

Comparaison théorique et numérique de méthodes d'échantillonnage en dynamique moléculaire.

Eric CANCES¹, Frédéric LEGOLL^{1,2}, [Gabriel STOLTZ](#)^{1,3}

¹ CERMICS, ENPC. Marne-la-Vallée

² IMA, Minneapolis

³ CEA/DAM, Bruyères-le-Châtel,
<http://cermics.enpc.fr/~stoltz/home.html>

- E. Cancès, F. Legoll, et G. Stoltz, Theoretical and numerical comparison of some sampling methods for molecular dynamics, IMA Preprint, **2040** (2005)
- Présentation disponible sur <http://cermics.enpc.fr/~stoltz/semCEA.pdf> (lien "Séminaire PMC du 17 juin 2005")

- Calcul de propriétés d'équilibre du système

$$\langle A \rangle = \int_{T^* \mathcal{M}} A(q, p) d\mu(q, p)$$

où μ est la mesure canonique (NVT)

$$d\mu(q, p) = Z^{-1} \exp(-\beta H(q, p)) dq dp,$$

avec $\beta = 1/k_B T$ et

$$H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q),$$

Ce qu'on veut faire (2)

- Calcul sur une variété \mathcal{M} . Par exemple, $\mathcal{M} = \mathbb{T}^{3N}$ si conditions de bord périodiques, auquel cas $T^*\mathcal{M} = \mathbb{T}^{3N} \times \mathbb{R}^{3N}$.
- Situations plus compliquées (Blue Moon sampling)
- Calcul de propriétés dynamiques selon

$$\langle B \rangle(t) = \int_{T^*\mathcal{M}} B(\Phi_t(q, p), (q, p)) d\mu$$

Comment on peut le faire

- Echantillonnage de la partie configurationnelle (partie cinétique triviale)

$$d\pi(q) = f(q) dq = Z_q^{-1} e^{-\beta V(q)} dq.$$

- Approximation de $\langle A \rangle$ de la forme

$$\frac{1}{N} \sum_{n=0}^{N-1} A(q^n, p^n) \simeq \int_{\mathcal{M}} A(q) d\mu(q)$$

- Calcul des $\{q^n\}$? Convergence et vitesse de convergence?

Comment on peut le faire (2)

Type 1 $(q^n)_{n \in \mathbb{N}}$ **variables aléatoires i.i.d.**, loi de densité

$$f(q) = Z_q^{-1} e^{-\beta V(q)}$$

Type 2 $(q^n)_{n \in \mathbb{N}}$ réalisation d'une **chaîne de Markov** à espace d'états continu, dont π est une mesure invariante (id. μ)

Type 3 $(q^n)_{n \in \mathbb{N}}$ est une approximation $(q_{t_n})_{n \in \mathbb{N}}$ d'une réalisation d'un **processus stochastique** $(q_t)_{t \geq 0}$ dont π est une mesure invariante (id. μ)

Type 4 $(q^n)_{n \in \mathbb{N}}$ est une approximation de $(q(t_n))_{n \in \mathbb{N}}$ où $(q(t), p(t), x(t))_{t \geq 0}$ trajectoire d'un **système déterministe étendu**, mesure invariante $d\rho$, dont la projection est $d\mu$

Comment on peut le faire (3)

- Une longue trajectoire
- Plusieurs trajectoires courtes
- Evoluer une distribution initiale
- Vitesse de convergence théorique et pratique
- Rappel et mise en garde : on veut faire de l'échantillonnage! **Pas d'interprétation physique** à faire des méthodes proposées...
- Méthodes "historiques" en DM : isocinétique, Andersen

Plan de la présentation

- Présentation des différentes méthodes d'échantillonnage (stochastiques, mixtes, déterministes, combinaison)
- Comparaison numérique sur un cas test
- Une application sous S-TAMP
- Des perspectives

- Variables aléatoires i.i.d. = méthode du **rejet**

Algorithme 1 (Rejet) *Pour $n \geq 0$,*

1. *générer un point $\tilde{q} \in \mathcal{M}$ selon g et calculer*

$$r = \frac{f(\tilde{q})}{cg(\tilde{q})};$$

2. *tirer $s \sim \mathcal{U}[0, 1]$*

3. *si $s > r$ retour en (1) (rejet proposition \tilde{q}), sinon aller en (4);*

4. *prendre $q^n = \tilde{q}$ (accepter proposition \tilde{q}); remplacer n par $n + 1$ et retour en (1).*

Méthodes stochastiques : Rejet (2)

- **Convergence** de la moyenne empirique garantie par LGN. On pose $S_N(A) = \sum_{n=0}^{N-1} A(q^n)$. Alors

$$\pi(|A|) < +\infty \Rightarrow \lim_{N \rightarrow \infty} \frac{1}{N} S_N(A) = \int_{\mathcal{M}} A d\pi \quad \text{a.s.}$$

- **Vitesse** de convergence garantie par TCL :

$$\pi(|A|^2) < +\infty \Rightarrow S_N(\bar{A}) \xrightarrow{N \rightarrow \infty} \mathcal{N}(0, 1),$$

avec $\gamma_A > 0$, $\bar{A} = A - \int_{\mathcal{M}} A d\pi$.

- Problèmes : demande bon **choix de g** (*rejection control*), peu d'acceptation, surtout si la dimension de l'espace est grande

- Version "chaîne de Markov" du rejet = MIS

Algorithme 2 (Metropolized independence sampling)

Soit une configuration initiale q^0 . Posant $w = f/g$ et pour $n \geq 1$,

1. tirer un point $\tilde{q} \in \mathcal{M}$ selon la densité g
2. tirer $s \sim \mathcal{U}[0, 1]$
3. si $s \leq \min \left\{ 1, \frac{w(\tilde{q})}{w(q^n)} \right\}$, on prend $q^{n+1} = \tilde{q}$, sinon $q^{n+1} = q^n$;
4. remplacer n par $n + 1$ et retour en (1).

Méthodes stochastiques : MIS (2)

- Convergence de la méthode : théorie des chaînes de Markov à espaces d'état continus (S.P. Meyn et R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer (1993))
- Convergence garantie de la **moyenne empirique pour une réalisation** quand
 1. invariance de la mesure canonique
 2. propriété d'irréducibilité du noyau de transition

$$\lambda^{\text{Leb}}(B) > 0 \Rightarrow P(q, B) > 0$$

- **Vitesse** de convergence de la moyenne empirique : analogue du TCL

Méthodes mixtes stochastiques/déterministes

- Dynamique Hamiltonienne sur la sous-variété

$$T^*\mathcal{M}(E_0) = \{(q, p) \in T^*\mathcal{M}; H(q, p) = E_0\}$$

avec énergie initiale $E_0 = H(q_0, p_0)$

- Besoin de **perturbations** pour échantillonner toutes les sous-variétés $T^*\mathcal{M}(E_0)$ de \mathcal{M}
- Plusieurs types de perturbations :
 1. perturbations fortes mais espacées (HMC)
 2. perturbation en continu (diffusion brownienne, Langevin)

- Chaîne de Markov de type Métropolis-Hastings avec fonction de proposition fondée sur DM

Algorithme 3 (Hybrid Monte Carlo) Soit $q^0 \in \mathcal{M}$ et $\tau > 0$. Pour $n \geq 0$,

1. tirer les moments p^n selon la distribution canonique, calculer $E^n = H(q^n, p^n)$
2. intégrer en dynamique NVE $\Phi_\tau(q^n, p^n) = (p^{n,\tau}, q^{n,\tau})$
3. calculer l'énergie $E^{n,\tau} = H(q^{n,\tau}, p^{n,\tau})$. Accepter position finale avec probabilité $\min\{1, e^{-\beta(E^{n,\tau} - E^n)}\}$
4. remplacer n par $n + 1$ et revenir en (1).

Méthodes mixtes : HMC (2)

- **Convergence** de la moyenne empirique sous des hypothèses difficiles à vérifier : C. Schütte, Habilitation thesis (1998).
- Une preuve simplifiée et qu'on comprend :
 1. **invariance** de la mesure canonique
 2. propriété d'**irréducibilité** montrée grâce à un principe de moindre action sous les hypothèses $V \in C^1$ borné supérieurement et ∇V globalement Lipschitz.
- Extension de la démonstration au cas **discrétisé**
- Choix du pas de temps en pratique afin que le taux de rejet soit faible. Choix de τ ?

Méthodes mixtes : HMC (3)

- Quid si potentiel pas borné supérieurement?
Exemple de l'oscillateur harmonique avec $\tau =$ période fondamentale : $q^n = q^0 \forall n !$
- Pour obtenir une dynamique ergodique il faut recourir à des **temps τ aléatoires**
- Preuve de convergence faite par perturbations de la dynamique à $\nabla V = 0$:

$$M\ddot{q}_\epsilon(t) = -\epsilon\nabla V(q_\epsilon(t))$$

i.e. limite temps courts/impulsions fortes.

- Equation différentielle stochastique

$$dq_t = -\nabla V(q_t)dt + \sigma dW_t,$$

où $(W_t)_{t \geq 0}$ est un processus de Wiener standard de dimension d et avec $\sigma = (2/\beta)^{1/2}$.

- Mesure canonique $d\pi$ **invariante** (solution stationnaire de l'équation de Fokker-Planck associée) et propriété d'irréducibilité (cf. processus elliptique), d'où

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(q_t^x) dt = \int_{\mathbb{R}^{3N}} A(q) d\pi \quad \text{a.s.}$$

Méthodes mixtes : diffusion brownienne (2)

- Implémentation numérique : différentes options
 1. Discrétisation "brutale" (Euler-Maruyama)

$$q^{n+1} = q^n - \frac{\Delta t^2}{2} \nabla V(q^n) + \beta^{-1/2} \Delta t R^n$$

avec $(R^n)_{n \in \mathbb{N}}$ suite de vecteurs gaussiens standards i.i.d.

2. Schémas **implicites** (difficile), voir *e.g.* Mattingly/Stuart/Higham
3. **Métropolisation** : la discrétisation type Euler-Maruyama fait office de fonction de proposition + procédure d'acceptation/rejet

Méthodes mixtes : diffusion brownienne (3)

- Convergence assurée sous hypothèses minimales pour la version métropolisée
- Pour les discrétisations sans Métropolisation, convergence vers une **mesure approchée** $\pi_{\Delta t}$ sous certaines hypothèses
- Un seul paramètre à choisir!
- Choix du pas de temps de temps en pratique (version métropolisée) : compromis entre **décorrélation et acception**. On montre (Roberts/Rosenthal) que le taux de rejet optimum est autour de 0.5

Méthodes mixtes : Langevin

- EDS hypo-elliptique (bruit sur p seulement)

$$\begin{cases} dq_t = M^{-1} p_t dt \\ dp_t = -\nabla V(q_t) dt - \xi M^{-1} p_t dt + \sigma dW_t \end{cases}$$

avec la relation de **fluctuation/dissipation**

$$\sigma = (2\xi/\beta)^{1/2}.$$

- Diffusion brownienne = limite $M \rightarrow 0$ de Langevin
- Mêmes preuves/résultats de convergence que pour diffusion brownienne

Méthodes mixtes : Langevin (2)

- Discrétisations analogues à diffusion brownienne (Métropolisation de la discrétisation d'Euler-Maruyama, schémas implicites ou explicites avec convergence vers mesure approchée)
- En pratique, **algorithme de BBK** = modification schéma de Verlet

$$\left\{ \begin{array}{l} p_i^{n+1/2} = p_i^n + \frac{\Delta t}{2} \left(-\nabla_{q_i} V(q^n) - \xi \frac{p_i^n}{m_i} + \frac{\sigma_i}{\sqrt{\Delta t}} R_i^n \right) \\ q_i^{n+1} = q_i^n + \Delta t \frac{p_i^{n+1/2}}{m_i} \\ p_i^{n+1} = \frac{1}{1 + \frac{\xi \Delta t}{2m_i}} \left(p_i^{n+1/2} - \frac{\Delta t}{2} \nabla_{q_i} V(q^{n+1}) + \sigma_i \frac{\sqrt{\Delta t}}{2} R_i^{n+1} \right) \end{array} \right.$$

Méthodes mixtes : Langevin (3)

- On montre qu'il faut modifier la relation de fluctuation/dissipation en pratique afin de maintenir la **bonne température cinétique** :

$$\sigma_i^{\Delta t} = \sqrt{\frac{2\xi}{\beta} \left(1 + \frac{\xi \Delta t}{2m_i} \right)}.$$

On retrouve la relation habituelle dans la limite $\Delta t \rightarrow 0$.

- Démonstration rigoureuse dans le cas $\nabla V = 0$, preuve numérique dans le cas général
- Paramètres à fixer : Δt , ξ (règle : $\frac{\xi \Delta t}{2m_i}$ "petit")

- Système étendu, variables (q, p, η, ξ)
- Système d'équations différentielles déterministes

$$\left\{ \begin{array}{l} \frac{dq_i}{dt} = \frac{p_i}{m_i}, \\ \frac{dp_i}{dt} = -\nabla_{q_i} V - \frac{p_i \xi}{Q}, \\ \frac{d\eta}{dt} = \frac{\xi}{Q}, \\ \frac{d\xi}{dt} = \sum_{i=1}^N \frac{p_i^2}{m_i} - dNk_B T, \end{array} \right.$$

- Invariance de la mesure $d\mu_{\text{NH}} = \exp(dN\eta) dq dp d\eta d\xi$
- $\tilde{H}_{\text{NH}} = \sum_{i=1}^N \frac{p_i^2}{2m_i} + V(q) + \frac{\xi^2}{2Q} + dNk_B T\eta$ conservé

Méthodes déterministes :

Nosé-Hoover (2)

- Dynamique NVE **globale** = préservation d'un invariant = dynamique non-ergodique sur l'espace tout entier! On se restreint donc à la sous-variété

$$\mathcal{M}(E_0) = \left\{ (q, p, \eta, \xi) \in \mathbb{R}^{2dN+2} \mid \tilde{H}_{\text{NH}}(q, p, \eta, \xi) = E_0 \right\}$$

La mesure $d\rho_{\text{NH}} = \frac{d\sigma_{\text{NH}}}{\|\nabla \tilde{H}_{\text{NH}}\|_2}$ est invariante.

- **Hypothèse** : la dynamique est ergodique pour la mesure $d\rho_{\text{NH}}$. Alors la dynamique projetée dans les variables (q, p) est ergodique pour la mesure canonique
- En particulier, il faut que l'énergie soit le **seul** invariant!

Méthodes déterministes :

Nosé-Hoover (3)

- Pas de preuve rigoureuse de convergence
- **Non-convergence** dans certains cas (oscillateur harmonique)
- Généralisations de Nosé-Hoover pour pallier à ces inconvénients : **Nosé-Hoover chain** (on thermostat le thermostat), **Recursive Multiple Thermostats** (on thermostat ensemble le thermostat et le système, dans un vrai formalisme Hamiltonien)
- Numériquement : algorithmes réversibles en temps, symplectiques ou préservant la mesure

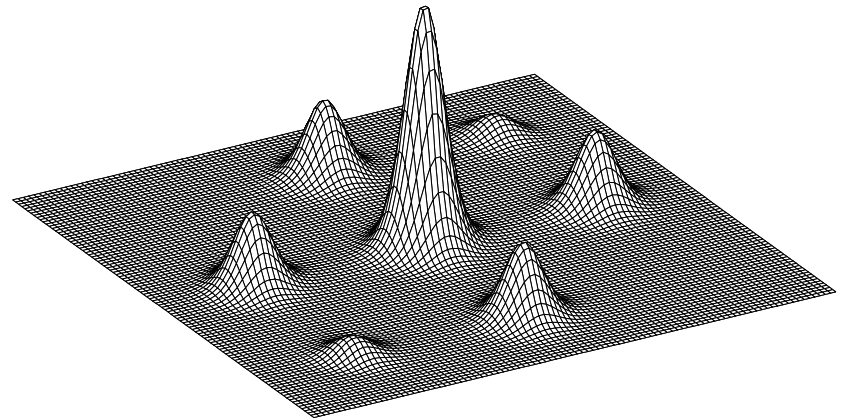
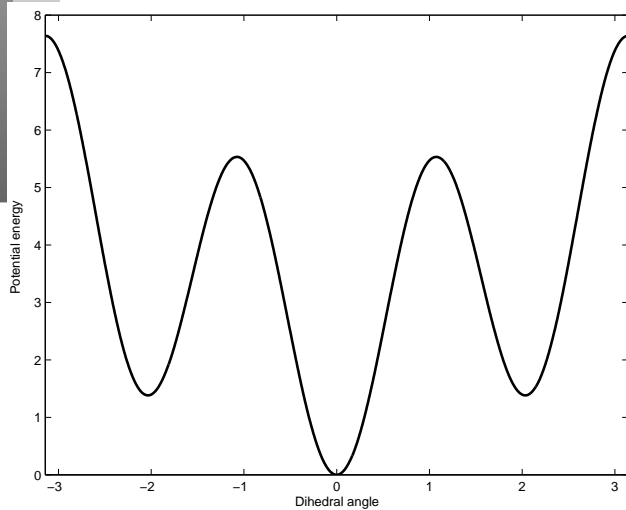
Plan de la présentation (rappel)

- Présentation des différentes méthodes d'échantillonnage (stochastiques, mixtes, déterministes)
- Comparaison numérique sur un cas test
- Une application sous S-TAMP
- Des perspectives

Le cas test numérique

- Modèle d'alcane linéaire $\text{CH}_3-(\text{CH}_2)_n-\text{CH}_3$ (Ryckaert/Belleman), "united-atom"
 - les carbones d'une liaison covalente sont liés par un potentiel harmonique raide
 - les carbones séparés par 3 liaisons ou plus interagissent via un potentiel de Lennard-Jones
 - trois carbones consécutifs $\text{C}_i-\text{C}_{i+1}-\text{C}_{i+2}$ sont liés par un potentiel à 3 corps (angle de valence)
 - quatre atomes consécutifs $\text{C}_i-\text{C}_{i+1}-\text{C}_{i+2}-\text{C}_{i+3}$ interagissent via un potentiel à 4 corps, fonction de l'angle dièdre ϕ_i

Le cas test numérique (2)



Gauche : potentiel de torsion des angles dièdres.

Droite : Distribution de probabilité empirique dans le plan (ϕ_1, ϕ_2) pour une molécule de pentane à 300 K (par Importance sampling avec $N = 10^9$ configurations).

Convergence numérique

- Indicateur de convergence ?
- Usuellement, tests sur la **partie cinétique** (moyenne, premiers moments)
- Ou convergence de certaines **observables physiques** (C_v, \dots)
- Ici, afin de rester en toute généralité, on propose un indicateur comparant **directement** les distributions **empiriques** générées et la distribution **canonique**
- Pour ce faire, on enlève les termes de Lennard-Jones

Convergence numérique (2)

- Mesure projetée sur (par exemple) deux angles dièdres

$$d\nu_{ij}(\phi_i, \phi_j) = Z_\phi^{-2} e^{-\beta V_4(\phi_i)} e^{-\beta V_4(\phi_j)} d\phi_i d\phi_j$$

- **Mesure écart** entre distribution empirique et ν

$$D_n(\{q^m\}) = \sup_{(\phi_i, \phi_j) \in [-\pi, \pi]^2} \left| \frac{1}{n} \sum_{m=0}^{n-1} \mathbf{1}_{\{\phi_i^m \leq \phi_i, \phi_j^m \leq \phi_j\}} - \int_{\{\psi_i \leq \phi_i, \psi_j \leq \phi_j\}} d\nu_{ij}(\psi_i, \psi_j) \right|$$

- Calculé en pratique sur une grille du plan (ϕ_i, ϕ_j)

Convergence numérique (3)

- Les méthodes **purement stochastiques** marchent très bien quand la dimension est petite (cf. bonne fonction de proposition, car système ordonné), sinon elles ne sont plus efficaces!
- Les méthodes **déterministes** ne marchent pas très bien, un peu mieux (en comparaison aux autres) quand la dimension de l'espace augmente
- Les méthodes **mixtes** ont une efficacité à peu près indépendante de la dimension de l'espace, Langevin et HMC étant les meilleurs

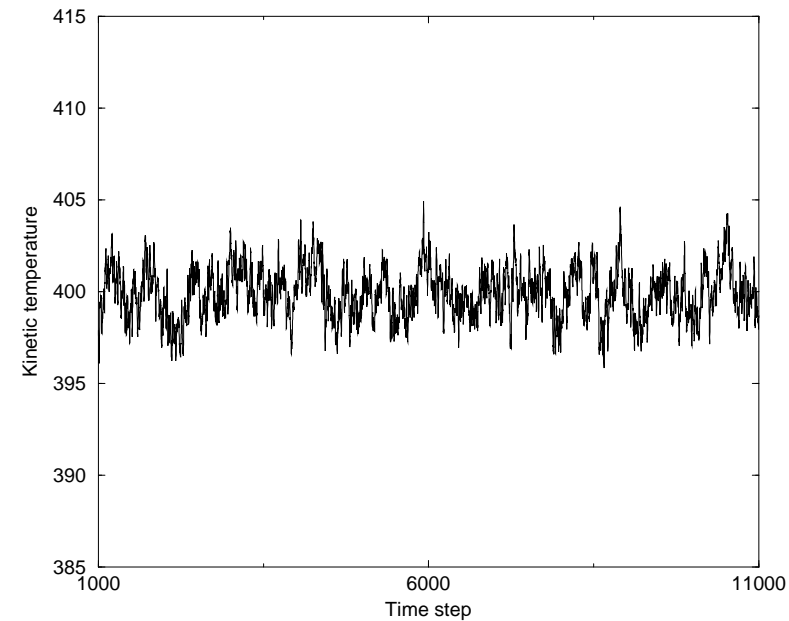
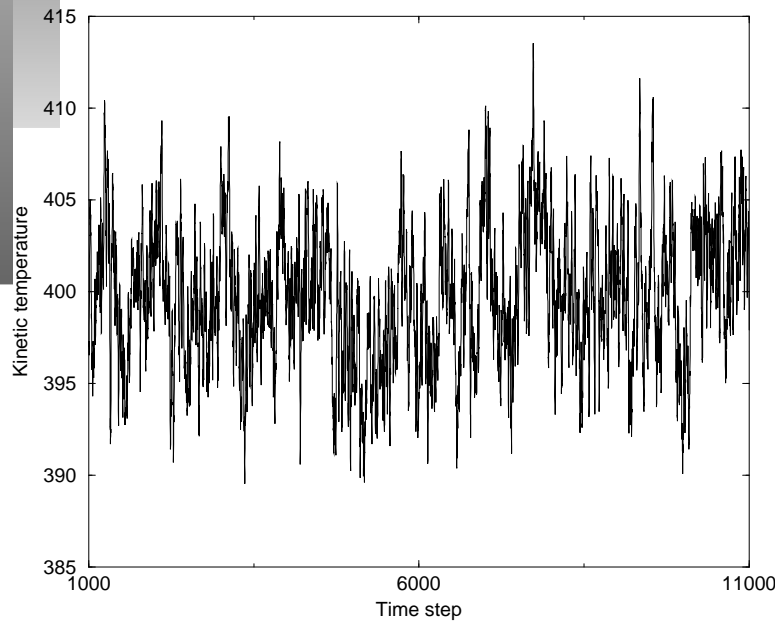
Convergence numérique (4)

- Réaliser plusieurs trajectoires courtes au lieu d'une seule très longue : amélioration modérée, en tout cas pas de dégradation des résultats tant que la longueur des trajectoires reste raisonnable (**parallélisation du pauvre**)
- Partir d'une distribution initiale approximant π (générée par tirage selon une mesure empirique engendrée par importance sampling) accélère la convergence
- Un sous-échantillonnage systématique ne dégrade pas les résultats (intéressant si calcul de propriétés **dynamiques** = peu de points mais bien placés!)

Plan de la présentation (rappel)

- Présentation des différentes méthodes d'échantillonnage (stochastiques, mixtes, déterministes)
- Comparaison numérique sur un cas test
- Une application sous S-TAMP
- Des perspectives

Implémentation dans S-TAMP : un peu de publicité



Evolution typique de la température cinétique au cours du temps, pour un fluide d'argon de densité $\rho = 1400 \text{ kg/m}^3$ à la température cible $T = 400 \text{ K}$. Gauche : $N = 8000$ particules. Droite : $N = 64000$. On note que les fluctuations de T_c diminuent avec N .

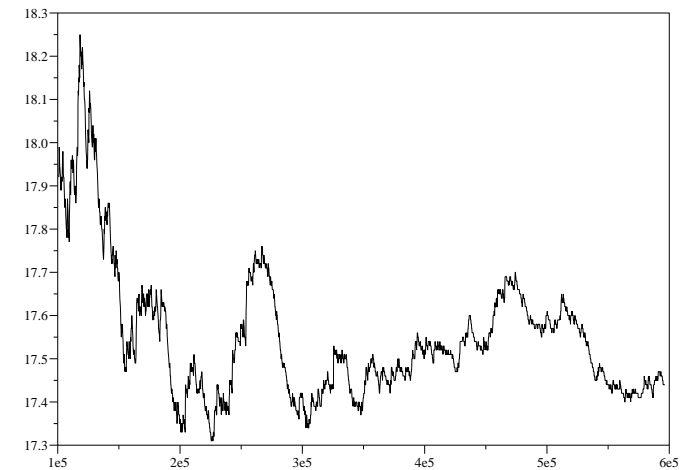
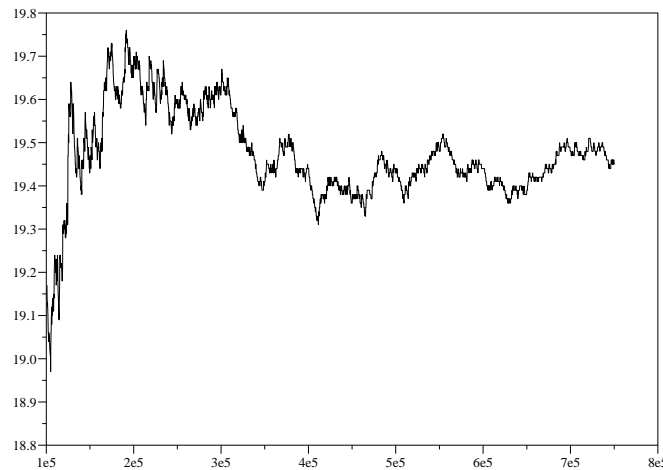
Implémentation dans S-TAMP : un peu de publicité (2)

- Chaleur spécifique à volume constant

$$C_v = \frac{\mathcal{N}_a}{Nk_B T^2} (\langle H^2 \rangle - \langle H \rangle^2)$$

- Quantité difficile à calculer par voie directe (fluctuations) = cas test sévère
- On a souvent recours à des approches alternatives (type dérivée numérique)

Implémentation dans S-TAMP : un peu de publicité (3)



Estimation de C_v en fonction du nombre de pas de temps pour un fluide d'argon de densité $\rho = 1400 \text{ kg/m}^3$ pour 64 000 particules. Gauche : $T = 200 \text{ K}$, maximum 800 000 pas de temps, valeur tabulée $C_v = 19,51 \text{ J/K/mol}$. Droite : $T = 600 \text{ K}$, maximum 600 000 pas de temps, valeur tabulée $C_v = 17,46 \text{ J/K/mol}$. La précision relative obtenue est de 0,01.

- Echantillonnage de type **Blue Moon** intervenant en particulier dans le calcul de différences d'énergie libre
- Pour les cas où la convergence est difficile : amélioration de méthodes de **recuit simulé** (systèmes de particules en interaction)
- **Couplage NVT/NVE**, conditions de bord stochastiques ("échantillonnage" des degrés de libertés manquants)