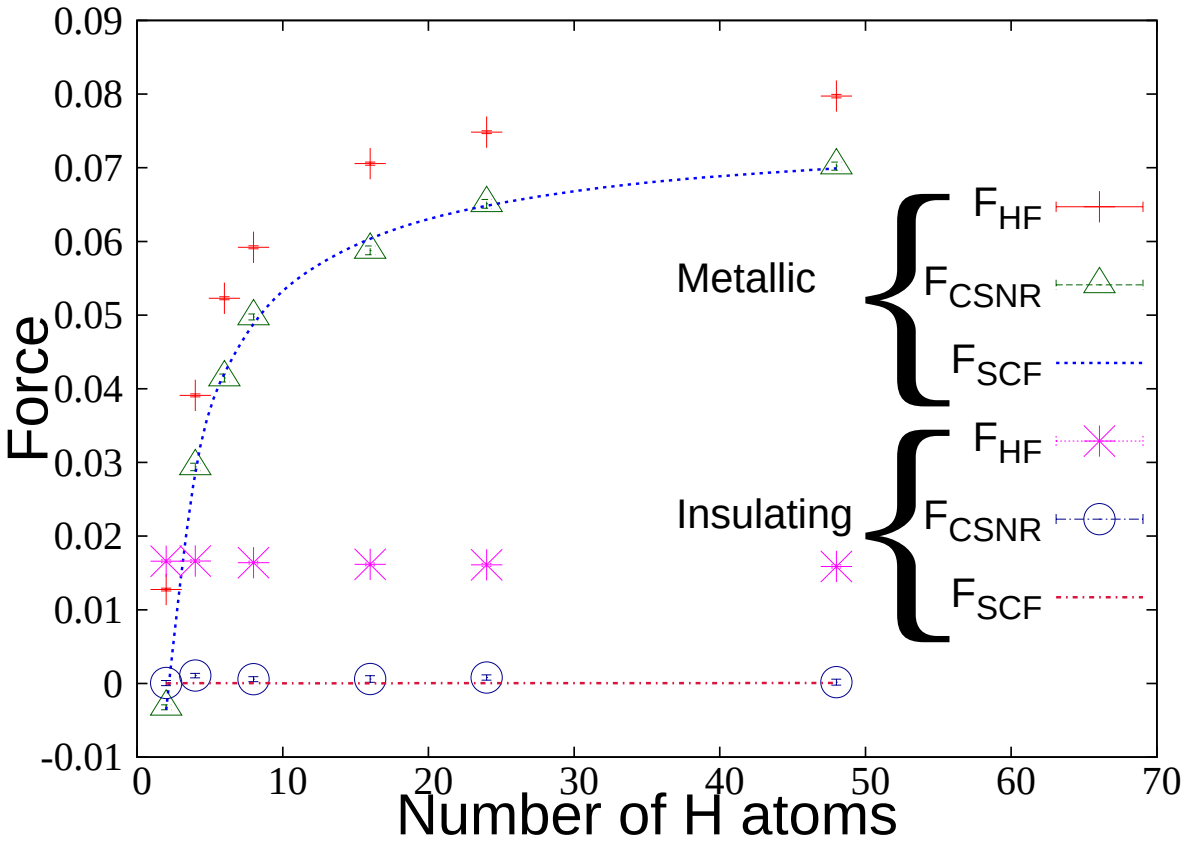

FIG.: Energy derivative with the correlated sampling with no reweighting

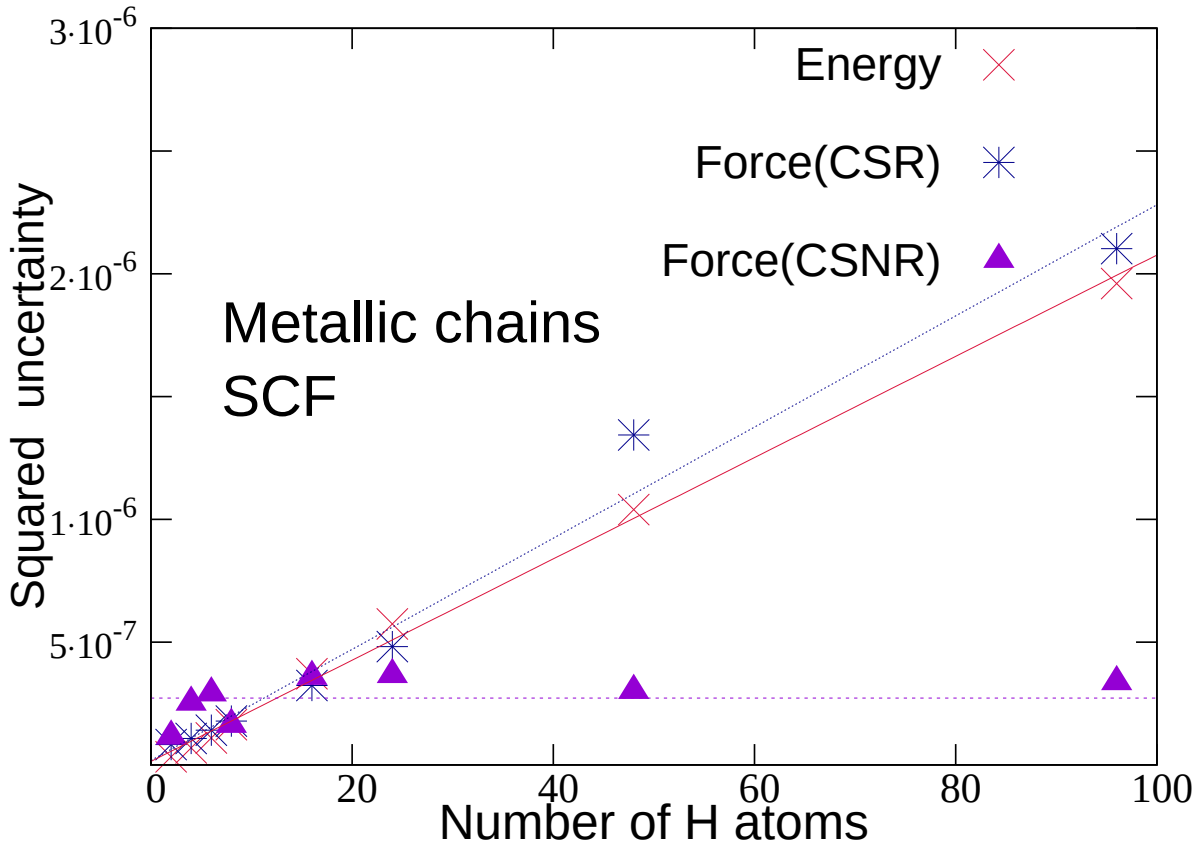

FIG.: Uncertainty as a function of N, metallic chains

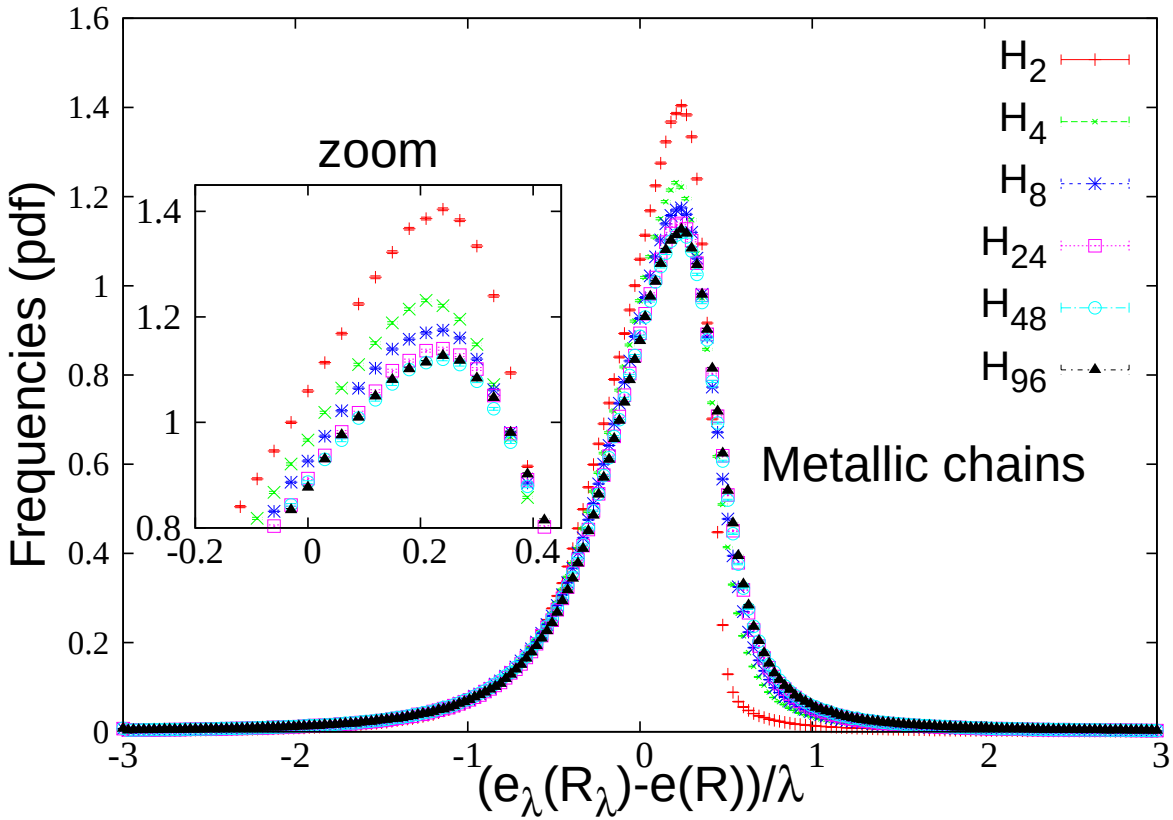

FIG.: Histogram of the correlated difference metallic chain

Locality of the algorithm

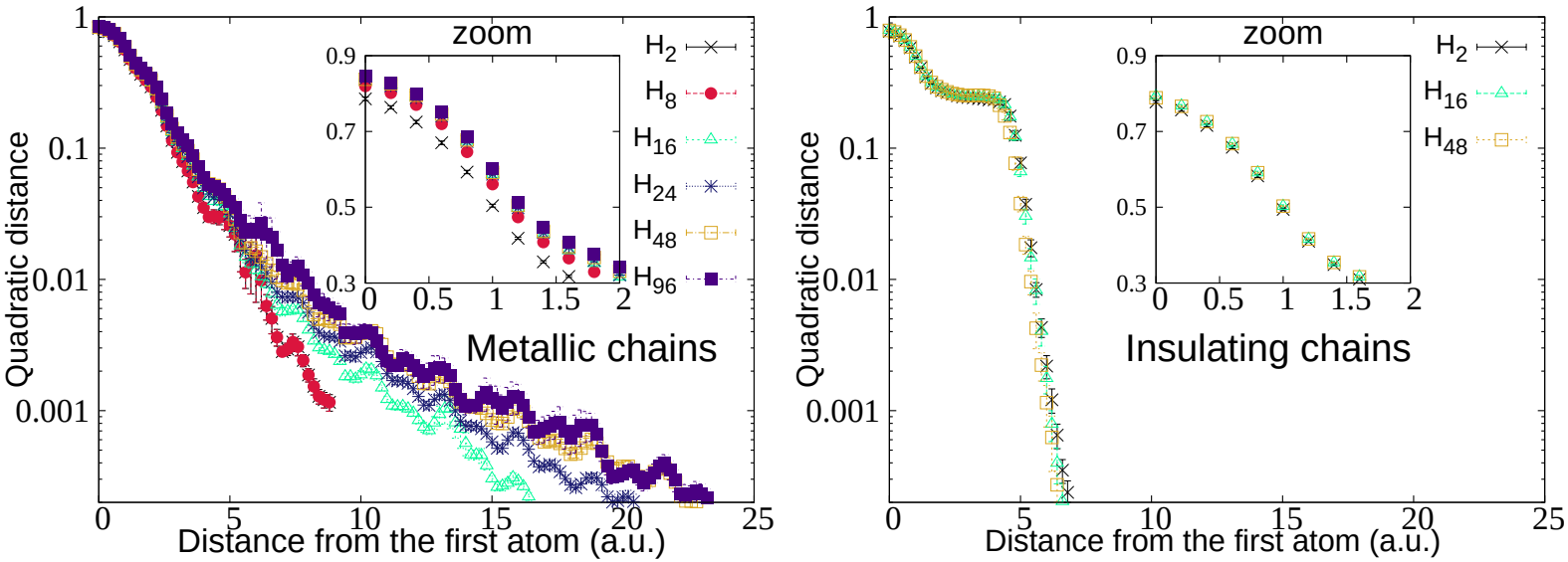

FIG.: Square average of the distance between two electrons belonging to $\mathbf{R}_0$ and $\mathbf{R}_1$ as a function of the distance to the first atom

Correlated sampling with no reweighting

- Seems to solve the small λ and the large N undesirable behaviors.
- Perspective to obtain small energy differences with comparable accuracy to the energy.

But relies on some particular dynamics (stability with respect to the chaos).

Illustration

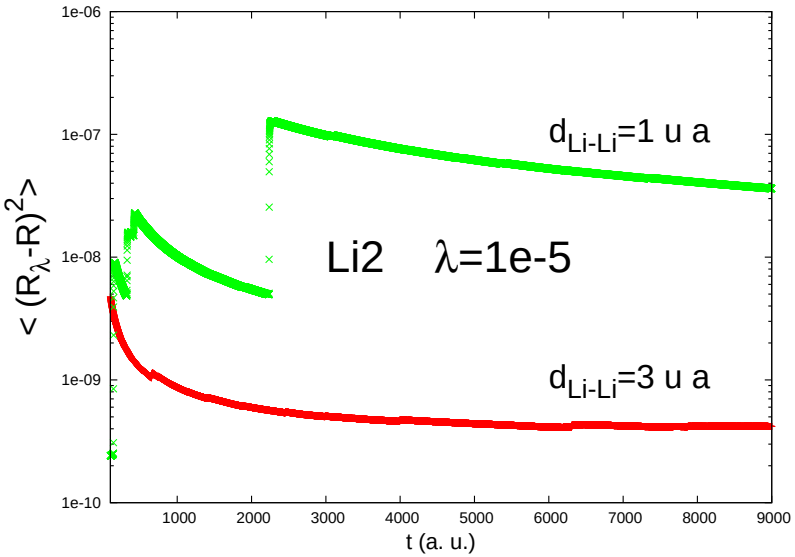

FIG.: Unstable situation versus a stable situation

$\Longrightarrow$ Necessity to better understand the chaotic properties of random processes and to develop stable dynamics (under way for larger systems).
Collaboration with T. Lelievre, B. Jourdain R. Roux.