

Ecole ECODOQUI: résumé du cours

Dynamique moléculaire et calcul d'énergie libre

T. Lelièvre

27 novembre 2008

Résumé

Ces notes portent sur l'étude des méthodes numériques pour le calcul de différences d'énergie libre en dynamique moléculaire [7, 26]. En dynamique moléculaire, les modèles consistent à décrire l'état d'un système par la position (et éventuellement la vitesse) de particules (typiquement les positions des noyaux dans un système moléculaire), qui interagissent au travers d'un potentiel. L'objectif est de calculer des moyennes par rapport à la mesure de Boltzmann-Gibbs associée à ce potentiel (moyennes dans l'ensemble canonique). Mathématiquement, il s'agit d'un problème d'échantillonnage de mesures métastables (ou multi-modales), en très grande dimension. La particularité de la dynamique moléculaire est que, bien souvent, on a quelques informations sur les "directions de métastabilité" au travers de coordonnées de réaction. En utilisant cette donnée, de nombreuses méthodes ont été proposées pour permettre l'échantillonnage de la mesure de Boltzmann-Gibbs. Dans une série de travaux [29, 9, 24, 15], nous avons analysé les méthodes basées sur des équations différentielles stochastiques avec contraintes (dont les solutions vivent sur des sous-variétés définies comme des lignes de niveaux de la coordonnée de réaction). Il s'agit en fait d'analyser des méthodes d'échantillonnage de mesures définies sur des sous-variétés de grande dimension. Plus récemment, nous avons étudié dans [30, 31, 28, 22] des méthodes adaptatives qui ont été proposées pour se débarrasser des métastabilités. Mathématiquement, il s'agit de méthodes d'échantillonnage d'importance, avec une fonction d'importance qui est calculée au cours de la simulation de manière adaptative. Nous avons étudié la vitesse de convergence vers la mesure d'équilibre pour ces méthodes adaptatives, en utilisant des méthodes d'entropie. Nous avons proposé de nouvelles méthodes numériques à la communauté appliquée pour utiliser au mieux ces idées, qui ont été adoptées par des praticiens du domaine [5].

On considère un système moléculaire de N particules, de position $(x_1, \dots, x_N) = x \in \mathbb{R}^{3N}$ et qui interagissent au travers d'un potentiel $V(x_1, \dots, x_N)$. On note $n = 3N$ la dimension totale. Dans l'ensemble canonique NVT, les positions des particules se répartissent suivant la mesure de probabilité de Boltzmann-Gibbs :

$$\mu(dx) = Z^{-1} \exp(-\beta V(x)) dx, \quad (1)$$

où $Z = \int \exp(-\beta V(x)) dx$ est la fonction de partition et $\beta = (kT)^{-1}$, avec k la constante de Boltzmann et T la température. Un des objectifs essentiels en dynamique moléculaire est de calculer des moyennes canoniques de la forme

$$\int \phi(x) \mu(dx) \quad (2)$$

pour certaines fonctionnelles ϕ , afin d'obtenir des "quantités macroscopiques" comme la vraisemblance d'une conformation moléculaire, ou l'avancement d'une réaction chimique.

Les domaines d'application concernent typiquement la biologie moléculaire (calcul de repliement de protéine, simulation de mécanismes moléculaires d'intérêt biologique) et la chimie computationnelle (calcul des processus réactionnels : adsorption sur un substrat, états de transitions, etc...). On renvoie aux monographies [1, 7, 16, 27] pour une présentation générale.

Remarque 1 *On utilise en général des approximations pour le potentiel V , appelées champs de force. Typiquement, V est la somme de potentiels modélisant l'interaction entre deux particules, trois particules et quatre particules :*

$$V = \sum_{i < j} V_1(x_i, x_j) + \sum_{i < j < k} V_2(x_i, x_j, x_k) + \sum_{i < j < k < l} V_3(x_i, x_j, x_k, x_l).$$

Par exemple, $V_1(x_i, x_j) = V_{LJ}(|x_i - x_j|)$ où $V_{LJ}(r) = 4\epsilon \left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right)$ est le potentiel de Lennard-Jones.

Le calcul de moyennes canoniques (2) est un problème en grande dimension (N est couramment de l'ordre de plusieurs centaines de milliers, jusqu'à plusieurs millions pour les applications en biologie) qui nécessite le recours à des méthodes de Monte Carlo, typiquement basées sur des chaînes de Markov (méthodes de type *Markov Chain Monte Carlo*, utilisées également en statistique par exemple).

Par exemple, pour échantillonner μ , on peut utiliser X_t solution de l'équation différentielle stochastique :

$$dX_t = -\nabla V(X_t) dt + \sqrt{2\beta^{-1}} dW_t. \quad (3)$$

La dynamique (3) est appelée, dans ce contexte, dynamique de gradient ou dynamique de Langevin amortie (*overdamped Langevin*). Elle est à la base de toutes les méthodes que nous considérons dans cette section. On rappelle que sous de bonnes hypothèses sur le potentiel V , on a la propriété d'ergodicité : pour μ -presque tout X_0 ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \phi(X_t) dt = \int \phi(x) \mu(dx). \quad (4)$$

Remarque 2 *La dynamique stochastique la plus utilisée par les physiciens est plutôt la dynamique de Langevin :*

$$\begin{cases} dX_t = M^{-1} P_t dt, \\ dP_t = -\nabla V(X_t) dt - \gamma M^{-1} P_t dt + \sqrt{2\gamma\beta^{-1}} dW_t, \end{cases} \quad (5)$$

où M est un tenseur de masse et γ un coefficient de friction. Dans la suite, on se restreint à la dynamique de Langevin amortie (3), qui est obtenue à partir de la dynamique de Langevin dans la limite $m \rightarrow 0$ ou $\gamma \rightarrow \infty$. La généralisation des résultats ci-dessous à la dynamique de Langevin n'est pas en général triviale car le générateur infinitésimal associé à cette dynamique n'est plus elliptique, ce qui rend l'analyse plus compliquée.

Ce qui rend en pratique la méthode de calcul (4) inopérante est la métastabilité du processus X_t . En effet, la convergence $T \rightarrow \infty$ est extrêmement longue à obtenir, du fait de puits dans le potentiel V dans lesquels le processus X_t reste piégé très longtemps. Pour donner une image intuitive on peut considérer le potentiel 2-dimensionnel représenté sur la Figure 1, pour lequel la première coordonnée du processus X_t^1 est une variable métastable (ou variable lente).

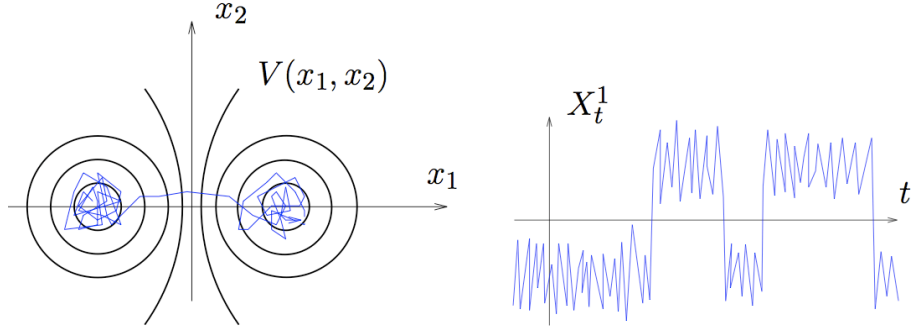


FIG. 1 – À gauche, les lignes de niveau d’un potentiel métastable, et à droite, une représentation schématique de l’évolution de la première coordonnée d’un processus stochastique évoluant selon la dynamique (3) dans ce potentiel.

Bien sûr, résoudre ce problème dans toute sa généralité est une tâche impossible (aucune méthode numérique ne peut explorer l’ensemble des puits d’un potentiel V en grande dimension, séparés par des barrières arbitrairement grandes, en un temps de calcul raisonnable). Nous allons donc ajouter une hypothèse supplémentaire, spécifique au cadre de la simulation moléculaire. Nous supposons dans la suite que l’on connaît les variables lentes du système, sous la forme d’une application

$$\xi : \mathbb{R}^n \rightarrow \mathbb{R} \text{ ou } \xi : \mathbb{R}^n \rightarrow \mathbb{T} \quad (6)$$

appelée coordonnée de réaction, ou variable collective. On note $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ le tore unidimensionnel. En pratique, on peut avoir $\xi : \mathbb{R}^n \rightarrow \mathbb{R}^p$ (ou $\xi : \mathbb{R}^n \rightarrow \mathbb{T}^p$) avec $p \ll n$, mais on suppose dans cette présentation $p = 1$ pour simplifier. L’ensemble des résultats sont généralisables au cas $p > 1$. Pour fixer les idées, sur l’exemple de la figure 1, la coordonnée de réaction serait $\xi : (x_1, x_2) \mapsto x_1$. En pratique, ξ est typiquement une distance signée entre deux (groupes de) particules (auquel cas ξ est à valeurs dans \mathbb{R}), par exemple pour indexer l’évolution d’une réaction de synthétisation d’un composé, ou un angle (auquel cas ξ est à valeurs dans \mathbb{T}), par exemple pour caractériser la conformation d’une molécule. La connaissance de ξ nécessite bien sûr une intuition sur le système considéré. La qualité des méthodes que nous présentons ci-dessous dépend crucialement du choix d’une bonne fonction ξ , et déterminer pour un système quelconque une bonne coordonnée de réaction reste un problème compliqué. Dans la suite, on suppose que ξ est une fonction régulière telle que $|\nabla\xi| > 0$.

À la coordonnée de réaction est associée un potentiel effectif, appelé énergie libre.

Définition 3 *L’image de la mesure μ par ξ est $Z^{-1} \exp(-\beta A(z)) dz$, où A est appelée l’énergie libre. On vérifie que*

$$A(z) = -\beta^{-1} \ln \left(\int_{\Sigma_z} \exp(-\beta V) |\nabla\xi|^{-1} d\sigma_{\Sigma_z} \right) = -\beta^{-1} \ln Z_{\Sigma_z}, \quad (7)$$

où $\Sigma_z = \{x, \xi(x) = z\}$ est une sous-variété (régulière) de \mathbb{R}^n , et σ_{Σ_z} est la mesure de Lebesgue sur Σ_z .

La formule (7) découle de la formule de la co-aire [2, 14], que l’on peut voir comme une généralisation de la formule de Fubini (intégration par tranches) : pour toute fonction régulière $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^n} \phi(x) |\nabla\xi|(x) dx = \int_{\mathbb{R}} \int_{\Sigma_z} \phi d\sigma_{\Sigma_z} dz. \quad (8)$$

La mesure $|\nabla\xi|^{-1}(x)\sigma_{\Sigma_z}(dx)$ est parfois notée $\delta_{\xi(x)-z}(dx)$ dans la littérature.

En utilisant la formule de la co-aire, on voit que l'on peut écrire une moyenne canonique sous la forme (c'est une formule de conditionnement) :

$$\int \phi d\mu = \int_{\mathbb{R}} \int_{\Sigma_z} \phi d\mu_{\Sigma_z} \exp(-\beta A(z)) dz \quad (9)$$

où la mesure de probabilité

$$d\mu_{\Sigma_z} = \frac{\exp(-\beta V)|\nabla\xi|^{-1}d\sigma_{\Sigma_z}}{Z_{\Sigma_z}} = \frac{\exp(-\beta V)|\nabla\xi|^{-1}d\sigma_{\Sigma_z}}{\exp(-\beta A(z))} \quad (10)$$

est la mesure canonique μ conditionnée à une valeur fixée de la coordonnée de réaction. Noter que $Z_{\Sigma_z} = \exp(-\beta A(z))$ est la fonction de partition associée à la mesure conditionnée μ_{Σ_z} .

On voit donc que le problème initial (calcul d'une moyenne canonique) est ramené aux problèmes suivants :

- (i) Calculer l'énergie libre A (à une constante additive près qui n'intervient plus après normalisation, cf. (9)–(10)) ;
- (ii) Échantillonner la mesure μ_{Σ_z} . Si ξ a été bien choisi, on peut espérer que cette tâche soit nettement plus simple que l'échantillonnage de la mesure μ (penser à nouveau au cas de la figure 1 pour lequel, dans les tranches Σ_z , le potentiel est convexe, donc sans métastabilité).

Dans plusieurs applications, la fonctionnelle ϕ d'intérêt est en fait une fonction de ξ , auquel cas le seul problème restant est le calcul de l'énergie libre (i). Par ailleurs, le calcul de l'énergie libre présente un intérêt en soi, car les physiciens utilisent le profil d'énergie libre pour guider leur intuition sur les mécanismes réactionnels (y compris sur des aspects dynamiques).

Remarque 4 *Il y a ici une question sous-jacente très importante pour la pratique : peut-on donner un sens dynamique à l'énergie libre A ? De manière plus précise, est-il possible d'écrire une dynamique effective sur la coordonnée de réaction $\xi(X_t)$ (où X_t satisfait la dynamique (3)) en fonction de A ? Cette question est l'objet d'un travail en cours [25].*

Dans la suite, on s'intéresse à trois grandes classes de méthodes pour le calcul de l'énergie libre A (problème (i)) : l'intégration thermodynamique (section 1), les méthodes hors équilibre (section 2), et les méthodes adaptatives (sections 3, 4 et 5). Pour une revue, on renvoie à [26] et au livre [7]. On discutera dans la section 1 l'échantillonnage d'une mesure sur une sous-variété (problème (ii)).

1 Intégration thermodynamique

L'intégration thermodynamique pour le calcul de l'énergie libre [23] est basée sur deux ingrédients.

Le premier ingrédient est que la dérivée de A par rapport à la coordonnée de réaction (appelée la force moyenne) peut être obtenue en échantillonnant la probabilité conditionnelle μ_{Σ_z} . En effet, en utilisant la formule de la co-aire, on montre que

$$\begin{aligned} A'(z) &= Z_{\Sigma_z}^{-1} \int_{\Sigma_z} \left(\frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \operatorname{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right) \right) \exp(-\beta V) |\nabla \xi|^{-1} d\sigma_{\Sigma_z} \quad (11) \\ &= Z_{\Sigma_z}^{-1} \int_{\Sigma_z} \frac{\nabla \xi}{|\nabla \xi|^2} \cdot \left(\nabla \tilde{V} + \beta^{-1} H \right) \exp(-\beta \tilde{V}) d\sigma_{\Sigma_z} \\ &= \int_{\Sigma_z} F d\mu_{\Sigma_z}, \end{aligned}$$

où

$$\tilde{V} = V + \beta^{-1} \ln |\nabla \xi|, \quad (12)$$

la force locale F est définie par

$$F = \frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \operatorname{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right) \quad (13)$$

et le vecteur courbure moyenne H est :

$$H = -\nabla \cdot \left(\frac{\nabla \xi}{|\nabla \xi|} \right) \frac{\nabla \xi}{|\nabla \xi|}. \quad (14)$$

Le deuxième ingrédient est que la dynamique contrainte :

$$\begin{cases} dX_t = -\nabla \tilde{V}(X_t) dt + \sqrt{2\beta^{-1}} dW_t + \nabla \xi(X_t) d\Lambda_t, \\ d\Lambda_t \text{ tel que } \xi(X_t) = z, \end{cases} \quad (15)$$

admet pour mesure invariante (et, sous de bonnes hypothèses, est ergodique par rapport à) la probabilité conditionnelle $d\mu_{\Sigma_z} = Z_{\Sigma_z}^{-1} \exp(-\beta \tilde{V}) d\sigma_{\Sigma_z}$. Noter qu'il faut bien utiliser le potentiel modifié \tilde{V} (défini par (12)) et non pas le potentiel original V . Plus précisément, l'équation différentielle stochastique contrainte (15) s'écrit :

$$dX_t = P(X_t) \left(-\nabla \tilde{V}(X_t) dt + \sqrt{2\beta^{-1}} dW_t \right) + \beta^{-1} H(X_t) dt, \quad (16)$$

où l'opérateur de projection orthogonale $P(x)$ est :

$$P(x) = \operatorname{Id} - n(x) \otimes n(x), \quad (17)$$

où $n(x) = \frac{\nabla \xi}{|\nabla \xi|}(x)$ est un vecteur normal à la sous-variété $\Sigma_{\xi(x)}$ au point x . Cette dynamique peut aussi s'écrire en utilisant le produit de Stratonovitch :

$$dX_t = -P(X_t) \nabla \tilde{V}(X_t) dt + \sqrt{2\beta^{-1}} P(X_t) \circ dW_t. \quad (18)$$

La formulation (15) avec un multiplicateur de Lagrange $d\Lambda_t$ associé à la contrainte $\xi(X_t) = z$ présente l'avantage de suggérer des schémas de discrétisation (cf. le schéma (21) ci-dessous).

De plus, on vérifie que

$$d\Lambda_t = d\Lambda_t^m + d\Lambda_t^f, \quad (19)$$

avec $d\Lambda_t^m = -\sqrt{2\beta^{-1}} \frac{\nabla \xi}{|\nabla \xi|^2}(X_t) \cdot dW_t$ et

$$d\Lambda_t^f = \frac{\nabla \xi}{|\nabla \xi|^2} \cdot \left(\nabla \tilde{V} + \beta^{-1} H \right) (X_t) dt = F(X_t) dt \quad (20)$$

si bien que

$$A'(z) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T d\Lambda_t = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T d\Lambda_t^f.$$

La valeur de la force moyenne en un point z et donc en pratique obtenue par une moyenne sur une trajectoire du multiplicateur de Lagrange $d\Lambda_t$ associé à la contrainte $\xi(x) = z$ pour la dynamique contrainte (15) (cf. l'équation (22) ci-dessous).

Le profil d'énergie libre est ensuite obtenu par une intégration en dimension 1 :

$$A(z) - A(0) = \int_0^z A'(z) dz \simeq \sum_{i=0}^K \omega_i A'(z_i),$$

où z_i désigne des points de Gauss d'une formule d'intégration, et ω_i les poids associés. Des deux problèmes (i) et (ii) mentionnés en introduction, on se ramène donc simplement au problème (ii) : l'échantillonnage d'une mesure sur une sous-variété. Remarquer que le fait que la partie à variation finie du processus Λ_t s'exprime en fonction de la force moyenne locale F (cf. (20)) permet de calculer $A'(z)$ sans calculer les dérivées compliquées dans l'expression (13) définissant F .

Dans [9], nous démontrons rigoureusement l'ensemble de ces résultats, qui étaient pour la plupart connus des physiciens [32, 34, 13, 8]. Nous proposons également une méthode de réduction de variance pour le schéma numérique associé. Précisons cela. Un schéma de discrétisation naturel pour (15) est le suivant :

$$\begin{cases} X_{n+1} = X_n - \nabla \tilde{V}(X_n)\delta t + \sqrt{2\beta^{-1}}\Delta W_n + \lambda_n \nabla \xi(X_n), \\ \text{où } \lambda_n \in \mathbb{R} \text{ est tel que } \xi(X_{n+1}) = z, \end{cases} \quad (21)$$

où $\Delta W_n = W_{(n+1)\delta t} - W_{n\delta t}$ désigne un incrément brownien. Il s'agit d'un schéma de type prédicteur correcteur, où l'on fait d'abord un pas du schéma d'Euler sur l'équation différentielle stochastique non contrainte (3), puis un pas de projection sur la sous-variété Σ_z . On montre que :

$$\lim_{T \rightarrow \infty} \lim_{\delta t \rightarrow 0} \frac{1}{T} \sum_{n=1}^{T/\delta t} \lambda_n = A'(z). \quad (22)$$

On peut réduire grandement la variance de l'estimateur $\frac{1}{T} \sum_{n=1}^{T/\delta t} \lambda_n$ en remarquant que, conformément à la formule (19), cet estimateur contient une partie fluctuante de moyenne nulle correspondant à la discrétisation de la martingale Λ_t^m . Une méthode de réduction de variance consiste alors à utiliser l'estimateur $\frac{1}{T} \sum_{n=1}^{T/\delta t} \frac{\lambda_n + \tilde{\lambda}_n}{2}$, où $\tilde{\lambda}_n$ est défini par (remarquer le changement de signe devant l'incrément brownien par rapport à (21)) :

$$\begin{cases} \tilde{X}_{n+1} = X_n - \nabla \tilde{V}(X_n)\delta t - \sqrt{2\beta^{-1}}\Delta W_n + \tilde{\lambda}_n \nabla \xi(X_n), \\ \text{où } \tilde{\lambda}_n \in \mathbb{R} \text{ est tel que } \xi(\tilde{X}_{n+1}) = z. \end{cases}$$

Remarque 5 *Il y a bien sûr plusieurs manières de projeter la dynamique (3) sur la sous-variété Σ_z , et on a choisi une méthode de projection qui se discrétise de manière très naturelle. Il est important de remarquer que la statistique (i.e. la mesure invariante) de la dynamique projetée dépend de la méthode de projection. Par exemple, une projection utilisant le potentiel de pénalisation $V_k(x, z) = V(x) + k(z - \xi(x))^2$ (dans la dynamique (3)) permet d'échantillonner la mesure invariante μ_{Σ_z} , dans la limite $k \rightarrow \infty$, sans modification du potentiel V en \tilde{V} . Dans le cas d'une équation différentielle ordinaire (sans mouvement brownien), les deux approches (projection "rigide" (16) et pénalisation) sont équivalentes. Ceci est relié à une discussion dans la littérature physique sur les différences statistiques entre l'utilisation d'une barre rigide et l'utilisation d'un ressort infiniment raide pour imposer une contrainte (cf. [37, 19]).*

Remarque 6 *On pourrait penser à utiliser un algorithme de projection différent dans l'algorithme (21). Par exemple, on pourrait considérer (noter le changement de $\nabla \xi(X_n)$ à $\nabla \xi(X_{n+1})$ par rapport à (21))*

$$\begin{cases} X_{n+1} = X_n - \nabla \tilde{V}(X_n)\delta t + \sqrt{2\beta^{-1}}\Delta W_n + \lambda_n \nabla \xi(X_{n+1}), \\ \text{où } \lambda_n \in \mathbb{R} \text{ est tel que } \xi(X_{n+1}) = z, \end{cases} \quad (23)$$

qui est l'équation d'Euler Lagrange associée à la discrétisation :

$$\begin{cases} \tilde{X}_{n+1} = X_n - \nabla \tilde{V}(X_n)\delta t + \sqrt{2\beta^{-1}}\Delta W_n, \\ X_{n+1} = \arg \min_{Y \in \Sigma_z} \|\tilde{X}_{n+1} - Y\|^2. \end{cases}$$

On montre dans [9, 24] que cette méthode de discrétisation, de même que la discrétisation (21) sont consistantes (dans la limite $\delta t \rightarrow 0$) avec la dynamique projetée (15). Ceci dit, nous montrons dans [24], qu'une discrétisation du type (23) n'est pas nécessairement consistante avec la dynamique projetée associée pour une équation différentielle stochastique générale (avec un terme de diffusion devant le mouvement brownien qui n'est pas une constante fois l'identité). Autrement dit, les opérations de projection de la dynamique et de discrétisation ne commutent pas en général. On renvoie à [24] pour plus de détails.

Nous terminons cette section en évoquant un résultat de l'article [15]. Une question naturelle concernant le schéma de discrétisation (21) est de savoir quelle est l'erreur induite par la discrétisation en temps sur la mesure invariante de la dynamique. Autrement dit, quel est l'écart entre la mesure μ_{Σ_z} invariante par la dynamique (15) et la mesure $\mu_{\Sigma_z}^{\delta t}$ invariante par la dynamique (21), en fonction de δt . Il s'agit donc d'étendre des résultats bien connus de vitesse de convergence faible pour la discrétisation par le schéma d'Euler des équations différentielles stochastiques [36, 35], à des équations différentielles stochastiques projetées. Dans [15], nous montrons, dans un cadre plus général, que l'erreur faible est d'ordre 1 : il existe une constante positive δt_0 telles que, pour tout pas de temps $\delta t < \delta t_0$, pour toute fonction régulière $g : \Sigma_z \rightarrow \mathbb{R}$, il existe $C > 0$ tel que

$$\left| \int_{\Sigma_z} g d\mu_{\Sigma_z}^{\delta t} - \int_{\Sigma_z} g d\mu_{\Sigma_z} \right| \leq C \delta t.$$

La preuve s'applique à une équation différentielle stochastique générale ergodique pour une mesure μ_{Σ} de support une sous-variété Σ de \mathbb{R}^n , et pour tout schéma numérique consistant tel que le processus discrétisé est ergodique par rapport à une mesure $\mu_{\Sigma}^{\delta t}$ portée par Σ . La démonstration repose sur l'utilisation d'un système de coordonnées orthogonales adapté à la sous-variété Σ . On renvoie à [15] pour des résultats détaillés appliqués à divers schémas numériques.

2 Méthodes hors équilibre

Les méthodes hors équilibre pour calculer l'énergie libre ont été introduites récemment [21]. Le principe est de forcer la coordonnée de réaction à bouger selon une évolution déterministe $z : [0, T] \mapsto \mathbb{R}$ prescrite *a priori* pour évaluer la différence d'énergie libre $(A(z(t)) - A(z(0)))_{0 \leq t \leq T}$. On parle de dynamique hors équilibre, car on force la coordonnée de réaction à avancer en temps (en imposant la contrainte $\xi(X_t) = z(t)$) à une vitesse bien supérieure à celle à laquelle elle évoluerait pour la dynamique de base (3). La question est alors la suivante : comment établir des relations entre des quantités thermodynamiques à l'équilibre (comme la différence d'énergie libre) et des quantités calculées hors équilibre ? Dans la littérature de physique, de tels liens ont été découverts par C. Jarzynski dans [21]. D'un point de vue mathématique, il s'agit de formules de Feynman-Kac. Détaillons cela.

Sur le modèle de (15), on introduit la dynamique hors équilibre suivante :

$$\begin{cases} dX_t = -\nabla \tilde{V}(X_t) dt + \sqrt{2\beta^{-1}} dW_t + \nabla \xi(X_t) d\Lambda_t, \\ d\Lambda_t \text{ tel que } \xi(X_t) = z(t). \end{cases} \quad (24)$$

Plus précisément, l'équation différentielle stochastique (24) s'écrit (comparer avec (18)) :

$$\begin{cases} dX_t = -P(X_t) \nabla \tilde{V}(X_t) dt + \sqrt{2\beta^{-1}} P(X_t) \circ dW_t + \nabla \xi(X_t) d\Lambda_t^{\text{ext}}, \\ d\Lambda_t^{\text{ext}} = \frac{z'(t)}{|\nabla \xi(X_t)|^2} dt, \end{cases} \quad (25)$$

où \circ désigne un produit de Stratonovitch. Comme dans le cas de la dynamique (15), on vérifie que

$$d\Lambda_t = d\Lambda_t^m + d\Lambda_t^f + d\Lambda_t^{\text{ext}},$$

avec $d\Lambda_t^m = -\sqrt{2\beta^{-1}} \frac{\nabla \xi}{|\nabla \xi|^2}(X_t) \cdot dW_t$, $d\Lambda_t^f = F(X_t) dt$ et $d\Lambda_t^{\text{ext}}$ défini dans (25).

En utilisant une formule de Feynman-Kac, on montre dans [29] l'égalité suivante :

$$A(z(t)) - A(z(0)) = -\beta^{-1} \ln(\mathbb{E}(\exp(-\beta \mathcal{W}(t)))) ,$$

où $\mathcal{W}(t)$ est un poids (un travail ou une énergie) associé à la trajectoire $(X_s)_{0 \leq s \leq t}$:

$$\mathcal{W}(t) = \int_0^t F(X_s) z'(s) ds = \int_0^t z'(s) d\Lambda_s^f.$$

A partir de cette formule, nous proposons un estimateur de $A(z(t)) - A(z(0))$ de la forme :

$$A(z(t)) - A(z(0)) \simeq -\beta^{-1} \ln \left(\frac{1}{M} \sum_{m=1}^M \exp \left(-\beta \mathcal{W}_{T/\delta t}^m \right) \right), \quad (26)$$

où $(\mathcal{W}_n^m)_{0 \leq n \leq T/\delta t}$ est le poids d'une m -ième réplique d'une trajectoire (discrétisée) de (24) (et $1 \leq m \leq M$, M désignant le nombre total de répliques). Plus précisément, pour m fixé, $(\mathcal{W}_n^m)_{0 \leq n \leq T/\delta t}$ est défini par $\mathcal{W}_0^m = 0$ et, $\forall n \in \{0, \dots, T/\delta t - 1\}$

$$\mathcal{W}_{n+1}^m = \mathcal{W}_n^m + \frac{z(t_{n+1}) - z(t_n)}{t_{n+1} - t_n} \lambda_n^{f,m}$$

avec $t_n = n\delta t$,

$$\lambda_n^{f,m} = \lambda_n^m - \frac{z(t_{n+1}) - z(t_n)}{|\nabla \xi(X_n^m)|^2} + \sqrt{2\beta^{-1}} \frac{\nabla \xi}{|\nabla \xi|^2}(X_n^m) \cdot \Delta W_n^m,$$

et (sur le modèle du schéma de discrétisation (21))

$$\begin{cases} X_{n+1}^m = X_n^m - \nabla V(X_n^m) \delta t + \sqrt{2\beta^{-1}} \Delta W_n^m + \lambda_n^m \nabla \xi(X_n^m), \\ \text{où } \lambda_n^m \in \mathbb{R} \text{ est tel que } \xi(X_{n+1}^m) = z(t_{n+1}). \end{cases}$$

Dans ces équations, $\Delta W_n^m = W_{(n+1)\delta t}^m - W_{n\delta t}^m$, où $(W_t^m)_{t \geq 0, 1 \leq m \leq M}$ désigne M mouvements Browniens indépendants. On renvoie à [29] pour des illustrations numériques de cette approche.

On peut signaler qu'en pratique, l'efficacité de ces méthodes hors équilibre pour le calcul de l'énergie libre n'a pas été vraiment démontrée. Il apparaît en effet que la variance de l'estimateur (26) explose si la variation en temps de l'application $(z(t))_{0 \leq t \leq T}$ est trop rapide. On est donc amené à utiliser des variations assez lente, ce qui se rapproche alors d'une méthode d'intégration thermodynamique (cf. section 1). Mentionnons tout de même le travail [33], où les auteurs proposent un algorithme comprenant une étape de reconfiguration stochastique pour mieux contrôler la variance de l'estimateur.

3 Méthodes adaptatives : algorithmes

Nous abordons dans cette section les méthodes qui ont été proposées le plus récemment pour calculer les différences d'énergie libre. Il s'agit des méthodes adaptatives, du type Wang-Landau [38], *Adaptive Biasing Force* [10, 18], ou *metadynamics* [20], par exemple. Dans [30], nous proposons un cadre unifié pour toutes ces méthodes, que nous résumons maintenant.

Le principe des méthodes adaptatives est le suivant. Il s'agit de modifier le potentiel vu par le processus X_t en fonction des configurations visitées, avec deux objectifs :

- (i) explorer de manière efficace l'espace des configurations,
- (ii) calculer le profil d'énergie libre.

Le potentiel modifié dépend du temps et est de la forme, pour tout $x \in \mathbb{R}^n$,

$$\mathcal{V}_t(x) = V(x) - A_t(\xi(x)),$$

où A_t est une approximation de l'énergie libre à l'instant t , au vu des configurations visitées par le système jusqu'à cet instant. On cherche à faire évoluer A_t de manière à ce que $\lim_{t \rightarrow \infty} A_t = A$ (cf. objectif (ii) ci dessus).

Pour comprendre comment mettre à jour A_t , faisons l'hypothèse un instant que le processus au temps t est immédiatement à l'équilibre par rapport au potentiel \mathcal{V}_t , *i.e.*, que

$$\psi(t, \cdot) = \psi^{\text{eq}}(t, \cdot) \propto \exp(-\beta \mathcal{V}_t)$$

où ψ désigne la densité de probabilité de la variable aléatoire X_t . On peut alors calculer une énergie libre observée, en remplaçant V par \mathcal{V}_t dans la formule (7) :

$$A_t^{\text{obs}}(z) = -\beta^{-1} \ln \left(\int_{\Sigma_z} \exp(-\beta \mathcal{V}_t) |\nabla \xi|^{-1} d\sigma_{\Sigma_z} \right)$$

et une force moyenne observée, en remplaçant V par \mathcal{V}_t dans la formule (11) :

$$\begin{aligned} (A_t^{\text{obs}})'(z) &= \frac{\int_{\Sigma_z} \left(\frac{\nabla \mathcal{V}_t \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \text{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right) \right) \exp(-\beta \mathcal{V}_t) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \exp(-\beta \mathcal{V}_t) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}, \\ &= \frac{\int_{\Sigma_z} F \exp(-\beta \mathcal{V}_t) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \exp(-\beta \mathcal{V}_t) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}} - A_t'(z), \end{aligned}$$

où F désigne la force moyenne locale (13). Ici et dans toute la suite, on note par $'$ une dérivée par rapport à la variable z (la valeur de la coordonnée de réaction). On vérifie facilement que

$$A_t^{\text{obs}} = A - A_t \text{ et } (A_t^{\text{obs}})' = A' - A_t'. \quad (27)$$

Au vu de (27) et pour que $\lim_{t \rightarrow \infty} A_t = A$, il semble naturel de proposer les deux types de mise à jour suivantes (où l'on ne suppose plus cette fois le processus à l'équilibre : $\psi(t, \cdot)$ désigne la densité de probabilité de X_t) :

- Mise à jour du potentiel A_t (méthode de type *Adaptive Biasing Potential* (ABP))

$$\frac{\partial A_t}{\partial t}(z) = -\frac{1}{\tau} \beta^{-1} \ln \int \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}, \quad (28)$$

où τ désigne un paramètre numérique homogène à un temps ;

- Ou bien mise à jour de la force moyenne A_t' (méthode de type *Adaptive Biasing Force* (ABF))

$$\frac{\partial A_t'}{\partial t}(z) = \frac{1}{\tau} \left(\frac{\int F \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}} - A_t'(z) \right). \quad (29)$$

En effet, si on suppose dans (28) que $\psi = \psi^{\text{eq}}$, on obtient l'évolution suivante pour A_t : $\frac{\partial A_t}{\partial t} = \frac{1}{\tau} (A - A_t)$ et donc A_t converge bien vers A . Un raisonnement similaire tient pour (29). Si le processus était à l'équilibre à chaque instant ($\psi = \psi^{\text{eq}}$), on obtiendrait donc l'énergie

libre (ou sa dérivée, la force moyenne) en temps long. On espère que cela sera encore le cas sans cette hypothèse d'équilibre instantané. C'est l'objet de la section 4.

Il est important de noter que les deux approches (ABP et ABF) sont bien différentes, au sens où la dérivée de l'équation (28) par rapport à z n'est pas l'équation (29) (ce serait le cas si $\psi = \psi^{\text{eq}}$).

Remarquer également que la méthode est construite de telle sorte que si un état stationnaire est atteint, *i.e.* si $(\psi(t, \cdot), A_t)$ converge en temps long vers un couple (ψ_∞, A_∞) , alors, nécessairement, $A_\infty = A$ (à une constante additive près). Il suffit pour s'en rendre compte d'écrire les formules (28) et (29) avec $(\psi(t, \cdot), A_t) = (\psi_\infty, A_\infty)$.

Pour le moment, nous n'avons pas précisé la dynamique sous-jacente pour X_t : notre discussion est valable aussi bien pour une dynamique de Langevin (5), que pour une dynamique de gradient (3). Pour fixer les idées, écrivons le système complet pour une dynamique adaptative de type ABF (29) (en prenant formellement $\tau = \infty$: mise à jour instantanée de A'_t), construite à partir d'une dynamique de gradient (3) :

$$\begin{cases} dX_t = -\nabla(V - A_t \circ \xi)(X_t) dt + \sqrt{2\beta^{-1}} dW_t, \\ A'_t(z) = \mathbb{E}(F(X_t) | \xi(X_t) = z), \end{cases} \quad (30)$$

où \circ désigne ici l'opérateur de composition et $\mathbb{E}(F(X_t) | \xi(X_t) = z)$ désigne une espérance conditionnelle. En terme de la densité $\psi(t, \cdot)$ de X_t , on a l'équation de Fokker-Planck associée :

$$\begin{cases} \partial_t \psi = \text{div}(\nabla(V - A_t \circ \xi)\psi + \beta^{-1} \nabla \psi), \\ A'_t(z) = \frac{\int_{\Sigma_z} F \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}. \end{cases} \quad (31)$$

On vérifie que les équations (30) et (31) sont bien équivalentes, car la mesure conditionnée de la loi de probabilité $\psi(x) dx$ sachant que $\xi(x) = z$ est donnée par $\frac{\psi |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int \psi |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}$ (ceci découle de la formule de la co-aire (8)).

Cette dynamique est connue des physiciens [10, 18] sous le nom de *Adaptive Biasing Force method*, sous une forme "proche" où on approxime l'espérance conditionnelle par une moyenne ergodique (cf. la section suivante).

L'intérêt du cadre que nous proposons pour les algorithmes adaptatifs est qu'il permet de comprendre un grand nombre de méthodes adaptatives dans un formalisme similaire, et d'imaginer de nouvelles méthodes, notamment en jouant sur la manière de mettre à jour A_t et A'_t (on a ici donné la méthode la plus simple, pour simplifier l'exposé). On renvoie à [30] pour plus de détails. De plus, ce cadre permet d'étudier la (vitesse de) convergence de certains algorithmes adaptatifs (cf. section 4) et de proposer de nouvelles méthodes de discrétisation (cf. section 5).

Remarque 7 *D'un point de vue probabiliste, les méthodes adaptatives que nous avons décrites peuvent être vues comme des méthodes de réduction de variance par fonction d'importance, avec une fonction d'importance adaptative. En effet, en partant du principe que pour un bon choix de ξ , la dynamique (30) converge plus vite vers l'équilibre que la dynamique (3) (ce que l'on va démontrer dans la section suivante), on dispose d'une méthode qui permet d'échantillonner des mesures métastables. Cette remarque donne à ces méthodes une portée assez générique, du moins pour les applications dans lesquelles on peut intuitiver une bonne "direction de métastabilité". On donne cette "direction", et l'algorithme calcule automatiquement une bonne fonction d'importance "dans cette direction".*

4 Méthodes adaptatives : étude théorique

L'objectif de cette section théorique est de prouver, pour la dynamique (30), que A'_t converge vers A' (objectif (ii) de la section précédente), et de déterminer à quelle vitesse la convergence a lieu (objectif (i) de la section précédente). Autrement dit, on cherche à déterminer en quoi est-ce que la dynamique (30) converge plus vite vers un état stationnaire que la dynamique de gradient (3) dont on a dit qu'elle était inutilisable en pratique car métastable. Cette section s'appuie sur l'article [31].

Pour montrer que la dynamique (30) va plus vite vers un état stationnaire que (3), il faut décider d'une mesure de la vitesse de convergence. Il y a plusieurs manières de quantifier un comportement métastable, comme par exemple le temps de sortie des puits de potentiel, la variance asymptotique des estimateurs considérés, le temps de "décorrélation" du processus X_t ou la vitesse de convergence de la loi de X_t vers μ . Dans la suite, c'est ce quatrième critère que l'on choisit.

Commençons par rappeler une analyse classique de la vitesse de convergence vers l'équilibre pour (3). On rappelle qu'associée à la dynamique (3), on peut écrire une équation de Fokker-Planck (linéaire) sur la densité ϕ de X_t solution de (3) :

$$\partial_t \phi = \operatorname{div}(\nabla V \phi + \beta^{-1} \nabla \phi). \quad (32)$$

Une approche classique pour étudier le comportement en temps long de cette équation consiste à introduire l'entropie de ϕ par rapport à la solution stationnaire $\phi_\infty = Z^{-1} \exp(-\beta V)$ de (32) :

$$H(\phi|\phi_\infty) = \int_{\mathbb{R}^d} \ln \left(\frac{\phi}{\phi_\infty} \right) \phi.$$

La fonction ϕ_∞ est la densité de la mesure μ . Remarquer que $H \geq 0$ et que $H = 0$ si et seulement si $\phi = \phi_\infty$. On rappelle l'inégalité (de Csiszar-Kullback)

$$\int_{\mathbb{R}^d} |\phi - \phi_\infty| \leq \sqrt{2H(\phi|\phi_\infty)}. \quad (33)$$

Contrôler l'entropie permet donc de contrôler la distance (en norme L^1) entre ϕ et ϕ_∞ . En utilisant le fait que ϕ_∞ est une solution stationnaire, on vérifie que

$$\frac{dH(\phi|\phi_\infty)}{dt} = -\beta^{-1} \int \left| \nabla \ln \left(\frac{\phi}{\phi_\infty} \right) \right|^2 \phi. \quad (34)$$

Le membre de droite est négatif, et l'entropie H décroît donc au cours du temps.

Pour obtenir la convergence de H vers 0 (et donc de ϕ vers ϕ_∞ par (33)) on a ensuite besoin d'une inégalité de Sobolev logarithmique, dont il faut vérifier si elle est satisfaite pour la fonction ϕ_∞ que l'on considère : il existe une constante $R > 0$ telle que pour toute fonction ψ positive et d'intégrale 1,

$$\int \ln \left(\frac{\psi}{\phi_\infty} \right) \psi \leq \frac{1}{2R} \int \left| \nabla \ln \left(\frac{\psi}{\phi_\infty} \right) \right|^2 \psi. \quad (35)$$

En utilisant (35) dans (34), on en déduit la convergence exponentielle de H vers 0 :

$$H(\phi(t, \cdot)|\phi_\infty) \leq H(\phi(0, \cdot)|\phi_\infty) \exp(-2\beta^{-1} R t).$$

Il nous reste donc à expliquer pourquoi ϕ_∞ satisfait génériquement l'inégalité fonctionnelle (35) dans les applications qui nous intéressent ici.

Avant de rappeler quelques critères bien connus pour établir (35) on donne une définition précise de l'inégalité de Sobolev logarithmique, pour de futures références.

Définition 8 On dit que la mesure ν satisfait une inégalité de Sobolev logarithmique de constante $\rho > 0$ (noté par la suite $ISL(\rho)$) si et seulement si, pour toute mesure μ telle que μ est absolument continue par rapport à ν (noté $\mu \ll \nu$ dans la suite)

$$H(\mu|\nu) \leq \frac{1}{2\rho} I(\mu|\nu)$$

où $H(\mu|\nu) = \int \ln \left(\frac{d\mu}{d\nu} \right) d\mu$ est l'entropie de la mesure μ par rapport à la mesure ν et $I(\mu|\nu) = \int \left| \nabla \ln \left(\frac{d\mu}{d\nu} \right) \right|^2 d\mu$ l'information de Fisher de la mesure μ par rapport à la mesure ν .

Avec un léger abus de notation, on dit qu'une densité de probabilité satisfait une inégalité de Sobolev logarithmique si la mesure associée satisfait une inégalité de Sobolev logarithmique et pour deux densités de probabilité ψ et ϕ , on note l'entropie des mesures associées $H(\psi|\phi)$.

On rappelle le critère dit de Bakry-Emery :

Proposition 9 Soit Π une fonction α -convexe, au sens où pour tout vecteur X , $X^T \nabla^2 \Pi X \geq \alpha |X|^2$ (où $\nabla^2 \Pi$ désigne la Hessienne de Π). Alors la densité de probabilité $\psi_\infty \propto \exp(-\Pi)$ satisfait une inégalité de Sobolev logarithmique pour une constante $\rho \geq \alpha$.

On rappelle également le critère perturbatif de Holley-Strook :

Proposition 10 Soit Π une fonction telle que la densité de probabilité $\psi_\infty \propto \exp(-\Pi)$ satisfait une inégalité de Sobolev logarithmique de constante ρ . Soit une fonction $\tilde{\Pi}$ bornée et la densité de probabilité $\tilde{\psi}_\infty \propto \exp(-\Pi + \tilde{\Pi})$. Alors $\tilde{\psi}_\infty$ satisfait également une inégalité de Sobolev logarithmique de constante $\tilde{\rho} \geq \rho \exp(-\text{osc } \tilde{\Pi})$ où $\text{osc } \tilde{\Pi} = \sup \tilde{\Pi} - \inf \tilde{\Pi}$.

Donc, dès que le potentiel V est une perturbation bornée d'un potentiel α -convexe, ϕ_∞ satisfait une inégalité de Sobolev logarithmique.

En résumé, si ϕ_∞ satisfait $ISL(R)$, on a :

$$H(\phi(t, \cdot), \phi_\infty) \leq H(\phi(0, \cdot), \phi_\infty) \exp(-2\beta^{-1}Rt) \quad (36)$$

Par l'inégalité de Csiszar-Kullback (33), on en déduit la convergence vers 0 de $\int |\phi(t, \cdot) - \phi_\infty|$ à vitesse exponentielle, avec un taux $\beta^{-1}R$. On considère donc dans la suite que (3) à un comportement métastable si et seulement si R est proche de 0. Cette définition est raisonnable au sens où : (i) on vérifie que R est effectivement très proche de 0 pour des potentiels comportant des puits séparés par des barrières très hautes (ce qui correspond à l'image intuitive d'un potentiel métastable); (ii) la "vraie" vitesse de convergence de la dynamique (3) serait donnée par le trou spectral de l'opérateur $\text{div}(\nabla V \cdot + \beta^{-1} \nabla \cdot)$, et on peut montrer que la dégénérescence du trou spectral pour cet opérateur implique la dégénérescence de l'inégalité de Sobolev logarithmique pour la mesure μ (cf. [11, 6]).

Intéressons-nous maintenant à la dynamique (30). Nous allons appliquer une méthode d'entropie pour analyser la vitesse de convergence de cette dynamique vers son état stationnaire. La difficulté par rapport à l'étude de (3) et (32) est que l'équation de Fokker-Planck (31) est non-linéaire. Pour simplifier, nous supposons dans toute la suite que ξ est à valeurs dans le tore :

$$\xi : \mathbb{R}^n \rightarrow \mathbb{T}.$$

Les résultats se généralisent au cas où ξ est à valeurs dans un domaine non borné en ajoutant un potentiel confinant adéquat (cf. [31], un tel potentiel étant d'ailleurs utilisé en pratique pour les calculs d'énergie libre par de telles méthodes). De plus, nous allons

légèrement modifier la dynamique (30) pour simplifier l'étude. Nous considérons dans la suite de cette section la dynamique :

$$\begin{cases} dX_t = -\nabla(V - A_t \circ \xi - \beta^{-1} \ln(|\nabla\xi|^{-2}))(X_t)|\nabla\xi|^{-2}(X_t) dt + \sqrt{2\beta^{-1}}|\nabla\xi|^{-1}(X_t)dW_t, \\ A'_t(z) = \mathbb{E}(F(X_t)|\xi(X_t) = z), \end{cases} \quad (37)$$

associée à l'équation de Fokker-Planck :

$$\begin{cases} \partial_t \psi = \operatorname{div} (|\nabla\xi|^{-2} (\nabla(V - A_t \circ \xi)\psi + \beta^{-1}\nabla\psi)), \\ A'_t(z) = \frac{\int_{\Sigma_z} F \psi(t, \cdot) |\nabla\xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \psi(t, \cdot) |\nabla\xi|^{-1} d\sigma_{\Sigma_z}}. \end{cases} \quad (38)$$

Remarquer que dans le cas où $|\nabla\xi| = 1$ (i.e. ξ est la distance signée à Σ_0) la dynamique (37) est exactement (30). Dans le cas général, nous avons besoin de rajouter les termes fonctions de $|\nabla\xi|$ pour simplifier l'analyse. On renvoie à [31] pour des résultats similaires sur la convergence de la dynamique originale (30).

Nous devons introduire quelques notations. On note ψ^ξ la "densité marginale de ψ selon ξ " :

$$\psi^\xi(t, z) = \int_{\Sigma_z} \psi(t, \cdot) |\nabla\xi|^{-1} d\sigma_{\Sigma_z}. \quad (39)$$

On introduit également $\mu_t^\xi(dx|z)$ la mesure conditionnelle de $\psi(t, x) dx$ sachant que $\xi(x) = z$:

$$d\mu_t^\xi(dx|z) = \frac{\psi(t, \cdot) |\nabla\xi|^{-1} d\sigma_{\Sigma_z}}{\psi^\xi(t, z)}. \quad (40)$$

La formulation "densité marginale" est justifiée par le fait que, par la formule de la co-aire (8), on vérifie que l'image de la mesure $\psi(t, x) dx$ par la fonction ξ est la mesure $\psi^\xi(t, z) dz$. Dans le cas $\xi(x_1, \dots, x_n) = x_1$ par exemple, on a simplement (par Fubini) $\psi^\xi(t, z) = \int \psi(t, z, x_2, \dots, x_n) dx_2 \dots dx_n$ et $d\mu_t^\xi(dx|z) = \frac{\psi(t, z, x_2, \dots, x_n) dx_2 \dots dx_n}{\psi^\xi(t, z)}$.

Une remarque fondamentale est la suivante :

Lemme 11 *Soit ψ solution de (38). Alors la marginale ψ^ξ de ψ selon ξ définie par (39) satisfait l'équation de la chaleur :*

$$\partial_t \psi^\xi - \beta^{-1} \partial_{z,z} \psi^\xi = 0 \text{ sur } \mathbb{T}. \quad (41)$$

Il est très facile de vérifier ce lemme dans le cas où $\xi(x_1, \dots, x_n) = x_1$, en intégrant l'équation (38) par rapport à x_2, \dots, x_n . Dans le cas général, la preuve est basée sur la formule de la co-aire (8). Ce premier résultat montre déjà qualitativement que la dynamique (37) va plus vite vers son état stationnaire que la dynamique (3), du moins si ξ a été bien choisi. On voit en effet que l'équation sur ψ^ξ est une équation de la chaleur, correspondant à un potentiel constant : en utilisant la dynamique adaptative, on a "aplatis" le potentiel dans la direction de la coordonnée de réaction. Si cette direction est bien la direction de métastabilité principale pour (3), on a donc très fortement amélioré la vitesse de convergence en utilisant (37).

Nous allons maintenant quantifier cela. Nous avons besoin de quelques notations supplémentaires. On note ψ_∞ la densité stationnaire attendue :

$$\psi_\infty = \exp(-\beta(V - A \circ \xi)).$$

A la densité de probabilité ψ_∞ , on associe sa marginale selon ξ (de manière consistante avec (39)) :

$$\psi_\infty^\xi(z) = 1_{\mathbb{T}}(z),$$

et la mesure conditionnelle de $\psi_\infty(x) dx$ sachant que $\xi(x) = z$ (de manière consistante avec (40)) :

$$d\mu_\infty^\xi(\cdot|z) = Z_{\Sigma_z}^{-1} \exp(-\beta V) |\nabla \xi|^{-1} d\sigma_{\Sigma_z},$$

où on rappelle que (cf. (7))

$$Z_{\Sigma_z} = \int_{\Sigma_z} |\nabla \xi|^{-1} \exp(-\beta V) d\sigma_{\Sigma_z} = \exp(-\beta A(z)).$$

Pour ψ une solution de (38), on introduit l'entropie totale :

$$E(t) = H(\psi(t, \cdot) | \psi_\infty),$$

l'entropie macroscopique :

$$E_M(t) = H(\psi^\xi(t, \cdot) | \psi_\infty^\xi),$$

l'entropie locale à une valeur z de la coordonnée de réaction :

$$e_m(t, z) = H\left(\mu_t^\xi(\cdot|z) \middle| \mu_\infty^\xi(\cdot|z)\right) = \int_{\Sigma_z} \ln\left(\frac{\psi(t, \cdot)}{\psi^\xi(t, z)} / \frac{\psi_\infty}{\psi_\infty^\xi(z)}\right) \frac{\psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\psi^\xi(t, z)},$$

et enfin l'entropie microscopique

$$E_m(t) = \int_{\mathcal{M}} e_m(t, z) \psi^\xi(t, z) dz.$$

On vérifie facilement que

$$E(t) = E_M(t) + E_m(t).$$

Noter qu'en utilisant le Lemme (11), on sait déjà que ψ^ξ converge vers son état stationnaire ψ_∞^ξ à vitesse exponentielle. En particulier :

$$E_M(t) \leq E_M(0) \exp(-2\beta^{-1}rt), \quad (42)$$

où

$$r = 4\pi^2 \quad (43)$$

est appelé dans la suite taux de convergence macroscopique. Pour étudier la convergence de ψ vers ψ_∞ , il suffit donc d'étudier la convergence de E_m vers zéro.

Les dénominations "macroscopique" et "microscopique" font ici référence au fait que ξ est considéré comme une fonction qui à une configuration associe une grandeur macroscopique (variable lente). Les variables "orthogonales à ξ " (dans une feuille Σ_z) sont au contraire des grandeurs microscopiques (variables rapides).

Dans [31], on montre le résultat suivant :

Théorème 12 *On suppose :*

- (H1) $\xi : \mathbb{R}^n \rightarrow \mathbb{T}$ est une fonction régulière telle que $|\nabla \xi| > 0$.
- (H2) V et ξ sont des fonctions régulières telles que $\|\nabla \xi\|_{L^\infty} \leq m < \infty$ et $\|\nabla_{\Sigma_z} F\|_{L^\infty} \leq M < \infty$ (où $\nabla_{\Sigma_z} = P\nabla$ (cf. (17)) désigne le gradient surfacique),
- (H3) V et ξ sont telles que $\exists \rho > 0, \forall z \in \mathbb{T}$, la mesure conditionnelle $\mu_\infty^\xi(\cdot|z)$ satisfait une inégalité de Sobolev logarithmique de constante ρ .

Alors, on montre que

1. L'entropie microscopique E_m converge à vitesse exponentielle vers zéro : $\exists C > 0$, $\exists \lambda > 0$, $\forall t \geq 0$,

$$\sqrt{E_m(t)} \leq C \exp(-\lambda t). \quad (44)$$

Plus précisément, si $\rho m^{-2} \neq r$, alors

$$\lambda = \beta^{-1} \min(\rho m^{-2}, r). \quad (45)$$

Dans le cas particulier $\rho m^{-2} = r$, pour tout $\lambda < \beta^{-1} \min(\rho m^{-2}, r)$, il existe $C > 0$ tel que (44) est vérifié.

2. La racine carrée de l'entropie totale \sqrt{E} , les quantités $\|\psi(t, \cdot) - \psi_\infty\|_{L^1(\mathbb{R}^n)}$ et $\|A'_t - A'\|_{L^2(\mathbb{T})}^2$ convergent exponentiellement vite vers zéro au taux λ .

La preuve est basée sur un calcul explicite de $\frac{dE_m}{dt}$ et sur une majoration de $|A'_t(z) - A'(z)|$ par l'entropie relative de $\mu_t^\xi(\cdot|z)$ par rapport à $\mu_\infty^\xi(\cdot|z)$, basée sur une inégalité de Talagrand.

Ces résultats montrent donc que, essentiellement sous une hypothèse de couplage borné (H2) (cf. remarque 13 pour une justification de ce terme) et une hypothèse d'ergodicité des variables microscopiques (H3), A'_t converge exponentiellement vite vers A' au taux $\lambda = \beta^{-1} \min(\rho m^{-2}, r)$, qui doit être comparé à $\beta^{-1}R$ dans (36). Le taux de convergence λ est en fait seulement limité par la constante ρ de l'inégalité de Sobolev logarithmique pour les mesures conditionnelles $\mu_\infty^\xi(\cdot|z)$. En effet, on peut rendre en pratique r aussi grand que l'on veut de deux façons : en restreignant le domaine de calcul de l'énergie libre, ou en utilisant une méthode de sélection pour accélérer la convergence de ψ^ξ vers ψ_∞^ξ (cf. section 5). La valeur de ρ dépend évidemment du choix de la coordonnée de réaction. Dans notre langage, une "bonne coordonnée de réaction" est telle que ρ est aussi grand que possible.

Pour une généralisation de ces résultats (cas où ξ est à valeurs dans \mathbb{T}^p ou \mathbb{R}^p , cas de la dynamique ABF non-modifiée (30), discussion d'autres jeux d'hypothèses possibles), nous renvoyons à [31].

Remarque 13 *Il est naturel à ce stade de se poser la question suivante : sous les hypothèses du théorème 12, que peut-on dire de la constante de Sobolev logarithmique R pour la mesure "totale" μ . Dans [28], nous généralisons un résultat de [17] à une fonction ξ non-linéaire en montrant le critère suivant (appelé "critère à double-échelle") pour prouver une inégalité de Sobolev logarithmique pour la mesure μ . Si on suppose : (i) la marginale de μ selon ξ satisfait une ISL(\bar{r}), (ii) les mesures conditionnelles de μ à $\xi(x) = z$ fixé satisfont une ISL(ρ) (ce qui est (H3) dans le théorème 12), (iii) V et ξ sont tels que $\|\nabla_{\Sigma_z} F\|_{L^\infty} \leq M < \infty$ et $|\nabla \xi|^2 \geq m > 0$ (ce qui est relié à (H2) dans le théorème 12), alors μ satisfait une ISL de constante R avec*

$$R \geq \frac{1}{2} \left(\bar{r}m + \frac{M^2 m}{\rho} + \rho - \sqrt{\left(\bar{r}m + \frac{M^2 m}{\rho} + \rho \right)^2 - 4\bar{r}m\rho} \right).$$

On sait prouver l'égalité pour des lois normales et une fonction ξ linéaire. Noter que le membre de droite est majoré par $\min(\bar{r}m, \rho)$, qui correspond au cas d'un couplage nul ($M = 0$). L'hypothèse (iii) est appelée une hypothèse de couplage borné, car dans le cas particulier d'une loi normale et d'une fonction ξ linéaire, $\|\nabla_{\Sigma_z} F\|_{L^\infty}$ mesure la covariance entre les composantes microscopiques et les composantes macroscopiques.

Pour revenir à la question posée au début de cette remarque, on voit que la constante de Sobolev logarithmique R pour la mesure de Boltzmann-Gibbs μ est typiquement beaucoup plus petite que λ car la constante de Sobolev logarithmique pour la marginale de μ selon

ξ (notée \bar{r} ci-dessus) est typiquement beaucoup plus petite que la constante de Sobolev logarithmique pour la marginale de ψ_∞ selon ξ (notée r ci-dessus). On peut à nouveau penser à l'exemple très simple de la Figure 1 pour lequel \bar{r} et R sont très petits pour une grande barrière de potentiel, comparativement aux constantes r et ρ .

5 Méthodes adaptatives : discrétisation

Dans cette section, on discute la discrétisation des méthodes adaptatives, en considérant (37) pour fixer les idées. Les schémas de discrétisation proposés sont généralisables à d'autres dynamiques adaptatives comme celles évoquées dans la section 3, y compris quand elles sont construites à partir d'une dynamique de Langevin (5).

Dans la communauté "dynamique moléculaire", les physiciens utilisent plus volontiers des moyennes ergodiques sur de longues trajectoires que des moyennes empiriques sur plusieurs marcheurs pour calculer des moyennes. Ainsi, plutôt qu'une espérance conditionnelle, une méthode du type (37) utiliserait typiquement une mise à jour :

$$A'_t(z) = \frac{\int_0^t F(X_s) \delta^\epsilon(\xi(X_s) - z) ds}{\int_0^t \delta^\epsilon(\xi(X_s) - z) ds},$$

où δ^ϵ désigne une approximation de l'identité. Dans [30], nous proposons d'utiliser plutôt des systèmes particulières, avec plusieurs marcheurs en interaction et des moyennes empiriques. Ainsi, une approximation particulière de (37) s'écrit : pour $1 \leq n \leq N$, N désignant le nombre total de marcheurs,

$$\left\{ \begin{array}{l} dX_t^{n,N} = -\nabla(V - A_t^N \circ \xi - \beta^{-1} \ln(|\nabla \xi|^{-2}))(X_t^{n,N}) |\nabla \xi|^{-2}(X_t^{n,N}) dt \\ \quad + \sqrt{2\beta^{-1}} |\nabla \xi|^{-1}(X_t^{n,N}) dW_t^n, \\ (A_t^N)'(z) = \frac{\sum_{m=1}^N F(X_t^{m,N}) \delta^\epsilon(\xi(X_t^{m,N}) - z)}{\sum_{m=1}^N \delta^\epsilon(\xi(X_t^{m,N}) - z)}, \end{array} \right. \quad (46)$$

où $(W_t^n)_{t \geq 0, n \geq 1}$ désigne une famille de mouvements browniens indépendants. Cette approche est plus facile à analyser mathématiquement, très facilement parallélisable, et permet en outre d'ajouter des mécanismes de sélection pour sélectionner les "bons marcheurs" et tuer les "mauvais" afin d'accélérer la convergence (cf. ci-dessous). En pratique, on utilise plutôt une discrétisation par histogramme de $(A_t^N)'$ qu'une approximation de l'identité, mais la discrétisation (46) reste représentative des méthodes standards, qui comportent typiquement trois paramètres de discrétisation : le pas de temps, le nombre de marcheurs N , et un paramètre ϵ donnant la taille caractéristique de discrétisation de la variable coordonnée de réaction.

Dans [22], nous analysons la vitesse de convergence du système particulière (46) dans le cas particulier où $\xi(x_1, \dots, x_n) = x_1$ et l'espace de configuration est \mathbb{T}^n ou $\mathbb{T} \times \mathbb{R}^{n-1}$. Nous montrons notamment que l'erreur comporte un terme $\frac{C(T, \epsilon)}{\sqrt{N}}$, avec une constante $C(T, \epsilon)$ qui dépend du temps maximum T et du paramètre ϵ . L'expression de C que nous obtenons est cependant sous-optimale par rapport aux résultats numériques.

Nous terminons cette section en explicitant un exemple de mécanisme de sélection possible, proposé dans [30]. On modifie la dynamique (38) en considérant : pour une constante $c > 0$

$$\left\{ \begin{array}{l} \partial_t \psi = \operatorname{div}(|\nabla \xi|^{-2} (\nabla(V - A_t \circ \xi) \psi + \beta^{-1} \nabla \psi)) + c \left(\frac{\partial_{z,z} \psi^\xi}{\psi^\xi} \right) \circ \xi \psi, \\ A'_t(z) = \frac{\int_{\Sigma_z} F \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}{\int_{\Sigma_z} \psi(t, \cdot) |\nabla \xi|^{-1} d\sigma_{\Sigma_z}}. \end{array} \right. \quad (47)$$

On vérifie facilement que cette dynamique modifiée est encore telle que, si un état stationnaire (ψ_∞, A'_∞) est atteint, alors $A'_\infty = A'$. De plus, par un calcul similaire à celui permettant de prouver le Lemme 11, on montre que la marginale ψ^ξ de ψ suivant ξ satisfait alors (comparer avec (41)) :

$$\partial_t \psi^\xi - (\beta^{-1} + c) \partial_{z,z} \psi^\xi = 0 \text{ sur } \mathbb{T}.$$

L'ajout du terme supplémentaire dans Fokker-Planck permet donc d'améliorer la "vitesse de convergence macroscopique", qui est celle reliée au paramètre r de la section précédente (cf. (43)). La diffusion dans l'espace de la coordonnée de réaction est plus rapide.

Le système particulaire associé à (47) consiste à modifier (46) en associant un poids

$$w_{n,N}(t) = \exp \left(c \int_0^t \frac{\partial_{z,z} \psi^\xi}{\psi^\xi} (\xi(X_s^{n,N})) ds \right)$$

à la particule $(X_t^{n,N})$ et à effectuer une moyenne pondérée par ces poids pour le calcul de $(A_t^N)'$ dans (46). En pratique, on effectue de temps en temps une étape de reconfiguration stochastique, en dupliquant (respectivement tuant) les marcheurs qui ont le poids le plus élevé (respectivement le plus faible), en gardant un nombre total de marcheurs constant (cf. [12] pour une présentation de telles méthodes).

La quantité $\frac{\partial_{z,z} \psi^\xi}{\psi^\xi}$ est discrétisée par une méthode d'histogramme pour approximer ψ^ξ et des différences finies :

$$\begin{aligned} \frac{\partial_{z,z} \psi^\xi}{\psi^\xi}(z) &\simeq \frac{\psi^\xi(z + \delta z) - 2\psi^\xi(z) + \psi^\xi(z - \delta z)}{\psi^\xi(z) \delta z^2}, \\ &\simeq \frac{3}{\psi^\xi(z) \delta z^2} \left(\frac{\psi^\xi(z + \delta z) + \psi^\xi(z) + \psi^\xi(z - \delta z)}{3} - \psi^\xi(z) \right). \end{aligned}$$

Ainsi, la méthode donne plus de poids aux particules qui sont dans des régions moins explorées que les régions voisines.

On renvoie à [30] pour des expériences numériques montrant l'intérêt de cette méthode sur un exemple concret. En pratique, on observe qu'il existe un optimum pour le paramètre c : si c devient trop grand, la variance des résultats augmente ce qui dégrade le résultat (noter qu'augmenter c revient à augmenter la disparité des poids et donc à diminuer le nombre effectif de marcheurs contribuant à la moyenne donnant $(A_t^N)'(z)$).

Références

- [1] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids*. Oxford Science Publications, 1987.
- [2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford science publications, 2000.
- [3] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. Société Mathématique de France, 2000. In French.
- [4] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Part. Diff. Eq.*, 26 :43–100, 2001.
- [5] V. Babin, C. Roland, and C. Sagui. Adaptively biased molecular dynamics for free energy calculations. *J. Chem. Phys.*, 128 :134101, 2008.

- [6] D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New trends in stochastic analysis*, pages 43–75. World Scientific, 1997.
- [7] C. Chipot and A. Pohorille, editors. *Free Energy Calculations*, volume 86 of *Springer Series in Chemical Physics*. Springer, 2007.
- [8] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem*, 6(9) :1809–1814, 2005.
- [9] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds : Application to mean force computation. *Commun. Pur. Appl. Math.*, 61(3) :371–408, 2008.
- [10] E. Darve and A. Pohorille. Calculating free energy using average forces. *J. Chem. Phys.*, 115 :9169–9183, 2001.
- [11] E.B. Davies. Spectral properties of metastable Markov semigroups. *J. Funct. Anal.*, 52 :315–329, 1983.
- [12] P. Del Moral. *Feynman-Kac Formulae : Genealogical and Interacting Particle Systems with Applications*. Springer-Verlag, 2004.
- [13] W. E and E. Vanden-Eijnden. Metastability, conformation dynamics, and transition pathways in complex systems. In *Multiscale modelling and simulation*, volume 39 of *Lect. Notes Comput. Sci. Eng.*, pages 35–68. Springer, Berlin, 2004.
- [14] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, 1992.
- [15] E. Faou and T. Lelièvre. Conservative stochastic differential equations : Mathematical and numerical analysis. *Mathematics of computation*, 2008. to appear.
- [16] D. Frenkel and B. Smit. *Understanding molecular simulation : from algorithms to applications*. Academic Press, 2002.
- [17] N. Grunewald, F. Otto, C. Villani, and M.G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. H. Poincaré Probab. Statist.*, 2008. To appear.
- [18] J. Hénin and C. Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.*, 121 :2904–2914, 2004.
- [19] E.J. Hinch. Brownian motion with stiff bonds and rigid constraints. *J. Non-Newtonian Fluid Mech.*, 271 :219–234, 1994.
- [20] M. Iannuzzi, A. Laio, and M. Parrinello. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. *Phys. Rev. Lett.*, 90(23) :238302, 2003.
- [21] C. Jarzynski. Equilibrium free energy differences from nonequilibrium measurements : A master equation approach. *Phys. Rev. E*, 56(5) :5018–5035, 1997.
- [22] B. Jourdain, T. Lelièvre, and R. Roux. Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process. in preparation.
- [23] J.G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3 :300–313, 1935.
- [24] C. Le Bris, T. Lelièvre, and E. Vanden Eijnden. Analysis of some discretization schemes for constrained stochastic differential equations. *C.R. Acad. Sci. Paris, Ser. I.*, 346(7-8) :471–476, 2008.
- [25] F. Legoll and T. Lelièvre. Effective dynamics for the reaction coordinate. In preparation.

- [26] F. Legoll, T. Lelièvre, and G. Stoltz. Some remarks on sampling methods in molecular dynamics. *ESAIM Proceedings*, 22 :217–233, 2008.
- [27] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*. Cambridge University Press, 2004.
- [28] T. Lelièvre. A general two-scale criteria for logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 2008. to appear.
- [29] T. Lelièvre, M. Rousset, and G. Stoltz. Computation of free energy differences through nonequilibrium stochastic dynamics : The reaction coordinate case. *J. Comput. Phys.*, 222(2) :624–643, 2007.
- [30] T. Lelièvre, M. Rousset, and G. Stoltz. Computation of free energy profiles with adaptive parallel dynamics. *J. Chem. Phys.*, 126 :134111, 2007.
- [31] T. Lelièvre, M. Rousset, and G. Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21 :1155–1181, 2008.
- [32] W. K. den Otter and W. J. Briels. The calculation of free-energy differences by constrained molecular-dynamics simulations. *J. Chem. Phys.*, 109(11) :4139–4146, 1998.
- [33] M. Rousset and G. Stoltz. Equilibrium sampling from nonequilibrium dynamics. *J. Stat. Phys.*, 123(6) :1251–1272, 2006.
- [34] M. Sprik and G. Ciccoli. Free energy from constrained molecular dynamics. *J. Chem. Phys.*, 109(18) :7737–7744, 1998.
- [35] D. Talay. Approximation of upper Lyapunov exponents of bilinear stochastic differential systems. *SIAM J. Numer. Anal.*, 28(4) :1141–1164, 1991.
- [36] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4) :483–509 (1991), 1990.
- [37] N.G. van Kampen. Statistical mechanics for trimers. *App. Sci. Res.*, 37 :67–75, 1981.
- [38] F. Wang and D.P. Landau. Determining the density of states for classical statistical models : A random walk algorithm to produce a flat histogram. *Phys. Rev. E*, 64 :056101, 2001.