Towards computing efficently cumulants in Monte Carlo, exchange cluster estimators.

Roland Assaraf[1], Hilaire Chevreau

[1] email: assaraf@lct.jussieu.fr Sorbonne Université CNRS, Laboratoire de Chimie Théorique CC 137 - 4, place Jussieu F-75252 PARIS CEDEX 05 - France

Monte Carlo methods adapted to statistical physics or quantum physics.

Quantum physics or statistical physics

- $R \in \Omega$ is a configuration (time trajectory in quantum physics or set of positions (and sometimes velocities) in statistical physics.
- Physical properties from logarithmic derivatives of integrals.

$$Z = \int dR e^{S(R)}$$

Examples of perturbations

Statistical physics $S = \beta H$ (the Hamiltonian).

-
$$\langle H \rangle = \frac{\int H \, e^{-\beta H} \, dR}{\int e^{-\beta H} \, dR}$$

Examples of perturbations

Statistical physics $S = \beta H$ (the Hamiltonian).

-
$$\langle H \rangle = \frac{\int H\, e^{-\beta H}\, dR}{\int e^{-\beta H}\, dR} = -\frac{1}{\beta}\frac{d \ln Z}{d\beta}$$

- Perturbation, addition of a magnetic field B.

$$H(R) \to H(R) + \underbrace{B \int M(R)}_{\text{perturbation}}$$

where M is the spin.
First derivative with respect to B mean magnetization, second derivative susceptibility

- Analogous formulas in quantum physics.

Second order cumulants or covariances

$$\mathrm{cov}(\mathrm{U},\mathrm{V}) = \mathbb{E}(\mathrm{UV}) - \underbrace{\mathbb{E}(U)\mathbb{E}(V)}_{\text{0 if centered}}$$

Second order cumulants or covariances

$$\mathrm{cov}(\mathrm{U},\mathrm{V}) = \mathbb{E}(\mathrm{UV}) - \underbrace{\mathbb{E}(U)\mathbb{E}(V)}_{\text{0 if centered}}$$

Size extensivity

A large system can be usually approximated as a set of independent fragments.

$$U \simeq \sum_m U_m$$
$$V \simeq \sum_m V_m$$

$m \neq n \Longrightarrow U_m$ independent of U_n and V_n, V_m ind. of V_n.

First order derivative

$\mathbb{E}(U) = \sum_m \mathbb{E}(U_m) = O(N)$

First order derivative

$\mathbb{E}(U) = \sum_m \mathbb{E}(U_m) = O(N) \quad V(U) = \sum_m V(U_m) = O(N)$.

No signal / noise problem.

First order derivative

$\mathbb{E}(U) = \sum_m \mathbb{E}(U_m) = O(N) \quad V(U) = \sum_m V(U_m) = O(N)$.

No signal / noise problem.

Second order derivative

$\mathrm{cov}(\mathrm{U}, \mathrm{V}) \simeq \sum \mathrm{cov}(\mathrm{U_m}, \mathrm{V_m})$ is extensive.

$$UV = \sum_{mn} U_m V_n$$

First order derivative

$\mathbb{E}(U) = \sum_m \mathbb{E}(U_m) = O(N) \quad V(U) = \sum_m V(U_m) = O(N)$.
No signal / noise problem.

Second order derivative

$\mathrm{cov(U,V)} \simeq \sum \mathrm{cov(U_m, V_m)}$ is extensive.

$$UV = \sum_{mn} U_m V_n = \underbrace{\sum_m U_m V_m}_{O(N)\text{terms}} + \underbrace{\sum_{m \neq n} U_m V_n}_{O(N^2)\text{ terms}}$$

$O(N^2)$ terms not contributing to the expectation value but to the variance.

First order derivative

$\mathbb{E}(U) = \sum_m \mathbb{E}(U_m) = O(N) \quad V(U) = \sum_m V(U_m) = O(N)$.
No signal / noise problem.

Second order derivative

$\mathrm{cov(U,V)} \simeq \sum \mathrm{cov(U_m,V_m)}$ is extensive.

$$UV = \sum_{mn} U_m V_n = \underbrace{\sum_m U_m V_m}_{O(N)\text{terms}} + \underbrace{\sum_{m \neq n} U_m V_n}_{O(N^2)\text{ terms}}$$

$O(N^2)$ terms not contributing to the expectation value but to the variance.

Indeed $m \neq n$, $\mathrm{cov(U_m,V_n)} = 0$ but $V(U_m V_n) = V(U_m)V(V_n)$.

The variance of the estimator grows as $O(N^2)$ while the expectation value grows as $O(N)$

The scaling of the variance is even larger for higher order cumulants.

The scaling of the variance is even larger for higher order cumulants.

Finite perturbation $S \to S + P$

$$Z_P = \int e^{-S-P} = \frac{\int e^{-S} e^{-P}}{\int e^{-S}} Z$$

The scaling of the variance is even larger for higher order cumulants.

Finite perturbation $S \to S + P$

$$Z_P = \int e^{-S-P} = \frac{\int e^{-S} e^{-P}}{\int e^{-S}} Z = \mathbb{E}(e^{-P}) Z \tag{1}$$

Example, the sign problem !

Looking at a fermionic problem as perturbation of a bosonic problem.

$P = i\pi \int n$ (where $n \in (0,1)$).

$$\ln(Z_P) = \ln(Z) + \underbrace{\ln(\mathbb{E}(e^{-P}))}_{\text{Infinite sum of cumulants}} \tag{2}$$

Noise (exponential) / signal ($O(N)$) growing exponentially with system size, the so-called sign problem.

The size extensivity problem (cumulant problem or the sign problem) is formulated in the limit of independent fragments.

But in the limit of (explicitely) independent fragments there should not a be size extensive problem or a sign problem !

The size extensivity problem (cumulant problem or the sign problem) is formulated in the limit of independent fragments.

But in the limit of (explicitely) independent fragments there should not a be size extensive problem or a sign problem !

If each fragment can be treated independently, size extensivity of the fluctuations !

The size extensivity problem (cumulant problem or the sign problem) is formulated in the limit of independent fragments.

But in the limit of (explicitely) independent fragments there should not a be size extensive problem or a sign problem !

If each fragment can be treated independently, size extensivity of the fluctuations !

How to exploit (approximate) independence to compute cumulants with size extensive fluctuations ?

In the litterature

Dynamical sign problem solved

Uses the Markovian property in the time.
Story for $X_{t>0}$ depends on X_0 but not of $X_{t<0}$. This high degree of independence is used in the Inchworm algorithm.

No such strong explicit independence for particles but partial solutions.

Cluster algorithms

- Spin models, flipping domains or clusters of spins (e.g. Wolf).

  Reduces the scaling of the fluctuations for the covariances $(O(N))$ *Wolf, Nuc. Phys. B [1988]*

- Domain exchange algorithm use a pair of replicas of the system. Ising models (*Chayes, J. Stat. Phys. (1998)*) and lattice models with a Z_2 symetry. *M. Hasenbusch, Phys. Rev. E 97, 012119 (2018)*.

Generalization to Two-body interacting system

Pair wise interacting system

$$Z = \int p(r)dr \qquad \text{with} \quad p(r) = \prod_{i,j} w_{ij}(r_i, r_j)$$

Examples

Statistical physics

$$Z = \int e^{-\beta \sum_{i,j} v(r_i, r_j)}$$

Generalization to Two-body interacting system

Pair wise interacting system

$$Z = \int p(r)dr \qquad \text{with} \quad p(r) = \prod_{i,j} w_{ij}(r_i, r_j)$$

Examples

Statistical physics

$$Z = \int e^{-\beta \sum_{i,j} v(r_i, r_j) - \beta \sum_i \dot{r_i}^2}$$

Generalization to Two-body interacting system

Pair wise interacting system

$$Z = \int p(r)dr \qquad \text{with} \quad p(r) = \prod_{i,j} w_{ij}(r_i, r_j)$$

Examples

Statistical physics

$$Z = \int e^{-\beta \sum_{i,j} v(r_i,r_j) - \beta \sum_i \dot{r_i}^2}$$

Quantum physics $Z = \int e^{-\int dt\mathcal{L}(r,\dot{r})}$ (Feyman integral)

$$dt\mathcal{L}(r,\dot{r}) \underbrace{\simeq}_{\text{Trotter}} \frac{1}{2dt}\sum_i (r_i(t+dt) - r_i(t))^2 + dt\sum_{ij} v(r_i(t), r_j(t))$$

Exchange cluster algorithms

Pairwise probability density to be sampled p.

Defining an independent replicas

$r \in \Omega$

Exchange cluster algorithms

Pairwise probability density to be sampled p.

Defining an independent replicas

$r \in \Omega \to (r^1, r^2) \in \Omega^1 \times \Omega^2$

$$P(r^1, r^2) = p(r^1)p(r^2) \tag{3}$$

Building links between indices of the variables

w_{ij}^{11} interaction between r_i^1 and r_j^1
w_{ij}^{12} interaction between r_i^1 and r_j^2 (particle j of system 2 put in 1).
w_{ij}^{21} interaction between r_i^2 and r_j^1 (particle j of system 1 put in 2).

Probability to link i and j

$$1 - \min\left(\frac{w_{ij}^{12} w_{ij}^{21}}{w_{ij}^{11} w_{ij}^{22}}, 1\right) \tag{4}$$

Building domains

- A domain (cluster) is a list of linked indices.
- An indix $\Leftrightarrow$ pair of variables $\in \Omega^1 \times \Omega^2$.
- Domain (cluster) list of pairs of variables belonging to the two replicas.

Domains can be exchanged at will between the two replicas !

This operation leaves the joint density $P(r^1, r^2) = p(r^1)p(r^2)$ invariant.

proof: checking the detailed balance property

Intuitive and physical interpretation

Some illustrative properties

- Probability to unlink $(i,j) = 1 \iff w_{ij}^{12} w_{ij}^{21} \geq w_{ij}^{11} w_{ij}^{22}$
  $\iff$ favorizes exchanging one particle i or j.
  $\Longrightarrow$ If (i,j) are not indirectly linked they belong to different domains.
- If (i,j) not interacting in the two systems $\Longrightarrow$ probability to unlink $(i,j) = 1$

The more two fragments are independent the more frequent they can be separately replaced by another fragment belonging to the other replica

Domains in the Lennard Jones model

Lennard Jones model

- Particles in a 3-dimensional box.
- Interaction between particle i and j

$$u_{ij} = 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}] \quad (5)$$

where r_{ij} is the distance between particle i and j.
$\sigma = 3.4A$
$\frac{\epsilon}{k} = 1.00568$KJ. mol^{-1}
Density 1 particle for a sphere of radius $10A$

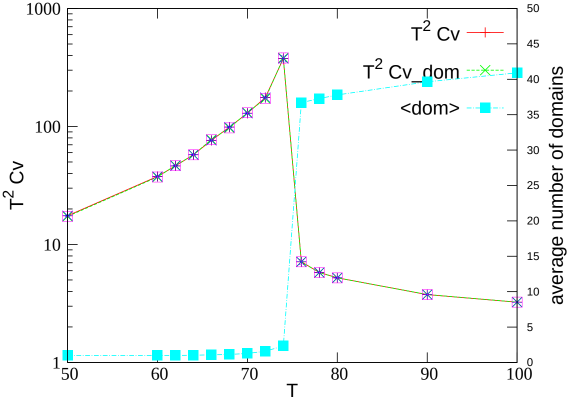

Figure: Average number of domains and heat capacity per particle (Lennard Jones model) N=50 particles in a 59×59 box

Sampling the exchange of domains improve ergodicity but is a tool to reduce the scaling of the variance

The exchange domain operators $\hat{D}$ form a commutative algebra of 2^{N_d} P invariant and self-adjoint operators, which can be used to build 2^{N_d} control variates.

$$\hat{D}(P) = P \Longrightarrow \mathbb{E}(\hat{D}(O)) = \mathbb{E}(O)$$

proof (I.P.P.) $\int OP = \int \hat{D}(P) O = \int P \hat{D}(O)$

Computation of covariances

$U = \frac{1}{2} \sum_{i,j} u_{ij}$ and $V = \frac{1}{2} \sum_{i,j} v_{ij}$

Computation of covariances

$U = \frac{1}{2}\sum_{i,j} u_{ij}$ and $V = \frac{1}{2}\sum_{i,j} v_{ij}$

$$\mathrm{cov(U,V)} \equiv \mathbb{E}(\mathrm{U,V}) - \mathbb{E}(\mathrm{U})\mathbb{E}(\mathrm{V})$$

Computation of covariances

$U = \frac{1}{2}\sum_{i,j} u_{ij}$ and $V = \frac{1}{2}\sum_{i,j} v_{ij}$

$$\mathrm{cov}(\mathrm{U},\mathrm{V}) \equiv \mathbb{E}(\mathrm{U},\mathrm{V}) - \mathbb{E}(\mathrm{U})\mathbb{E}(\mathrm{V})$$

$\frac{1}{2}(U^1 - U^2)(V^1 - V^2)$ unbiased estimator on the replicas.

$$\frac{1}{2}(U^1 - U^2)(V^1 - V^2) = \frac{1}{8}\sum_{i,j,k,l}(u_{ij}^{11} - u_{ij}^{22})(v_{kl}^{11} - v_{kl}^{22}) \qquad (6)$$

Computation of covariances

$U = \frac{1}{2} \sum_{i,j} u_{ij}$ and $V = \frac{1}{2} \sum_{i,j} v_{ij}$

$$\mathrm{cov(U,V)} \equiv \mathbb{E}(\mathrm{U,V}) - \mathbb{E}(\mathrm{U})\mathbb{E}(\mathrm{V})$$

$\frac{1}{2}(U^1 - U^2)(V^1 - V^2)$ unbiased estimator on the replicas.

$$\frac{1}{2}(U^1 - U^2)(V^1 - V^2) = \frac{1}{8} \sum_{i,j,k,l} (u_{ij}^{11} - u_{ij}^{22})(v_{kl}^{11} - v_{kl}^{22}) \qquad (6)$$

Basic Idea

D_{kl} beeing the minimal domain containing (k,l). If $D_{kl} \bigcap D_{ij} = \emptyset$

$$\frac{1}{2}(1 + \hat{D}_{kl})((u_{ij}^{11} - u_{ij}^{22})(v_{kl}^{11} - v_{kl}^{22})) = 0$$

The sum (6) is reduced to $O(N)$ terms, and the variance is $O(N)$ down from $O(N^2)$!

One simple improved estimator of the covariance. Let (m, n) be a pair of domains.

Interactions between two domains

$$U_{mn}^{11} \equiv \sum_{(i,j)\in D_m \times D_n} u_{ij}^{11}$$

One simple improved estimator of the covariance. Let (m, n) be a pair of domains.

Interactions between two domains

$$
\begin{aligned}
U_{mn}^{11} &\equiv \sum_{(i,j)\in D_m \times D_n} u_{ij}^{11} - u_{ij}^{22} \\
V_{mn}^{11} &\equiv \sum_{(i,j)\in D_m \times D_n} v_{ij}^{11} - v_{ij}^{22}
\end{aligned}
\tag{7}
$$

One simple improved estimator of the covariance. Let (m, n) be a pair of domains.

Interactions between two domains

$$\begin{aligned} U_{mn}^{11} &\equiv \sum_{(i,j)\in D_m\times D_n} u_{ij}^{11} - u_{ij}^{22} \\ V_{mn}^{11} &\equiv \sum_{(i,j)\in D_m\times D_n} v_{ij}^{11} - v_{ij}^{22} \end{aligned} \tag{7}$$

Interaction between one domain m and the other domains.

$$\begin{aligned} \mathcal{U}_m^{11} &\equiv \frac{U_{mm}^{11}}{2} + \sum_{p\neq m} U_{mp}^{11} \\ \mathcal{V}_m^{11} &\equiv \frac{V_{mm}^{11}}{2} + \sum_{p\neq m} V_{mp}^{11} \end{aligned}$$

Estimator of the covariance

$$\tilde{\chi} = \frac{1}{2}\sum_{m}\mathcal{U}_m^{11}\mathcal{V}_m^{11} - \frac{1}{2}\sum_{m<n} U_{mn}^{11} V_{mn}^{11} \tag{8}$$

$O(N)$ terms since $U_{mn} \to 0$ and $V_{mn} \to 0$ if D_m far from D_n.

Size extensivity of the variance of χ

Specific heat (Lennard Jones model)

$C_v \equiv \frac{k}{T^2}(< U^2 > - < U >^2) = \frac{k}{T^2}\mathrm{cov(U,U)}$

where U is the Lennard Jones potential

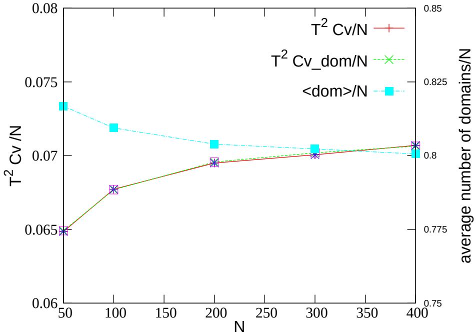

Figure: Average number of domains and heat capacity per particle (Lennard Jones model), $T = 100K$

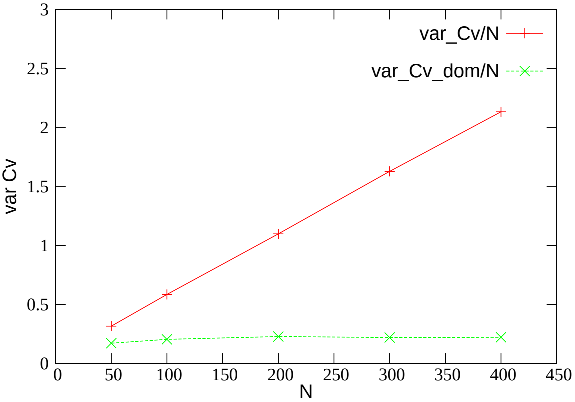

Figure: Variance of estimators of the heat capacity per particle

- Method for a general pairwise interacting variables model to compute covariances with size extensive variance $O(N)$ down from $O(N^2)$.
- Based on an exchange cluster algorithm, using an independent replica.
- Proof of concept on a Lennard Jones model (continous model with no Z_2 symmetry).

- Method for a general pairwise interacting variables model to compute covariances with size extensive variance $O(N)$ down from $O(N^2)$.
- Based on an exchange cluster algorithm, using an independent replica.
- Proof of concept on a Lennard Jones model (continous model with no Z_2 symmetry).

Work in progress

- Extension to higher order cumulants (with H. Chevreau).
- Extension to quantum bosonic systems.
- Other method applicable to non pair-wise systems (Variational and Diffusion Monte Carlo) with A. Bienvenu and J. Feldt. Using conditional expectation values (side walks).