# The generalized correlated sampling approach: toward an exact calculation of energy derivatives in Diffusion Monte Carlo

Roland Assaraf[a],Alexander Kollias[b] and Michel Caffarel[c]

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

[b)] Carnegie Institution of Washington, Geophysical Laboratory 5251 Broad Branch Rd., N.W., Washington, DC 20015, USA

[c)] Laboratoire de Chimie et Physique Quantiques, CNRS-UMR 5626, IRSAMC Université Paul Sabatier, 118 route de Narbonne 31062 Toulouse Cedex, France

# General perspective

**Quantum Monte Carlo (QMC) : stochastic technics used to solve the Schroedinger equation**

## In principle adapted to the many body problem (weak limitation in system sizes)

## In Practice

- Reference methods for groundstate energies of large systems.
- Less successful for other quantities.

# Quantities of physical interest ?

## Most of them can be obtained from total energies

- 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 ...

**Differences of energies of very close systems (small energy differences)**

# Paradigm : Calculation of an observable $O$

$$H_\lambda = H + \lambda O => \bar{O} = \frac{dE_\lambda}{d\lambda} \simeq \frac{E_\lambda - E_0}{\lambda} \quad (1)$$

## Direct calculation

$E_\lambda$, $E_0$ computed independently.

$$\delta\left(\frac{\mathbf{E_\lambda - E_0}}{\lambda}\right) \sim \frac{\delta \mathbf{E_0}}{\lambda} \underset{\lambda \to \mathbf{0}}{\longrightarrow} \infty$$

# System size dependency in a direct calculation

## Example : a one particle gap $\Delta = E(N+1) - E(N)$

$\delta E \sim N$ **and** $\Delta \sim 1$ **(best case)**

Independent calculation of energies :

$$\Longrightarrow \frac{\delta \Delta}{\Delta} \sim N$$

$1/N$ **plays the role of the small parameter** $\lambda$

In practice limiting factor on system sizes

Exploiting accurate QMC total energies to obtain accurate small differences is not as simple as in a deterministic method.

=> At the heart of practical limitations regarding properties one can compute and system sizes one can reach in QMC.

**Objective**

$$E_\lambda - E_0 \sim \lambda \Longrightarrow \delta(E_\lambda - E_0) \sim \lambda$$

$$\Longleftrightarrow$$

**Finite statistical error on energy derivatives.**

Exploiting accurate QMC total energies to obtain accurate small differences is not as simple as in a deterministic method.

=> At the heart of practical limitations regarding properties one can compute and system sizes one can reach in QMC.

## Objective

$$E_\lambda - E_0 \sim \lambda \Longrightarrow \delta(E_\lambda - E_0) \sim \lambda$$

$$\Longleftrightarrow$$

**Finite statistical error on energy derivatives.**

Exploiting accurate QMC total energies to obtain accurate small differences is not as simple as in a deterministic method.

=> At the heart of practical limitations regarding properties one can compute and system sizes one can reach in QMC.

## Objective

$$E_\lambda - E_0 \sim \lambda \Longrightarrow \delta(E_\lambda - E_0) \sim \lambda$$

$$\Longleftrightarrow$$

**Finite statistical error on energy derivatives.**

# Plan

# Plan

# Plan

# Plan

# Variational energy

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

## Average on a probability distribution

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

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

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

# Variational energy

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

## Average on a probability distribution

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

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

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

# Sampling of $\Psi^2$

## Dynamic over the configurations

$$\mathbf{R}(t+dt) = \mathbf{R}(t) + \mathbf{b}dt + d\mathbf{W} \quad (2)$$

- $\mathbf{b}(t) \equiv \frac{\nabla\Psi}{\Psi}$ (drift)
- $d\mathbf{W}$ gaussian random numbers (diffusion).
  $\langle dW_i dW_j \rangle = \sqrt{dt}\delta_{ij}$

A trajectory $\mathbf{R}(t)$ <=> Sample of $\Psi^2$.

# Why VMC can be a very accurate method ?

## Zero-variance - zero-bias property (ZVZB)

| | | |
|---|---|---|
| Bias | $E_V - E_0$ | $\lvert\Psi - \Phi\rvert^2$ |
| Variance | $\sigma^2(E_L)$ | $\lvert\Psi - \Phi\rvert^2$ |

$\phi$ Exact groundstate

Important since

Accuracy in QMC <=> $E_V - E_0$ systematic error

$+ \frac{\sigma(E_L)}{\sqrt{N}}$ statistical error.

# Diffusion Monte Carlo (DMC)

## Sampling the exact groundstate

$$e^{-tH}|\Psi\rangle = \Phi \tag{3}$$

Trotter formula : $e^{-tH} = e^{-\delta tH} e^{-\delta tH} \dots e^{-\delta tH}$

$$\langle \mathbf{R}'|\mathbf{e}^{-\delta \mathbf{t}\mathbf{H}}|\mathbf{R}\rangle = \mathbf{P}(\mathbf{R} \rightarrow \mathbf{R}')\mathbf{W}(\mathbf{R})$$

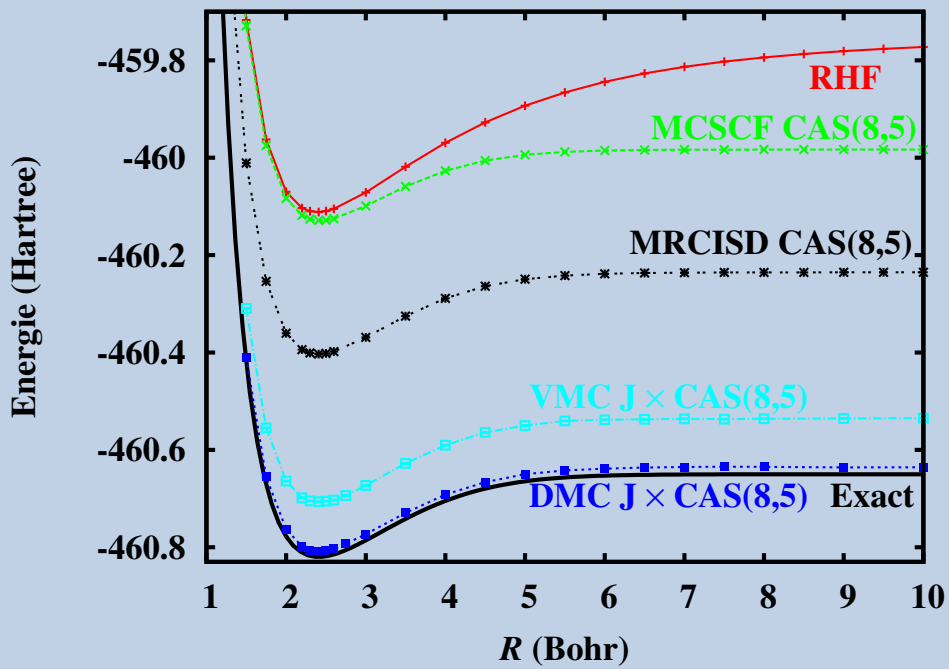
Overdamped Langevin Weight

In practice for fermions, Fixed node approximation :
$H \longrightarrow H_{FN} \Longrightarrow \Phi \longrightarrow \Phi_{FN}$ (variational solution in the space of functions having the same nodes as $\Psi$).

## Conclusion

We can sample something better than $\Psi$, $\Phi_{FN}$, even if its analytic form is unknown.

## Illustration




Dissociation curve for HCl (collaboration J. Toulouse)

**Can we compute accurate small differences or derivatives on the VMC curve, on the DMC curve ?**

# Observable and energy on a similar footing in VMC

## ZVZB improved estimators (R. Assaraf, M. Caffarel, 2003)

$$H_\lambda = H + \lambda O => \langle O \rangle_{\Phi^2} = \frac{dE_\lambda}{d\lambda} \simeq \frac{dE_V[\Psi_\lambda]}{d\lambda}$$

$$\Psi_\lambda = \Psi + \lambda\tilde{\Psi}$$

$$\frac{dE_V[\Psi_\lambda]}{d\lambda} = \langle \tilde{O} \rangle_{\Psi^2}$$

$$O \Longrightarrow \tilde{O}[\Psi, \tilde{\Psi}] = O + \frac{(H - E_L)\tilde{\Psi}}{\Psi} + 2(E_L - E_V)\frac{\tilde{\Psi}}{\Psi}$$

# Observable and energy on a similar footing in VMC

## ZVZB improved estimators (R. Assaraf, M. Caffarel, 2003)

$$H_\lambda = H + \lambda O => \langle O \rangle_{\Phi^2} = \frac{dE_\lambda}{d\lambda} \simeq \frac{dE_V[\Psi_\lambda]}{d\lambda}$$

$$\Psi_\lambda = \Psi + \lambda \tilde{\Psi}$$

$$\frac{dE_V[\Psi_\lambda]}{d\lambda} = \langle \tilde{O} \rangle_{\Psi^2}$$

$$O \Longrightarrow \tilde{O}[\Psi, \tilde{\Psi}] = O + \frac{(H - E_L)\tilde{\Psi}}{\Psi} + 2(E_L - E_V)\frac{\tilde{\Psi}}{\Psi}$$

# Energy/observable in VMC

| | Bias | Variance | Choice of |
|---|---|---|---|
| $E_V[\Psi]$ | $\delta\Psi^2$ | Idem | $\Psi$ |
| $E_V'_\lambda = \langle \tilde{O}[\Psi,\tilde{\Psi}]\rangle_{\Psi^2}$ | $\delta\Psi^2 + \delta\Psi\delta\Psi'$ | Idem | $(\Psi,\tilde{\Psi})$ |
| $\langle O\rangle_{\Psi^2}$ | $\delta\Psi$ | $O(1)$ | $(\Psi,\tilde{\Psi}=0)$ |

$$\delta\Psi = \Psi - \Phi$$

$$\delta\Psi' = \tilde{\Psi} - \frac{d\Phi_\lambda}{d\lambda}$$

# Electronic density

R. Assaraf, M. Caffarel, and A. Scemama, Phys. Rev. E. **75** 035701, (2007)

$$\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r}_i - \mathbf{r})$$

Regularized bare estimator :

Counting electrons in small boxes of size $\Delta R$ around $\mathbf{r}$ :

$=>$ Bias $O(\Delta R)$

$=>$ Variance $O(\frac{1}{\Delta R}^3)$.

**Poorest, in regions rarely or never visited by electrons**

## Simplest improved estimator

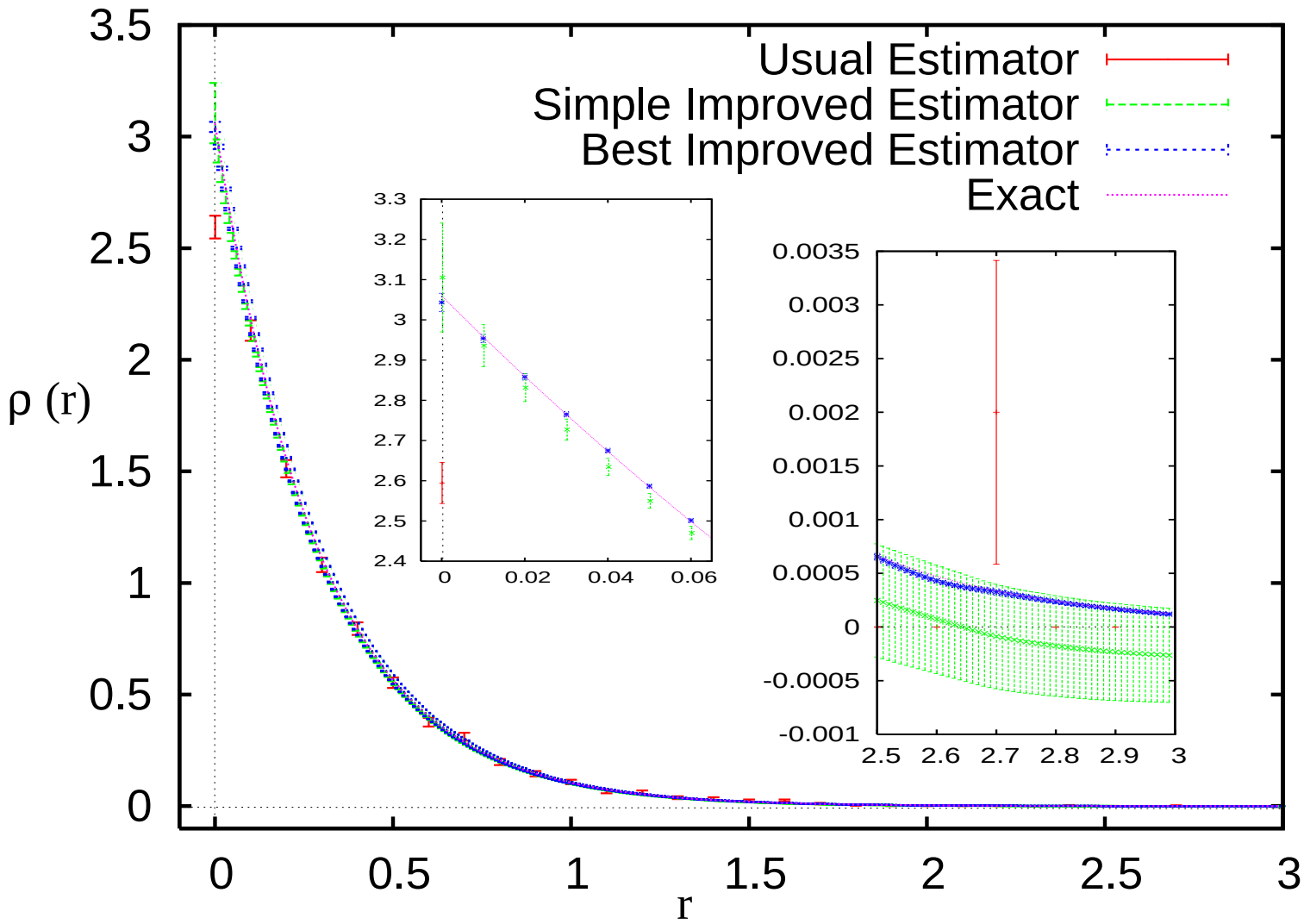
$\delta(\mathbf{r}_i - \mathbf{r}) = \frac{-1}{4\pi}\Delta\frac{1}{|\mathbf{r}_i - \mathbf{r}|}$ + Green formula

$$\tilde{\rho}_{ZV}(\mathbf{r}) = \left\langle -\frac{1}{4\pi}\sum_{i=1}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}|}\frac{\nabla_i^2(\Psi^2)}{\Psi^2}\right\rangle_{\Psi^2}.$$

Used by P. Langfelder [1] *et al* for the density at a nucleus, in combination with a modified important sampling (to cure the still infinite variance).

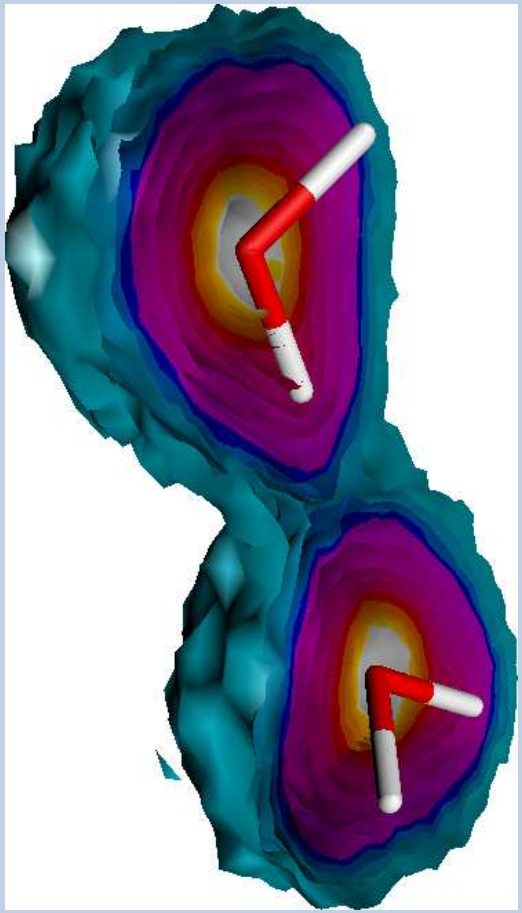
[1] P. Langfelder, S.M. Rothstein, J. Vrbik, J. Chem. Phys. **107** 8525 (1997)

# Illustration : Helium atom



# Application : water dimer

Bare regularized estimator



Improved estimator

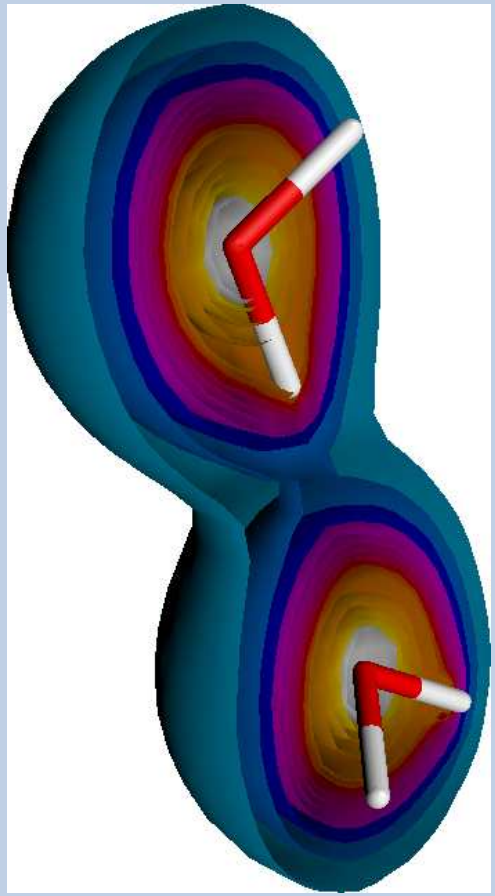


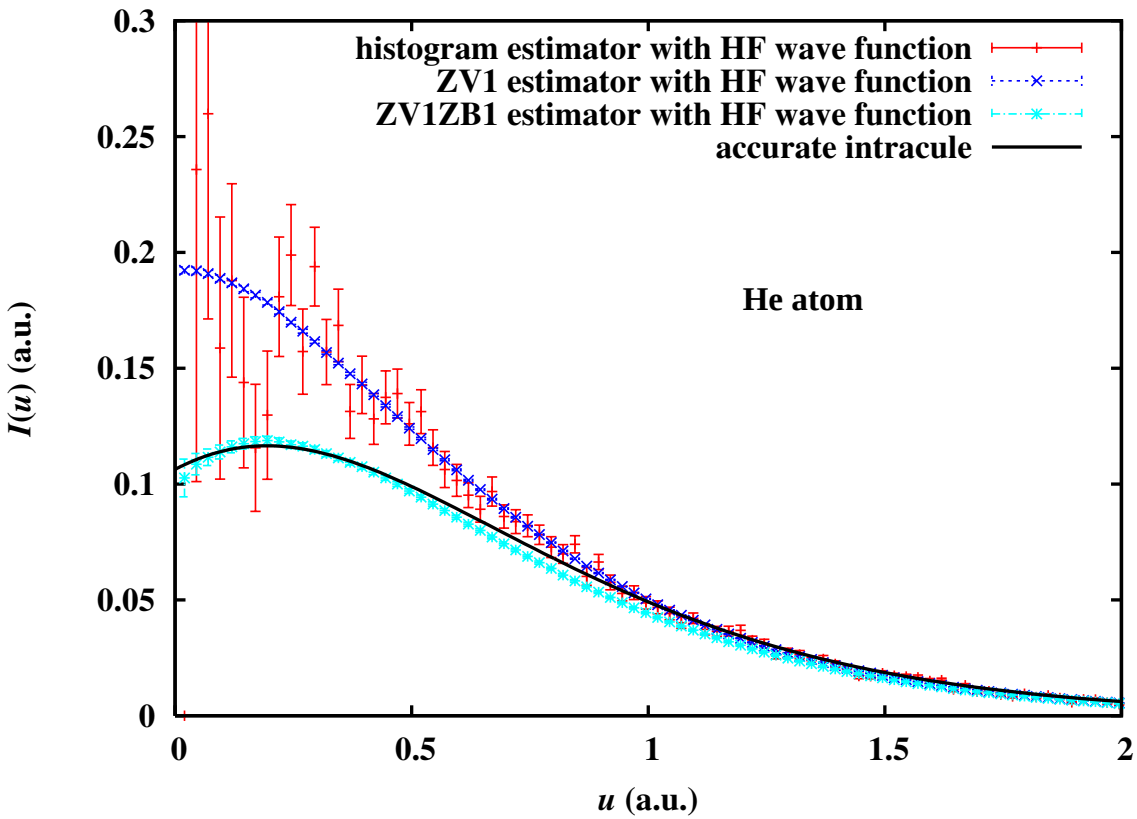
Fig.: Isodensity surfaces of the water dime

# Pair correlation function

[J. Toulouse, R. Assaraf, C. J. Umrigar. J. Chem. Phys. **126** 244112 (2007)]

$\hat{I}(u) = \sum_{i<j} \delta(r_{ij} - u)$

Probability density to find a pair of electrons at distance $u$



# Impact of the optimisation of $\Psi$

**Illustration : Intracule calculations (collaboration with J. Toulouse et C. Umrigar)**

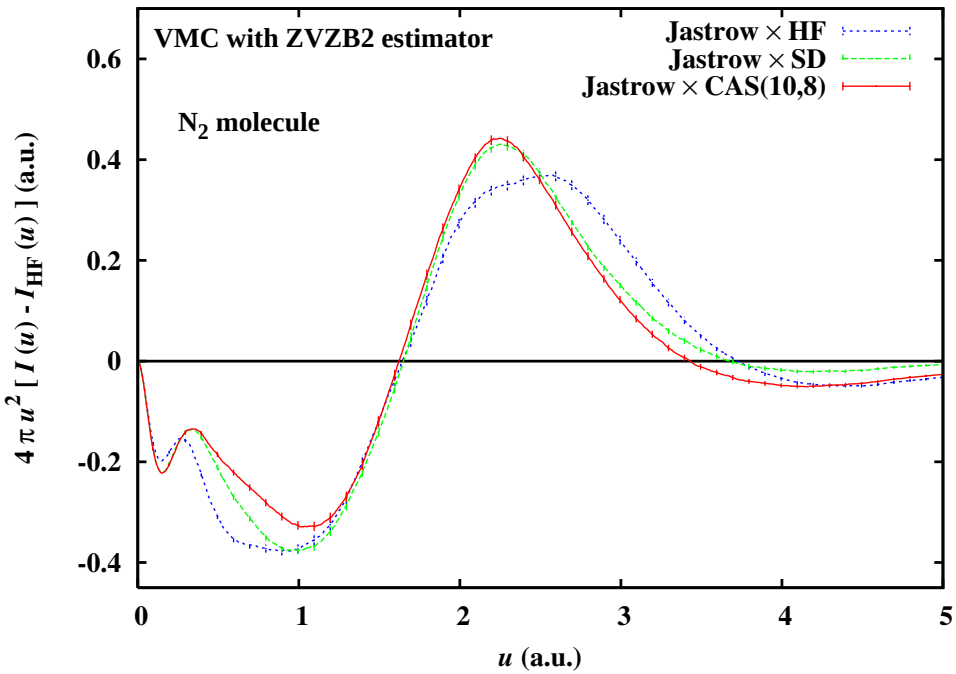



Fig.: Correlation part of the radial intracule in function of the distance electron-electron $u$ for the molecule $N_2$.

# In summary

**Derivatives (or small differences) of VMC energies : on the same footing as total VMC energies**

Strategy : Using improved estimators $\tilde{O}_{ZVZB}[\Psi, \tilde{\Psi}]$.

$=>$ control of the statistical error and the systematic error, at a reasonable computational cost.

**Current work :**

- Improving estimators for forces (geometry optimization).
- Application to a variety of other properties (dipole moments...).

# Generalization to DMC

## R. Assaraf, A. Kollias, M. Caffarel

### The DMC energy

$$E_{DMC} = \langle \Psi \mid H \mid \Phi \rangle = \langle E_L \rangle_{\Psi\Phi}$$

Trial function — Groundstate of $H$

$$\Psi\Phi = \Psi e^{-tH} \mid \Psi \rangle = \left\langle e^{-\int_0^t E_L(\mathbf{R}(s))ds} \mid \mathbf{R}(t) \right\rangle_{[\mathbf{R}(s)]} \quad (2)$$

$[\mathbf{R}(s)]$ Overdamped Langevin process with the drift $\mathbf{b} = \frac{\nabla\Psi}{\Psi}$

# Generalization to DMC

## R. Assaraf, A. Kollias, M. Caffarel

### The DMC energy

$$E_{DMC} = \langle \Psi \mid H \mid \Phi \rangle = \langle E_L \rangle_{\Psi\Phi}$$

Trial function — Groundstate of $H$

$$\Psi\Phi = \Psi e^{-tH} \mid \Psi \rangle = \left\langle e^{-\int_0^t E_L(\mathbf{R}(s))ds} \mid \mathbf{R}(t) \right\rangle_{[\mathbf{R}(s)]} \quad (2)$$

$[\mathbf{R}(s)]$ Overdamped Langevin process with the drift $\mathbf{b} = \frac{\nabla\Psi}{\Psi}$

# Generalization to DMC

## R. Assaraf, A. Kollias, M. Caffarel

### The DMC energy

$$E_{DMC} = \langle \Psi \mid H \mid \Phi \rangle = \langle E_L \rangle_{\Psi\Phi}$$

Trial function — Groundstate of $H$

$$\Psi\Phi = \Psi e^{-tH} \mid \Psi \rangle = \left\langle e^{-\int_0^t E_L(\mathbf{R}(s))ds} \mid \mathbf{R}(t) \right\rangle_{[\mathbf{R}(s)]} \quad (2)$$

$[\mathbf{R}(s)]$ Overdamped Langevin process with the drift $\mathbf{b} = \frac{\nabla\Psi}{\Psi}$

# Forward walking

$$H_\lambda = H + \lambda O => H_\lambda \Phi_\lambda = E_\lambda \Phi_\lambda$$

Computation of

$$\frac{dE_\lambda}{d\lambda} = \frac{d\langle E_{L\lambda}\rangle_{\Phi_\lambda \Psi}}{d\lambda}$$

## Forward walking

Same trial function for $H$ and $H_\lambda$ ($\Psi = \Psi_\lambda$).

$$\frac{d\Psi\Phi_\lambda}{d\lambda} = \left\langle \left[ -\int_0^{\mathbf{t}} \mathbf{dsO[R(s)]} \right] e^{-\int_0^t E_L(\mathbf{R}(s))ds} \mid \mathbf{R}(t) \right\rangle_{[\mathbf{R}(s)]}$$

In principle exact, but tractable only if $\mathbf{O}$ has a small variance.
Example for a force at the nucleus, variance infinite.

# Approximations

## Example : Hybrid improved estimator, in a FN approximation

$$\langle \tilde{O} \rangle_h \equiv 2\langle \tilde{O} \rangle_{DMC} - \langle \tilde{O} \rangle_{VMC}$$

- Systematic error : $O(|\Psi - \Phi|^2) + O(|\tilde{\Psi} - \Phi_1|.|\Phi_{FN} - \Phi|)$

  $O(|\Phi_{FN} - \Phi|)$ for $\langle O \rangle_h$

- Variance $O(|\Psi - \Phi|^2)$

  $O(1)$ for $\langle O \rangle_h$

Application to forces (R. Assaraf, M. Caffarel, 2003)

# Generalized correlated sampling

## First ingredient : two different guiding functions

$$\Psi_\lambda = \Psi + \lambda\tilde{\Psi} => \mathbf{b}_\lambda = \mathbf{b} + \lambda\tilde{\mathbf{b}}$$

$$E_\lambda - E_0 = \langle E_{L\lambda}\rangle_{\Psi_\lambda\Phi_\lambda} - \langle E_L\rangle_{\Psi\Phi}$$

## Second ingredient correlating the Overamped Langevin process

$$\mathbf{R}(t+dt) = \mathbf{R}(t) + \mathbf{b}dt + d\mathbf{W}$$
$$\mathbf{R}_\lambda(t+dt) = \mathbf{R}_\lambda(t) + \mathbf{b}_\lambda dt + d\mathbf{W}_\lambda$$

Wiener processes related by a unitary transform

$$d\mathbf{W}_\lambda = U_\lambda d\mathbf{W} \quad (3)$$

## Difference of energies : average on the process $[\mathbf{R}, \mathbf{R}_\lambda]$.

$$E_\lambda - E_0 = \left\langle E_{L\lambda}(\mathbf{R}_\lambda) e^{-\int_0^t ds E_{L\lambda}(\mathbf{R}_\lambda)} - E_L(\mathbf{R}) e^{-\int_0^t ds E_L(\mathbf{R})} \right\rangle_{[\mathbf{R},\mathbf{R}_\lambda]}$$

## Why a hope of improvement over usual forward walking ?

$$\int_0^t ds O[\mathbf{R}(s)] \longrightarrow \int_0^t ds \tilde{O}[\mathbf{R}(s)]$$

$$\tilde{O} = \frac{dE_{L\lambda}(\mathbf{R}_\lambda)}{d\lambda} = \lambda \tilde{O}_{ZV}(\mathbf{R}) + \quad \vec{T} \quad .\vec{\nabla} E_L(\mathbf{R})$$

$$\vec{T} \equiv \frac{\mathbf{R}_\lambda - \mathbf{R}}{\lambda}$$

ZV principle on $\tilde{O}$ (choice of $\tilde{\psi}$) ! !
Small fluctuations **if** $\langle \vec{T}^2 \rangle$ **finite !**

# Why can we hope $\langle T^2 \rangle$ finite ?

## Analysis in 1D $U = Id$

$$\begin{aligned} \frac{dT}{dt} &= \frac{\partial b}{\partial x} T \quad \text{Growth rate term} \\ &+ \tilde{b} \quad \text{Deviation term} \end{aligned} \tag{4}$$

## If $\tilde{b}$ turned off at time $t_0$

The accumulated deviation $T(t_0)$ will evolve as

$$T(t) = e^{\int_{t_0}^{t} \frac{\partial b}{\partial x}} T(t_0) \to 0$$

Since $\langle \frac{\partial b}{\partial x} \rangle < 0$ (the system lies in a finite region).

# Illustration in many dimensions

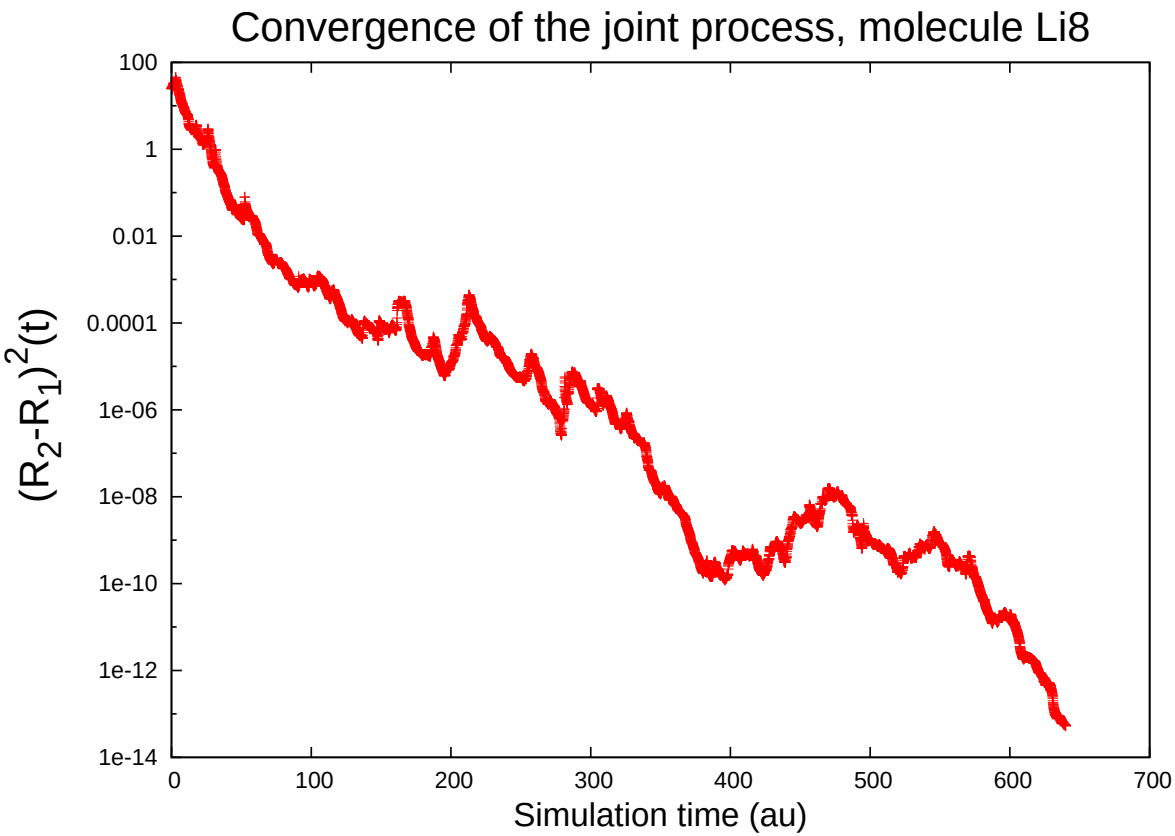



Fig.: Quadratic deviation of two correlated simulations on the $Li_8$ molecule

# Stability for the H2 molecule

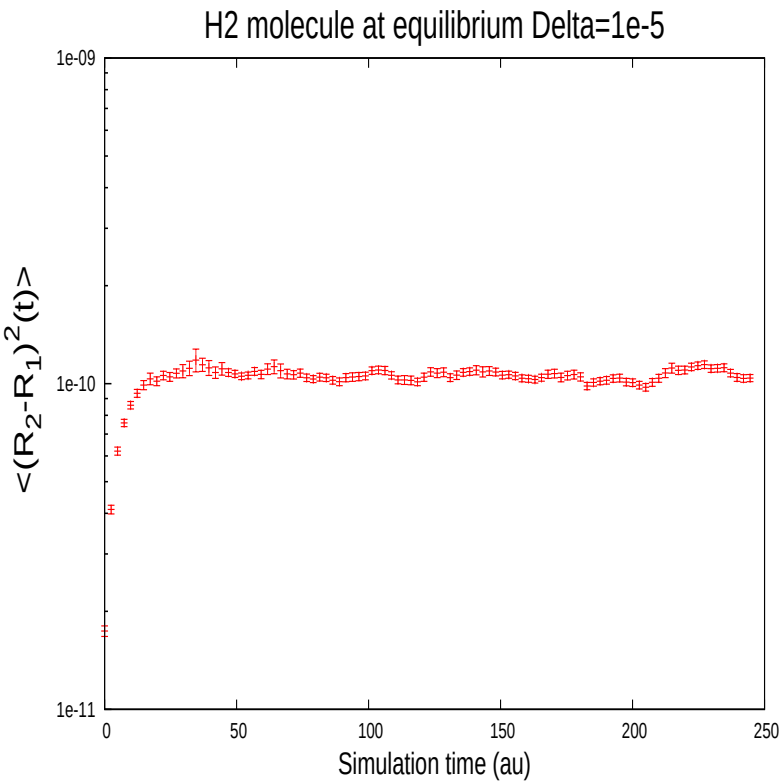
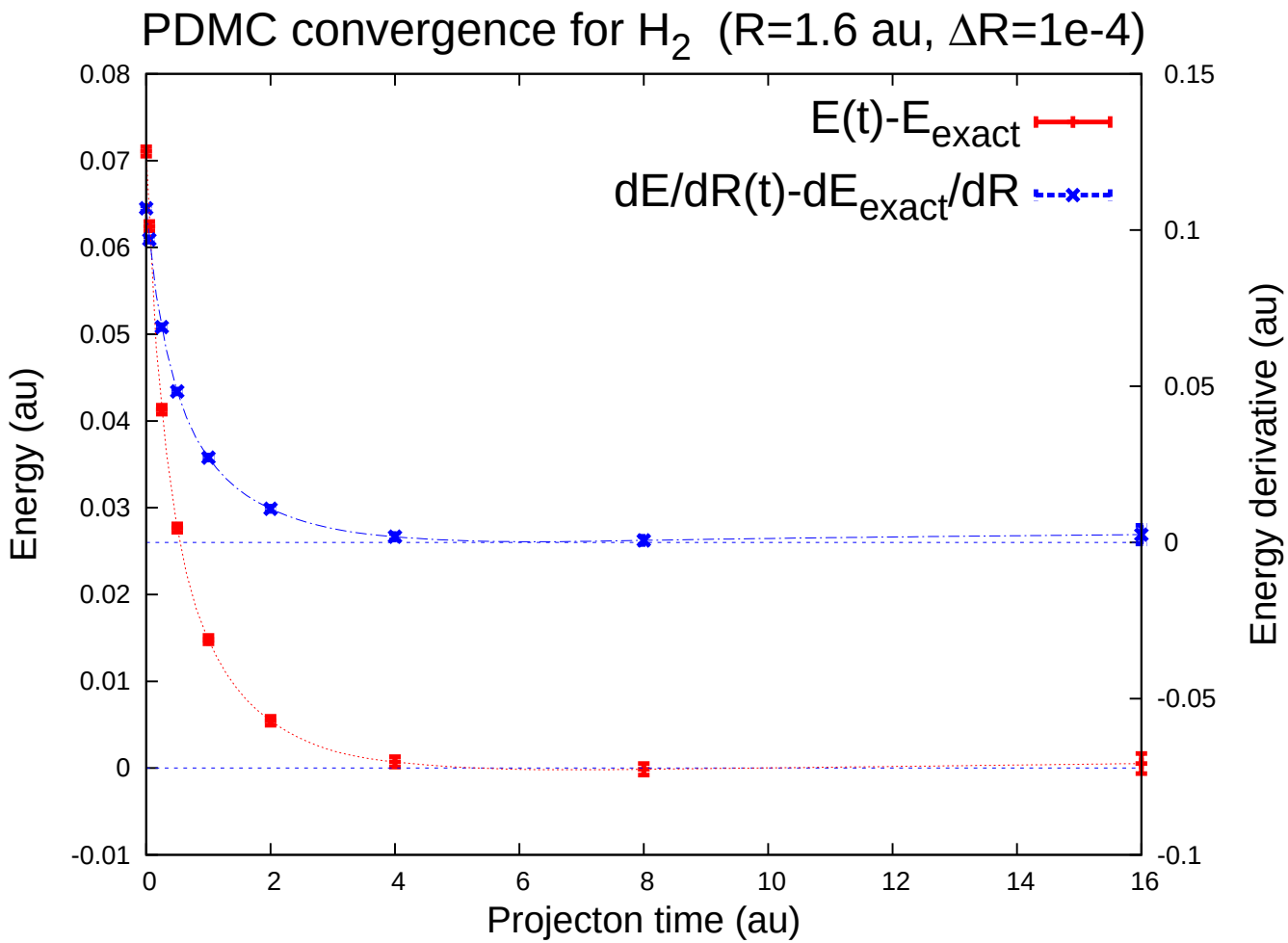



Fig.: Quadratic distance between the two different processes

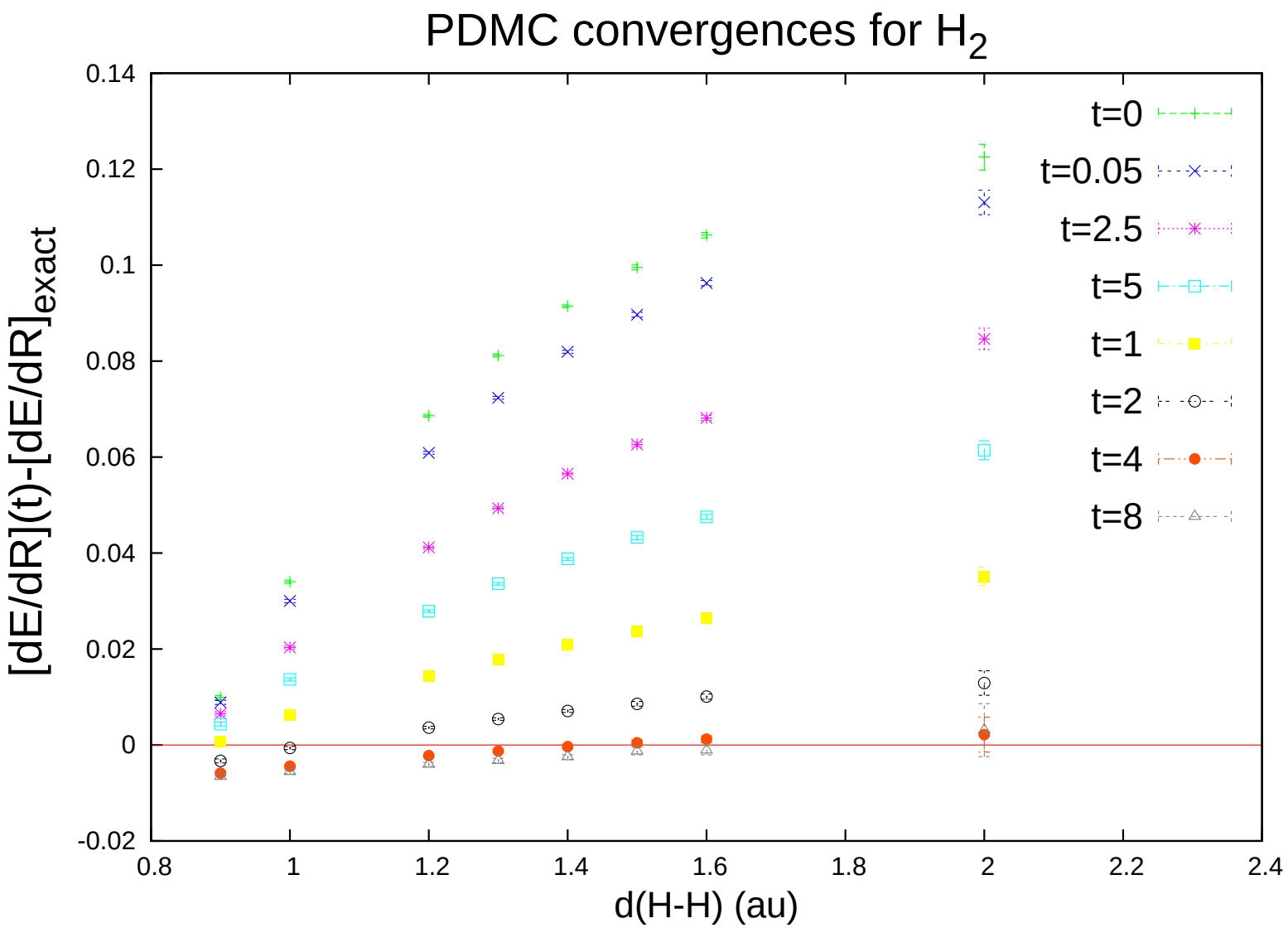
# DMC energy derivative in H2

## Crude $\Psi$ : Unoptimized CSF $\times$ Minimal jastrow (e-e cusp)

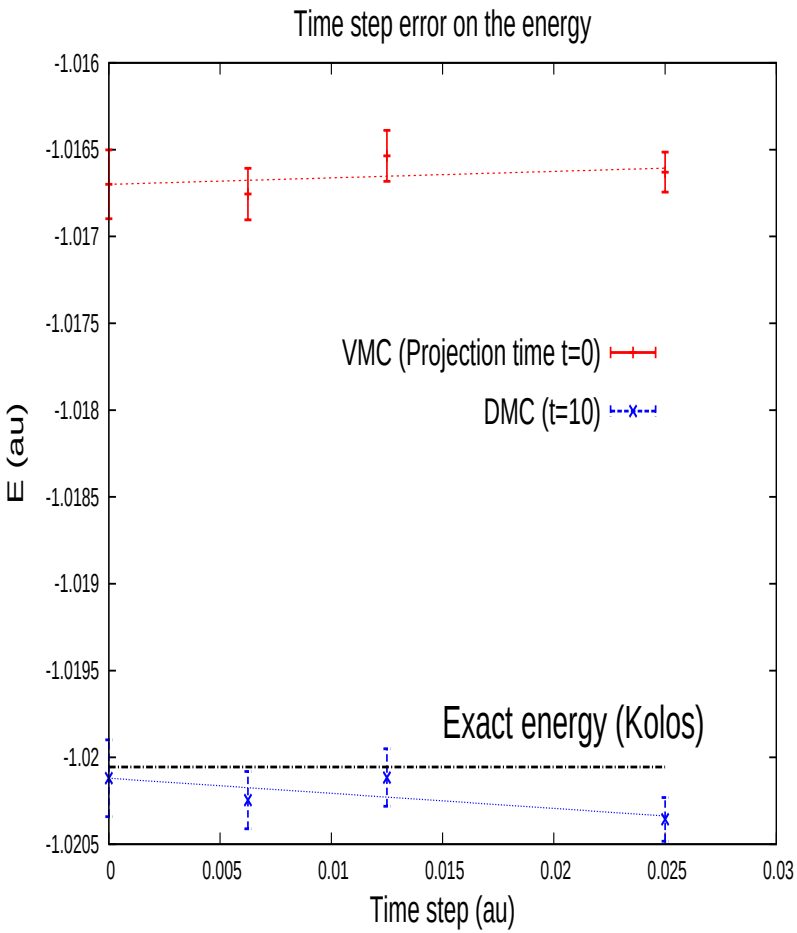


Plan





# Finite time step error



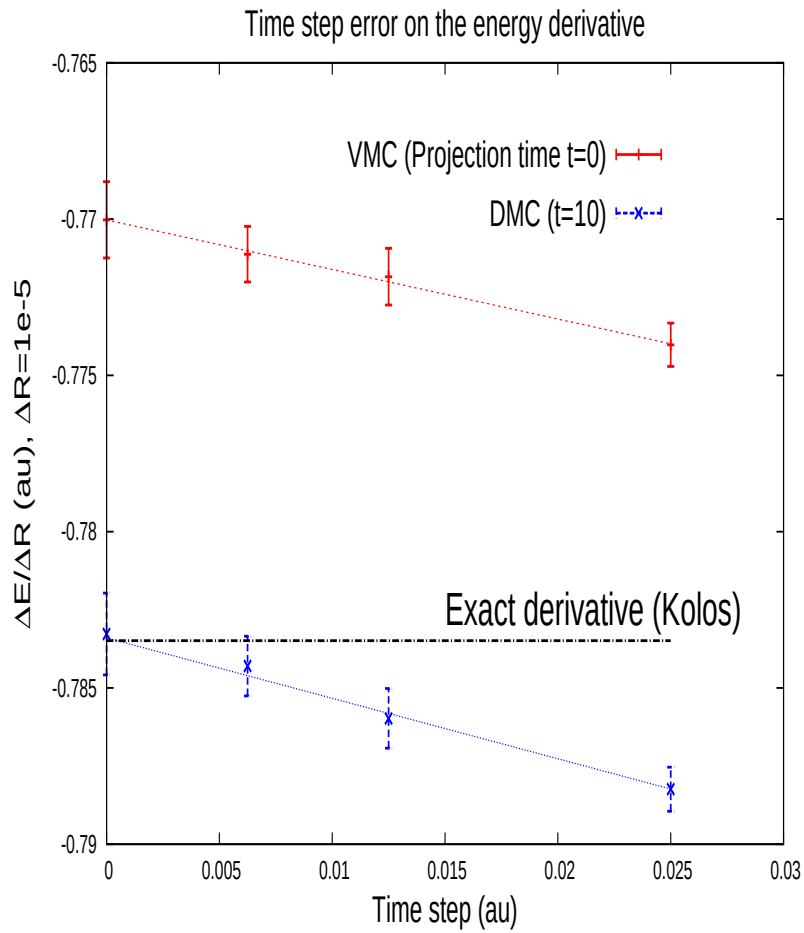
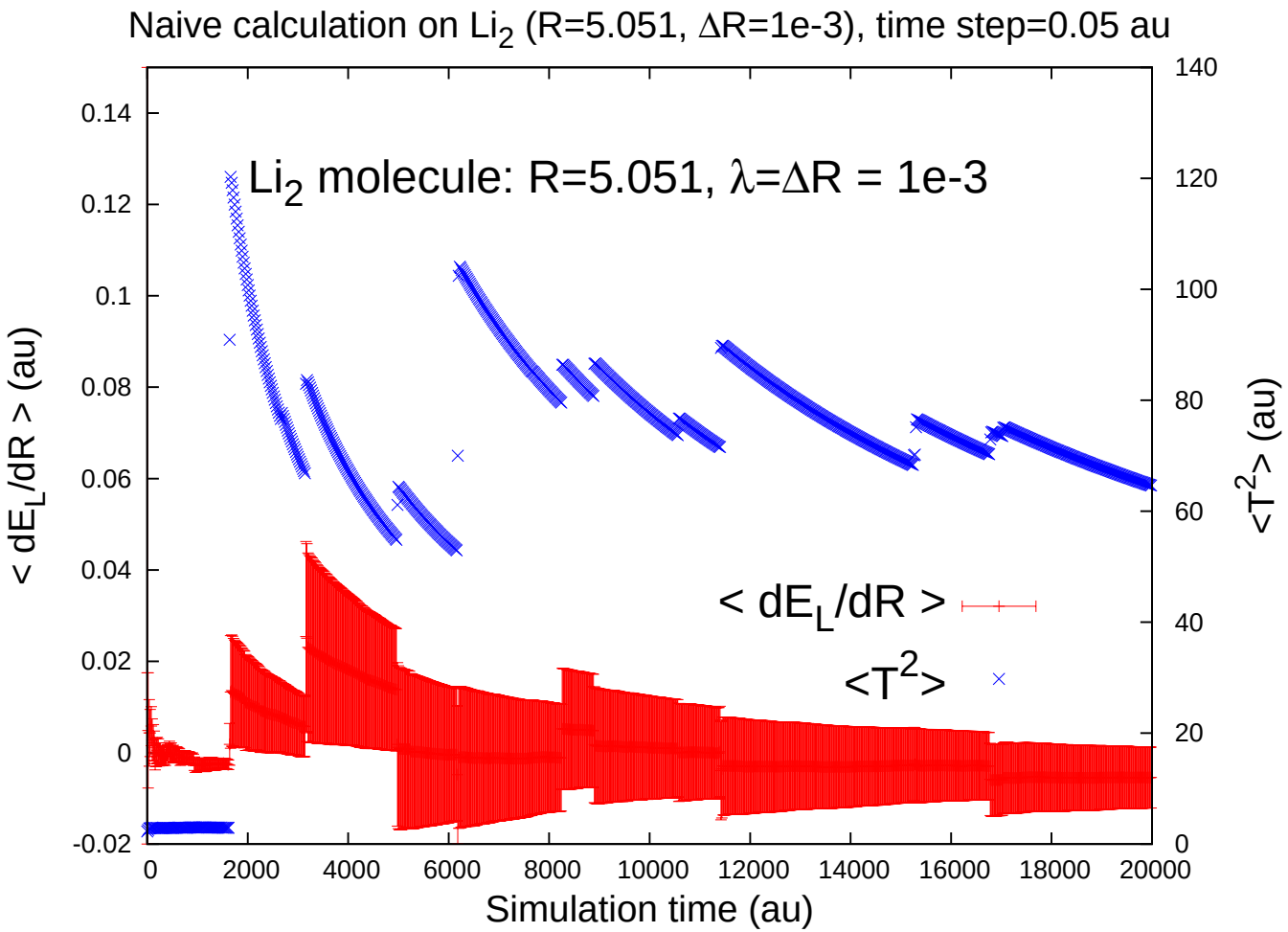




Fig.: Finite step error for the energy its derivative. $H_2$ molecule, $R = 0.8$. $\Psi$ = simple CSF $\times$ Full optimized jastrow

# Calculation on $Li_2$

2 VMC with same pseudo random numbers $(U = Id)$



# Origin of the instabilities for Li2 ?

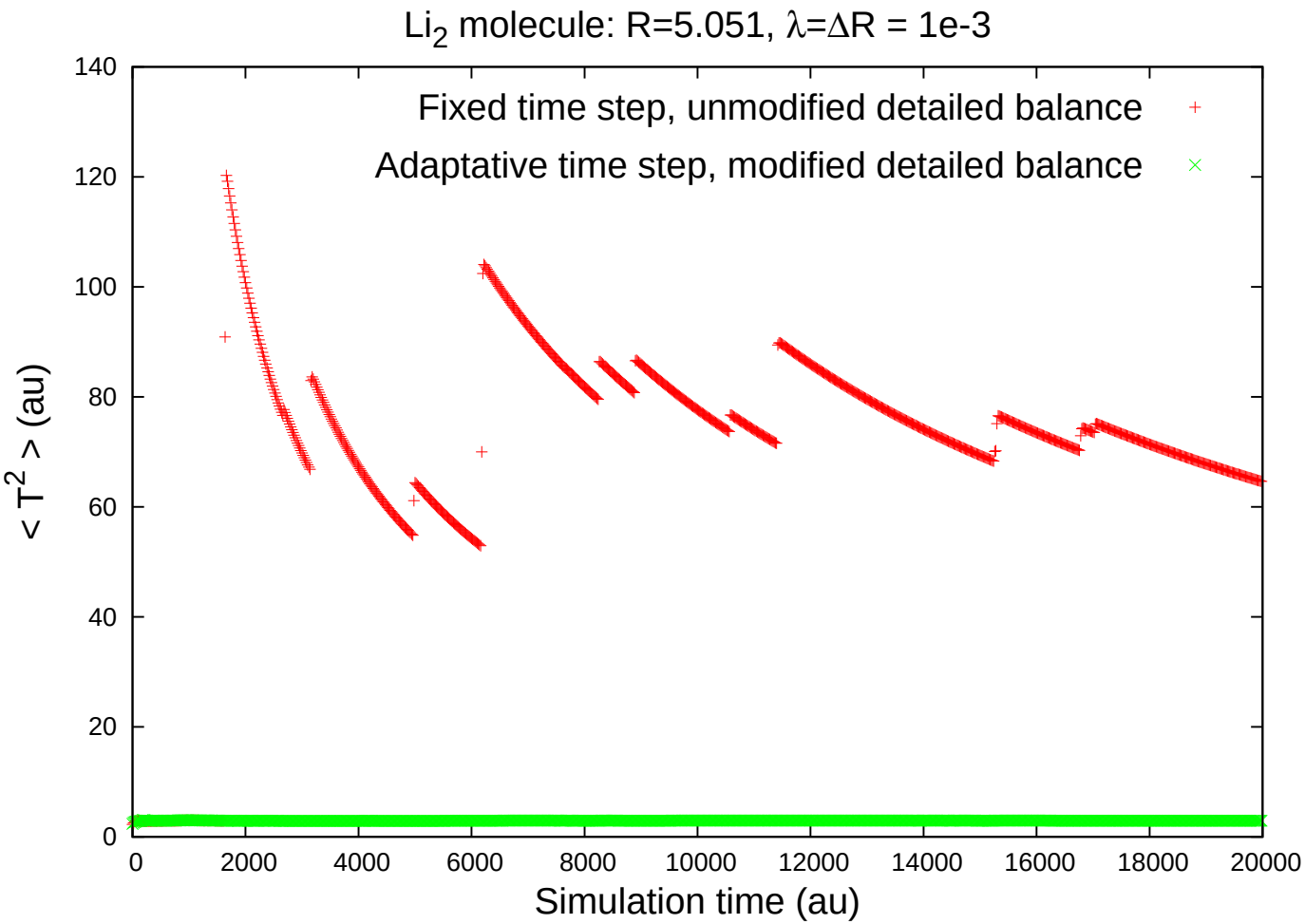
Finite time is a destabilization factor

## Proposition

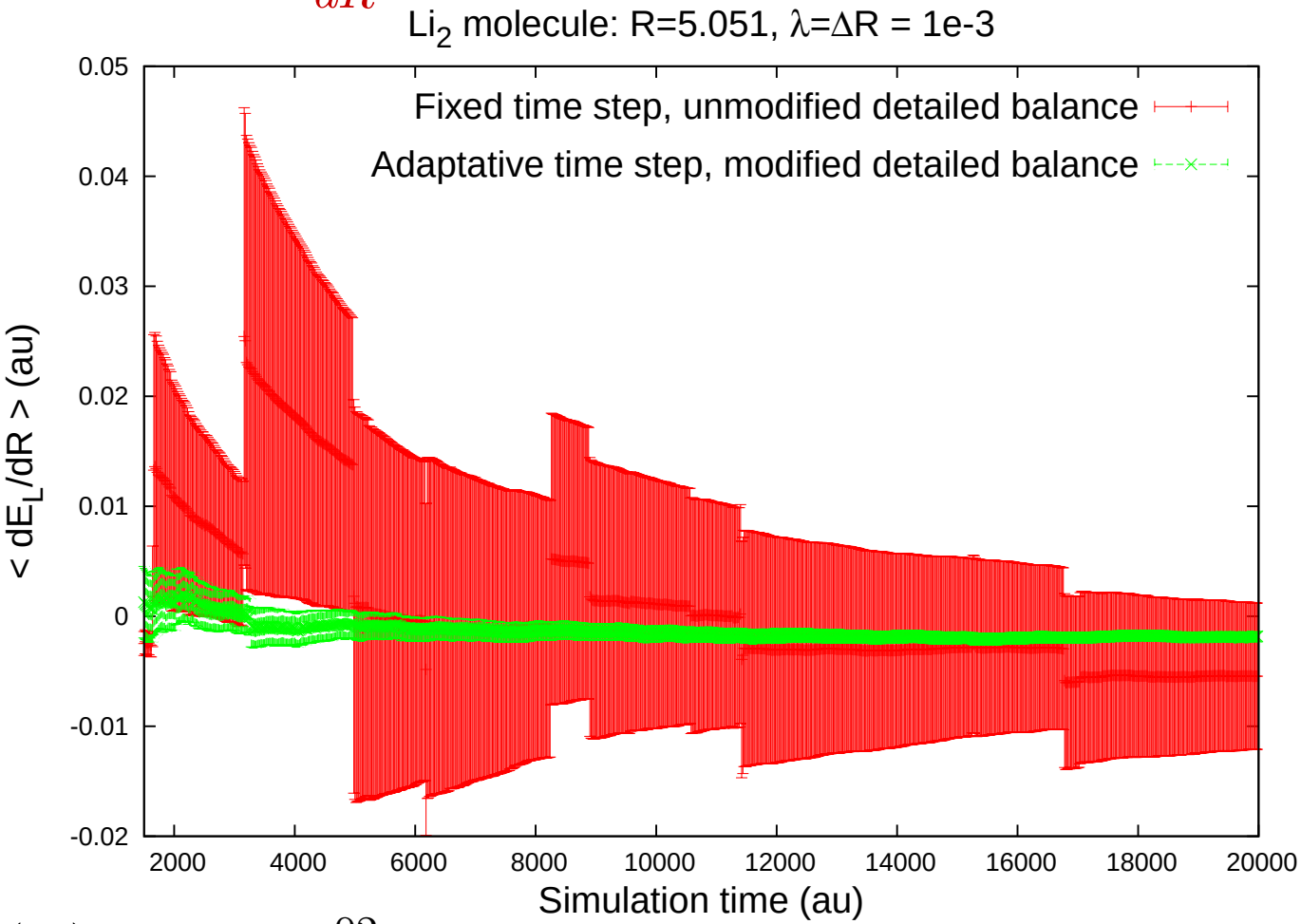
- $P_{acc}(\mathbf{R}, \mathbf{R}_\lambda) = \min[P_{acc}(\mathbf{R}), P_{acc}(\mathbf{R}_\lambda)]$
- Adaptative time step $\tau$ (smaller time steps in regions with large local kinetic energy)

Note : Finite time step error $O(\langle\tau\rangle)$.

# Effect on $\langle T^2 \rangle$



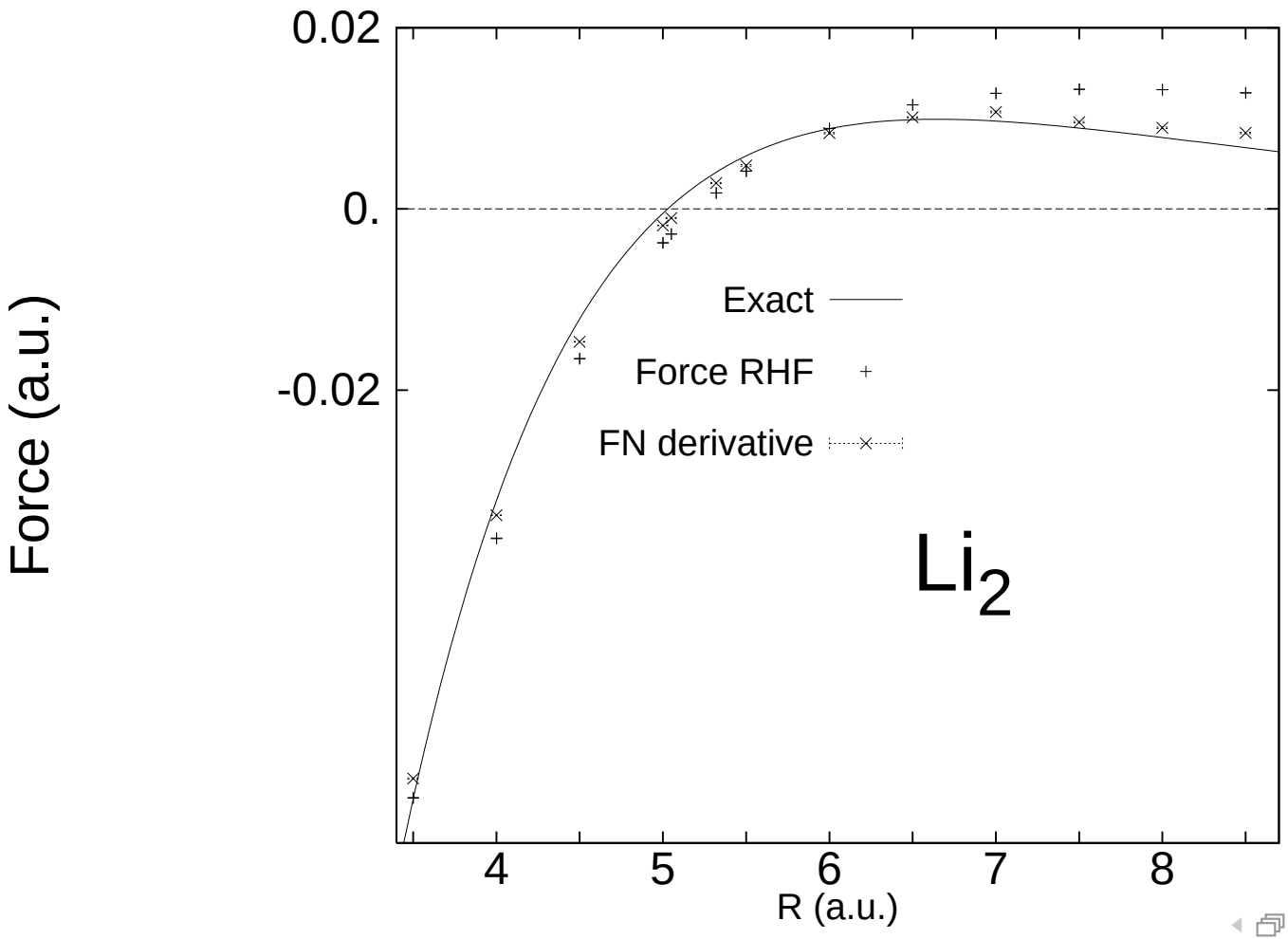
# Effect on $\frac{dE_L}{dR}$




$$\sigma(E) = 0.16e^{-02}$$

$$\sigma(\frac{\Delta E}{\Delta R}) = 0.43e^{-03} => \sigma(\Delta E) = 0.43e^{-06} << \sigma(E)\,!\,!$$

# Application : Force calculation in the dissociation of $Li_2$



# Impact of the unitary transform

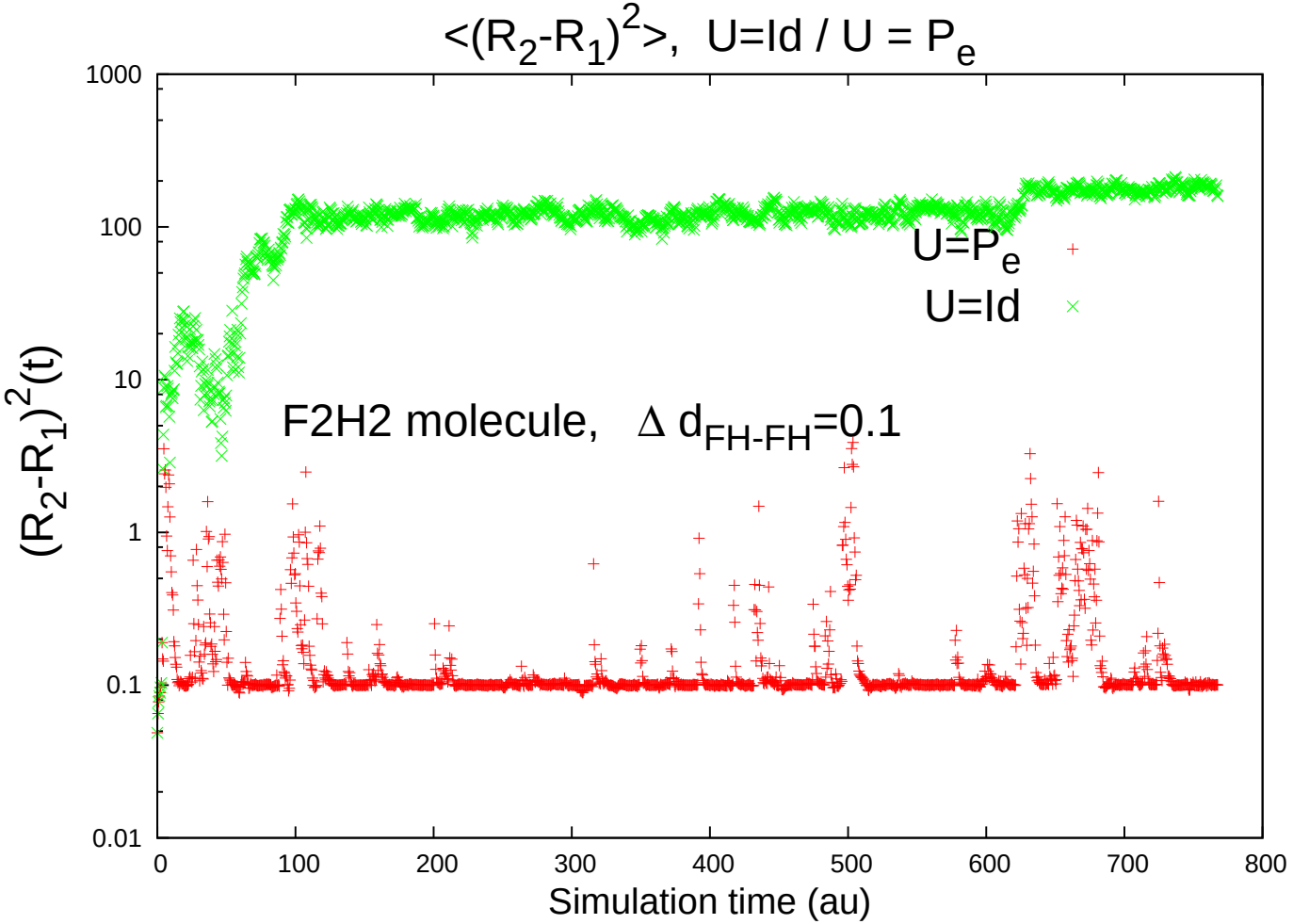



Fig.: F2H2, 2 geometries $\Delta d_{FH-FH} = 0.1$, $U = Id$ / $U = P_e$ permutation in the space of electrons (the closest same spin electrons have the same random numbers).

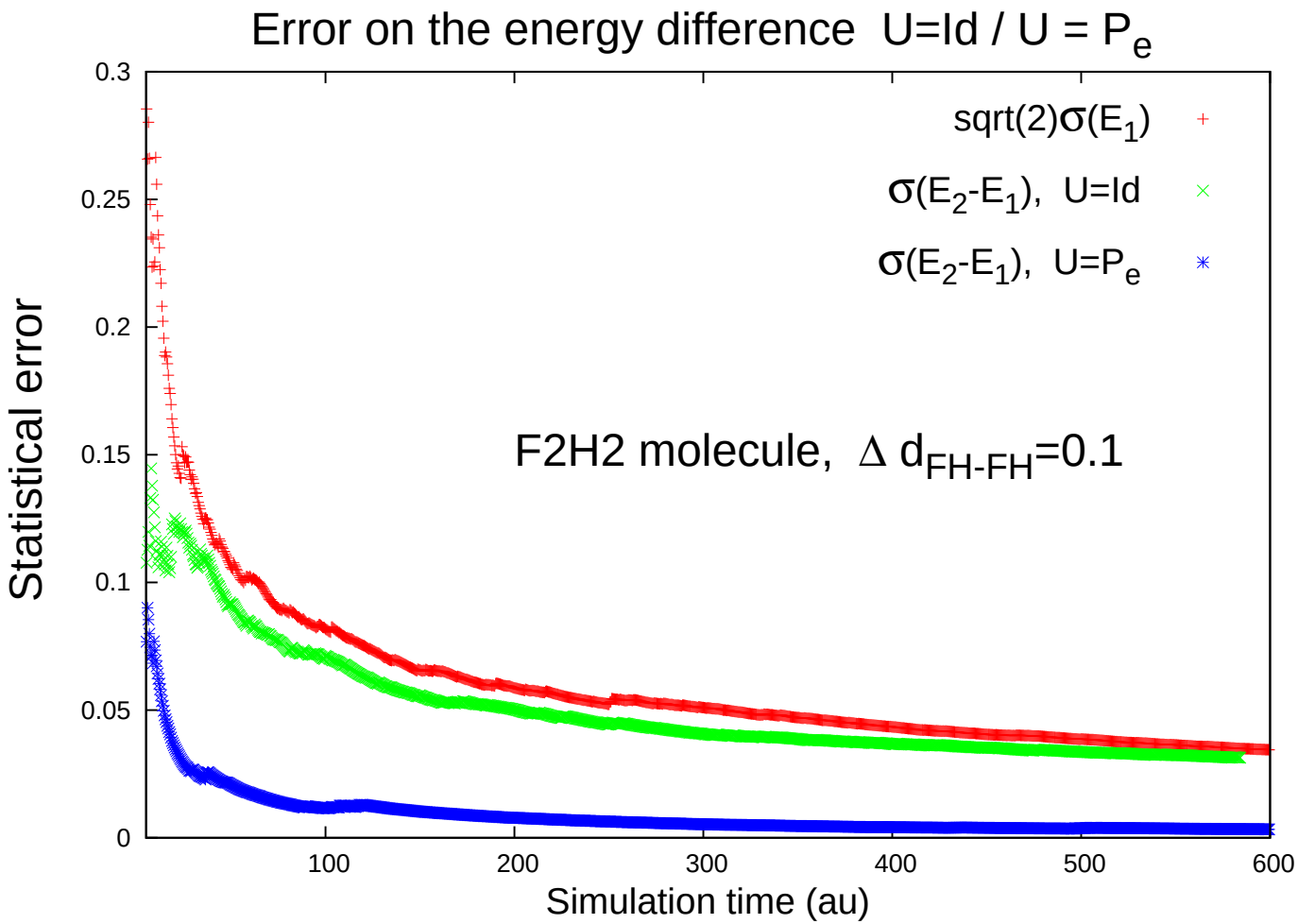



Fig.: F2H2, 2 geometries $\Delta d_{FH-FH} = 0.1$, $U = Id$ / $U = P_e$ permutation in the space of electrons (the closest same spin electrons have the same random numbers).

# Stable case

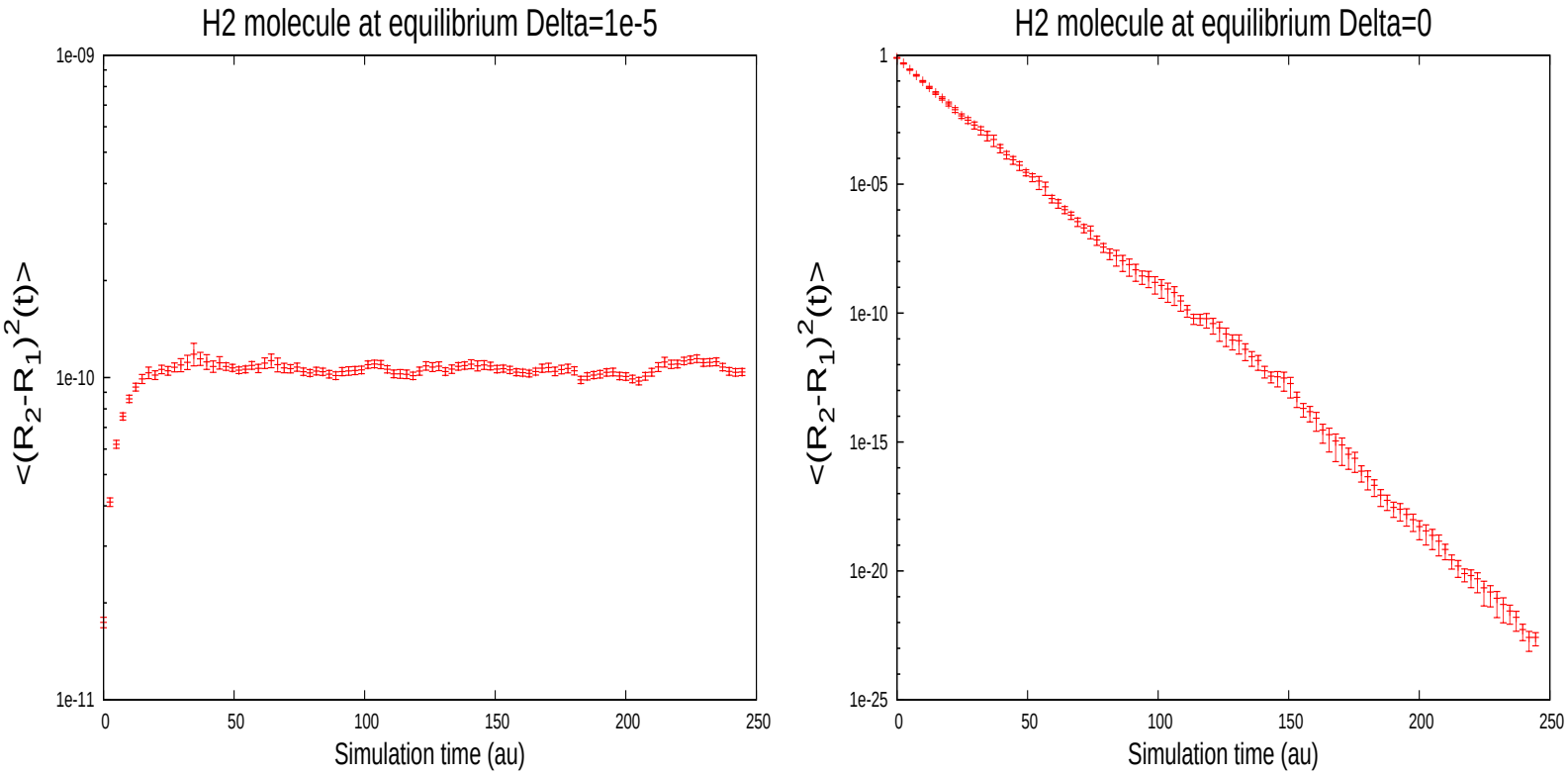



Fig.: Quadratic distance between the processes / sensitivity of the unperturbed process to initial conditions (H2 molecule).

# Unstable case

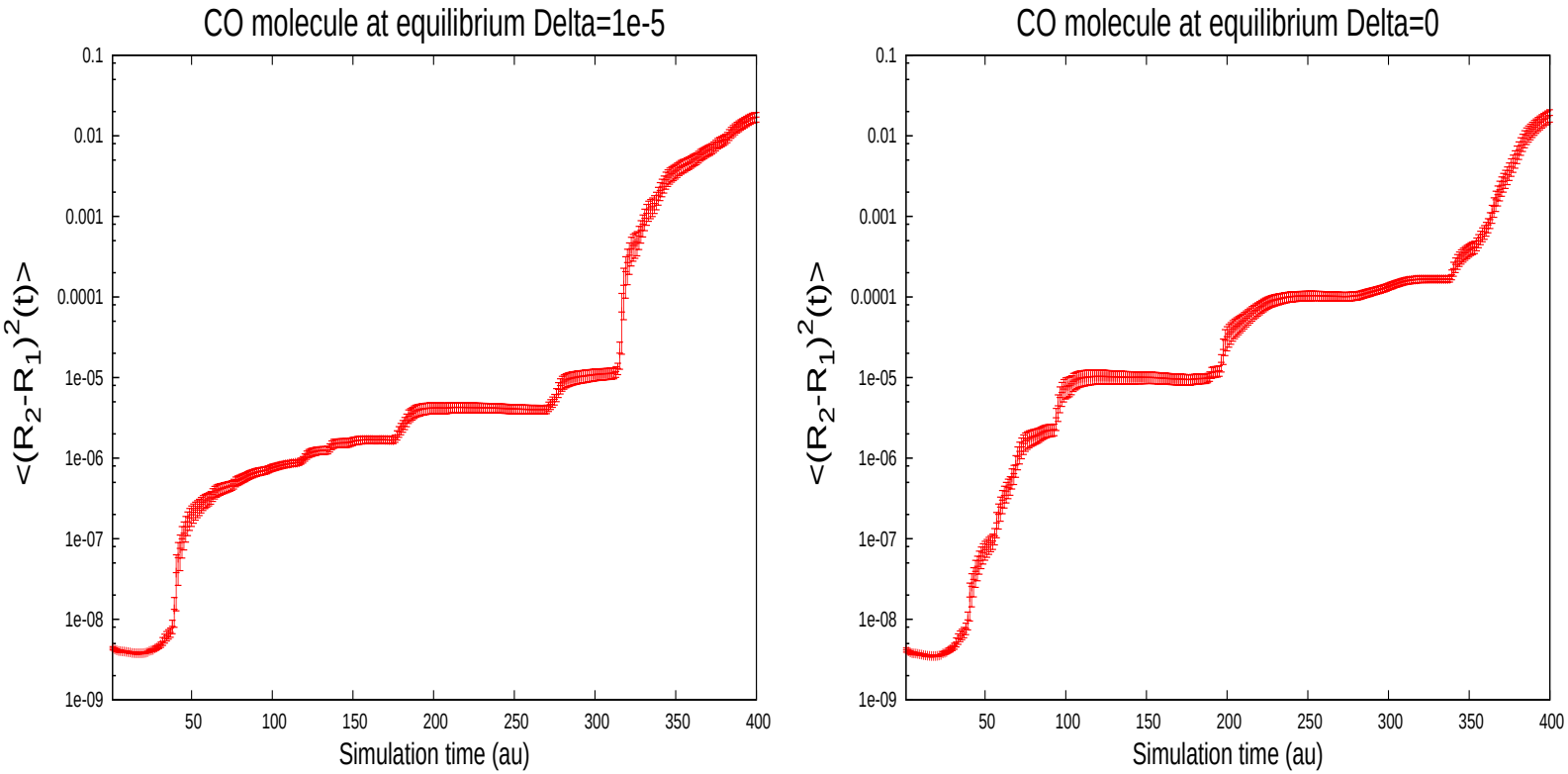



Fig.: Quadratic distance between the two different processes / sensitivity of the unperturbed process to initial conditions (CO molecule.

## Some comparisons DMC energies differences correlated / independent

Calculation of $\Delta E = E(d_{A-B} + \Delta R) - E(d_{A-B})$

| Molecule | $\Delta R$ (B) | Geometry | gain on $\sigma(\Delta)$ | cpu gain |
|---|---|---|---|---|
| Li8 | 1e-5 | eq | 7e4 | 0.5e10 |
| Li4 | 1e-5 | eq | 1e5 | 1e10 |
| Li2 | 1e-5 | 5.051 (eq) | 4e5 | 16e10 |
| Li2 | 1e-5 | 3.0 | 1.7e5 | 2.9e10 |
| Li2 | 0.1 | 3 | 42 | 1600 |
| Li2 | 0.1 | 1 | 3.6 | 13 |
| CO | 0.1 | 2.175 B (eq) | 3.9 | 15 |
| CO | 0.1 | 20 | 15 | 225 |
| Li4 | 0.1 | eq | 14 | 196 |
| F2H2 | 0.1 | eq | 10.3 | 106 |
| Li4 | 1 | eq | 5.6 | 31 |
| Li2 | 1 | eq | 5.25 | 26 |

# Conclusion and perspectives

## Small differences, derivatives of VMC energies

Basically a solved problem, though a lot of work has yet to be done (improving estimators).

## Generalized correlated sampling method for DMC small differences and derivatives

- Different guiding functions for the two systems.
- Pseudo-random numbers (Wiener process) related by a unitary transform.
- Finite time step stabilization (joint acceptation probability, adaptative time step).

## Main results

- Non trivial DMC energy derivatives obtained.
- DMC derivatives $\leftarrow$ chaotic properties of the dynamics.
- Substantial gain for small but finite differences ($\sim 0.1$) even for chaotic systems.
- Work still to be done to have $\langle T^2 \rangle$ always finite !

## Open questions

- Better insight needed into the chaotic properties of the Overdamped Langevin process.
- Stabilization techniques to be searched for.