A constrained hybrid Monte-Carlo algorithm and the problem of calculating the free energy in several variables

Carsten Hartmann* and Christof Schütte†
Freie Universität Berlin
Institute of Mathematics II
Arnimallee 2-6
14195 Berlin, Germany

We consider the problem of computing molecular free energy profiles along several orthogonal reaction coordinates by means of constrained simulations. The reaction coordinates define families of submanifolds, and the mean force along the reaction coordinates is the averaged force acting vertically to the submanifold. We give a rigorous justification for the calculation of the mean force along the constrained coordinates, and provide a concise geometrical interpretation of the different contributions to the mean force in terms of the extrinsic geometry of the submanifold. From this we are able to derive a hybrid Monte-Carlo based algorithm that can be used to compute expectation values from constrained simulations such as the mean force in the context of thermodynamic free energy statistics.

PACS numbers: 52.65.Yy,65.40.Gr,02.70.Tt
Keywords: molecular dynamics, holonomic constraints, free energy, mean force, affine connection, normal connection, mean curvature, metric entropy, hybrid Monte-Carlo.

I. INTRODUCTION

Many simulations in molecular-dynamics (MD) applications are devoted to the calculation of free energy profiles along selected essential coordinates. These coordinates may be slow degrees of freedom or parameterizations of certain paths along which a reaction takes place. We understand the term *reaction* in a very broad sense, and mostly assign it to any conformational transition in a molecule.

In principle the free energy observable can be easily computed from the marginal probability distribution of the reaction coordinate that the underlying dynamics generates. However, since conformational changes occur rather rarely, reliably sampling these parts of phase space is a relatively tedious issue [1]. Hence one way out is to constrain the system to fixed coordinate values which correspond to rare events, and to compute the average force acting upon these coordinates. The free energy is then obtained as the potential of mean force. This technique is known as thermodynamic integration [2]. The hope is that once one has successfully identified the essential coordinates, sampling in the remaining phase space should be comparably fast.

During the last few years many attempts have been made to derive expressions for the mean force along specified coordinates that can be computed on the fly during constrained simulations. Most authors proposed to exploit the force of constraint that conserves the constraint in the course of integration [3–6]. Nevertheless at the beginning, there was some disagreement about the correct expression for the reactive force as well as about the proper averaging procedure — which is the probability measure to be taken? There are even incommensurable definitions of free energy as it was pointed out in the review [7].

The problem of the correct expression for the force has been addressed in great detail, yet the question of how to compute the respective average has remained open. To address the second topic we first reconsider the problem of deriving an expression for the force in a manner which allows us to address two key issues: at first it provides a clear and concise interpretation of the mean force along a reaction coordinate, and even more important, it enables us to derive a hybrid Monte-Carlo (HMC) scheme for constrained MD simulations. With this tool at hand we are then able to compute free energy profiles along several reaction coordinates from a constrained MD trajectory. The main advantage of this novel method is that it is conceptually very simple and lucid, and that ergodicity is guaranteed under rather mild assumptions. Additionally it offers control over the numerical discretization error. We shall shortly sketch our approach:

* To whom correspondence should be addressed. E-mail: chartman@math.fu-berlin.de
† E-mail: schuette@math.fu-berlin.de

Free energy as the potential of mean force

Consider a molecular system whose motion is governed by the Lagrangian function $L : T\mathbf{R}^n \to \mathbf{R}$. The corresponding Hamiltonian function $H : T^*\mathbf{R}^n \to \mathbf{R}$ is as usual the sum of kinetic and potential energy. We introduce a multidimensional reaction coordinate by defining a vector-valued function $c : \mathbf{R}^n \to \mathbf{R}^m$ depending upon the position variables $q \in \mathbf{R}^n$ only. One possible way then is to define the *Fixman potential* or the thermodynamical *Helmholtz free energy* $F = U - TS$ from the probability density of the level sets $c(q) = c'$

$$\exp(-\beta F(c')) = \int_{c(q)=c'} \exp(-\beta H(q,p))\mathrm{d}\sigma\,, \tag{I.1}$$

where (q,p) are the canonical coordinates on the phase space $T^*\mathbf{R}^n = \mathbf{R}^n \times \mathbf{R}^n$, the parameter $\beta = 1/T$ is the inverse temperature, and $\mathrm{d}\sigma = \mathrm{d}\sigma(q,p)$ is the surface element of the level sets $\Sigma \times \mathbf{R}^n$ considered as submanifolds in phase space with $\Sigma = \{q \,|\, c(q) = c'\} \subset \mathbf{R}^n$. We shall not conceal that a different notion of free energy which is by far more prevalent in the literature is provided by

$$\begin{aligned}\exp(-\beta G(c')) &= \int_{\mathbf{R}^n\times\mathbf{R}^n} \exp(-\beta H(q,p))\delta(c(q)-c')\,\mathrm{d}q\mathrm{d}p \\ &= \int_{c(q)=c'} \exp(-\beta H(q,p))\mathrm{vol}(\mathbf{D}c(q))^{-1}\mathrm{d}\sigma\,,\end{aligned}$$

where $\mathrm{vol}(\mathbf{D}c) = \sqrt{\det(\mathbf{D}c\mathbf{D}c^T)}$ denotes the matrix volume of the rectangular Jacobian matrix. Though this is certainly the traditional definition of free energy it has the unpleasant property that it depends on the parameterization of Σ as can be readily checked. For that reason we adopt the first definition (I.1), as it does not suffer from this ambiguity; it is typically used in transition state theory. All basic ideas stay the same, and at the end of this paper we will shortly explain how both concepts are related to each other.

For convenience we require that the vectors $\operatorname{grad} c_i$ are mutually orthogonal. Not only is this restriction not too severe, since in molecular dynamics applications these coordinates will often be orthogonal coordinates like, for instance, torsion angles, but it can be easily relaxed at the price of introducing further notation. Indeed it can be shown by taking the derivatives with respect to the reaction coordinate on either sides of (I.1) that computing the free energy is equivalent to asking for the mean force acting along the reaction coordinate

$$\frac{\partial F}{\partial c'} = \left(\int_{c(q)=c'} \exp(-\beta H(q,p))\mathrm{d}\sigma\right)^{-1} \int_{c(q)=c'} \frac{\partial H}{\partial c}\Big|_{c(q)=c'} \exp(-\beta H(q,p))\mathrm{d}\sigma\,, \tag{I.2}$$

that is, the free energy is the potential of the thus defined mean force. In principle the force $f_c = -\partial H/\partial c$ inside the integral can be computed after pulling back the Hamiltonian to a tubular neighbourhood of Σ and endowing it with an appropriate set of coordinates. Results in this regard have appeared in [4] for the alternative definition of free energy, yet there was a lack of understanding of the different contributions to the mean force. We shall reconsider the known result to some extend to obtain some insight into the underlying geometry.

We shall emphasize that picking out a good reaction coordinate is in general a nontrivial task which requires a lot of physical or chemical previous knowledge. In this sense there is no natural candidate for the reaction coordinate as there would be in the presence of symmetry [8], or in reduction onto certain characteristic manifolds [9]. It is also important to note that rather than obtaining a reduced system of equations the primary goal of free energy calculations is the free energy landscape from which statistical molecular properties are analyzed.

Constrained and unconstrained vector fields

In order to develop an algorithmic concept for computing the mean force, it is very common to treat the reaction coordinate $c(q) = c'$ as a holonomic constraint, and to show that the force in demand is the constraint force plus some further terms [3, 5, 6, 10]. As we will show, these further terms are related to the connection 1-form of the normal bundle over the configuration submanifold $\Sigma \subset \mathbf{R}^n$. Switching to the Lagrangian view we find that the generalized force can be equally well expressed as $f_c = \partial L/\partial c$. This yields

$$f_c^i(q,v) = -\lambda_i(q,v_0) + \|\operatorname{grad} c_i(q)\|^{-1}\,\langle v_1, \mathrm{d}n_i[v_0]\rangle)\,, \quad q \in \Sigma\,,$$

where λ_i labels the constraint force for the i-th component of the constraint, $v = v_0 + v_1$ is the decomposition of a generic tangent vector $v = \dot{q}$ into components $v_0 \in T_q\Sigma$, and $v_1 \in N_q\Sigma$, where the normal space is spanned by an

orthonormal frame $\{n_1, \dots, n_m\}$. The expression $\langle v_1, \mathrm{d}n_i[v_0]\rangle$ is related to the connection 1-form associated with the normal frame; in some sense the last term can be understood as a corrective force that acts on the intrinsic motion on the constrained phase space $T\Sigma$.

We have to be a bit careful as the approach using holonomic constraints is very subtle. On the one hand we want to evaluate the force f at $c = c'$ regardless of $\dot{c}$, and on the other hand the constraint imposes the condition $\dot{c}(q) = 0$ on the velocities. For that reason the supposedly attractive Routh construction $L_\lambda = L - \lambda^i c_i$, or as well the Dirac bracket are bound to miss the point: either $\dot{c} = 0$ on both sides of the equals sign, or $\dot{c} \neq 0$ in case of which it is not clear why λ should be the constraint force.

Conditional averaging with hybrid Monte-Carlo

In principle the above expression for f_c allows for an averaging scheme using any probability distribution for the non-constrained coordinates. Regarding free energy calculations as opposed to classical averaging [11, 12] or homogenization in time [13], we consider averages over a submanifold with a weight given by the respective Gibbs distribution $\mu \propto \exp(-\beta E)$, where E denotes the total energy of L. Therefore it cannot be the case that reducing the dynamics to the free energy landscape will provide a pathwise approximation to the original dynamics. Since the free energy is a thermodynamical concept, we rather expect convergence in probability with respect to the Gibbs measure [14].

The velocity average of the force f_c can be computed analytically. It is known that the velocity contribution to the constraint force is related to the second fundamental form of the embedding, describing the curvature of the constraint surface in the surrounding space (extrinsic curvature). On average only the mean curvature of Σ in $\mathbf{R}^n$ contributes to the force, since the connection 1-form vanishes. We find

$$\bar{f}_c^i(q) = -\|\mathrm{grad}\, c_i(q)\|^{-1}(\langle n_i(q), \mathrm{grad}\, U(q)\rangle - \beta^{-1}\kappa(q))\,, \quad q \in \Sigma\,,$$

where κ denotes the mean curvature. The mean curvature stands for the metric entropy of the constraint surface, whereas the vanishing connection indicates that the average ambient space forcing on the surface is zero.

Once the velocities have been averaged out, the configuration space average can be efficiently computed from the constrained Gibbs density. Intriguingly the constrained density μ_Σ is related to the just described conditional density on $T\mathbf{R}^n|\Sigma = \Sigma \times \mathbf{R}^n$ in a very simple way

$$\mu_\Sigma(q, v) = \mu(q, v_0)\,, \quad (q, v_0) \in T\Sigma\,.$$

Thus μ_Σ is merely the full density μ restricted to the constrained phase space. This representation is very convenient because it does not rely on introducing local coordinates, and we can use it to set up the hybrid Monte-Carlo procedure which is presented herein quite efficiently. Our method is similar to the famous *blue moon sampling* but considerably simpler as there is no reweighting involved [15]. This simplicity is owed to the particular choice of the free energy but it can easily be carried forward to the traditional definition.

In principle, computing the mean force is not a big deal using any of the standard thermostat techniques. However they have to meet two major requirements: firstly the thermostat has to be consistent with constrained dynamics, and secondly the dynamics has to be ergodic with respect to the constrained Gibbs density in configuration space. It is well-known that the ordinary Nosé-Hoover thermostat suffers from ergodicity problems for certain classes of Hamiltonians [16, 17]. This pathology can be removed by using extensions to the single-oscillator chain or the Andersen constant temperature thermostat [18, 19]. But even then, the sampling works well *only if* the dynamics is ergodic, and conditions to guarantee ergodicity are still lacking. Additionally all these more sophisticated methods have in common that due to their complexity they are relatively hard to implement, and they require a careful adjustment of the parameters involved. Even worse, it is not clear a priori how these methods fit constrained symplectic integration; see [20] for a discussion on that topic. In particular in the Nosé-Hoover method the constraint force becomes dependent on the thermostat variables, which means that it can no longer be interpreted as the constraint force for the subsystem. A promising alternative is to use stochastic high friction Langevin dynamics since it is ergodic for the systems under considerations [7]. However since our approach is heavily based on deterministic dynamics we shall not discuss this method here.

We adopt the hybrid Monte-Carlo (HMC) technique. Belonging to the class of Metropolis Monte-Carlo it is conceptually very simple, and it is designed to handle symplectic integration, as one can use standard integrators for constrained Hamiltonian systems. Moreover it is proven that the HMC dynamics is ergodic with respect to the positional density under rather weak conditions which are fulfilled for our purposes. We show that HMC indeed generates the constrained Gibbs density with an acceptance procedure that is similar to the usual Metropolis acceptance step. We further demonstrate how HMC can be efficiently used in connection with ordinary constrained molecular dynamics

and show how the proposal can easily be generated. As an additional treat the acceptance procedure also controls the numerical error, for the HMC rejects those moves that have too large energy fluctuations.

II. SOME GEOMETRY

Reaction coordinates and the molecular Lagrangian

We start from the Lagrange function $L : T\mathbf{R}^n \to \mathbf{R}$ in an Euclidean configuration space that is defined by

$$L(q, v) = \frac{1}{2} \langle Mv, v\rangle - U(q)\,,$$

where M is the diagonal mass matrix, $v = \dot{q}$ and $U : \mathbf{R}^n \to \mathbf{R}$ is the molecular interaction potential. In order to keep a compact notation, we choose mass-scaled coordinates $q \mapsto M^{1/2}q$ which allows us to set $M = \mathbf{1}$ in the following. By no means this will affect our considerations; all computations can be carried out with respect to the scaled coordinates using the standard inner product of $\mathbf{R}^n$. A reaction coordinate is a smooth function $c : \mathbf{R}^n \to \mathbf{R}^m$ the level sets of which define a family of configurational submanifolds

$$\Sigma = \{q \in \mathbf{R}^n \,|\, c : \mathbf{R}^n \to \mathbf{R}^m,\, c(q) = c'\} \subset \mathbf{R}^n\,. \tag{II.1}$$

Note that a holonomic constraint $c(q) = c'$ further requires that $\dot{c}(q) = 0$. We assume that c is sufficiently smooth, and that for a given value c' the Jacobian $\mathbf{D}c(q)$ has maximum rank m everywhere on Σ, such that Σ is a proper submanifold of codimension m in $\mathbf{R}^n$. The normal bundle over Σ is defined as

$$N\Sigma = \{(\sigma, n) \,|\, \sigma \in \Sigma, n \in N_\sigma\Sigma\} \subset \mathbf{R}^n \times \mathbf{R}^n\,.$$

Then in a sufficiently small neighbourhood there is a natural diffeomorphism of $N\Sigma$ into $\mathbf{R}^n$ given by $(\sigma, n) \mapsto \sigma + n$, and we can pull back the Euclidean metric considering $N\Sigma$ as our classical configuration space. This is reasonable since we observe the system's behaviour only in the vicinity of the submanifold Σ. Let $(\sigma(t), n(t))$ be an arbitrary curve in $N\Sigma$ passing through (σ, n) at time $t = 0$. We identify the tangent space $T_{(\sigma,n)}N\Sigma$ with the n-dimensional subspace of $\mathbf{R}^n \times \mathbf{R}^n$ that is spanned by all vectors of the form $(X, Y) = (\dot{\sigma}(0), \dot{n}(0))$. Note that clearly $(X, Y) \mapsto X + Y$ maps the tangent vectors back into $\mathbf{R}^n$. Hence setting the mass matrix to unity the molecular Lagrangian reads

$$\tilde{L}(\sigma, n, \dot{\sigma}, \dot{n}) = \frac{1}{2} \langle \dot{\sigma} + \dot{n}, \dot{\sigma} + \dot{n}\rangle - U(\sigma + n)\,. \tag{II.2}$$

We introduce two projections: for each $\sigma \in \Sigma$ consider a decomposition $T_\sigma\mathbf{R}^n = T_\sigma\Sigma \oplus N_\sigma\Sigma$ into tangent and normal space. The corresponding orthogonal projections are denoted by H_σ, V_σ. By this we obtain a decomposition of $\mathbf{R}^n$, since we can naturally identify $T_\sigma\mathbf{R}^n$ with $\mathbf{R}^n$.

Coordinate expressions

Suppose $\sigma : \mathbf{R}^{n-m} \to \Sigma$ is a local embedding of $\mathbf{R}^{n-m}$ onto $\Sigma \subset \mathbf{R}^n$, and for each $\sigma \in \Sigma$ we are given an orthonormal frame $n_1(\sigma), \ldots, n_m(\sigma)$ which spans $N_\sigma\Sigma$. As a consequence of the maximum rank assignment for $\mathbf{D}c(q)$and the Implicit Function Theorem [21] we may introduce local coordinates $x^1, \ldots, x^{n-m}, y^1, \ldots, y^m$ in a tubular neighbourhood of Σ, such that $x^1, \ldots, x^{n-m}$ are coordinates on the manifold and $y^1, \ldots, y^m$ measure the distance from it in the frame given by the normalized gradients of the components of the map $c(q)$. Then in the tubular neighbourhood the relation between the Cartesian coordinates $q \in \mathbf{R}^n$ and the local coordinates is just

$$q(x, y) = \sigma(x) + y^i n_i(\sigma(x))\,, \tag{II.3}$$

whereas the reaction coordinate is related to the normal coordinates by

$$c_i(q(x, y)) - c'_i = \|\mathrm{grad}\, c_i(\sigma(x))\| y_i \tag{II.4}$$

without summation over i. Thus if the gradients $\mathrm{grad}\, c_i(\sigma)$ vanish nowhere, then Σ is completely determined by the condition $y = 0$. We endow the tangent spaces $T_{(\sigma,n)}N\Sigma$ with the standard bases $\partial/\partial x^1, \ldots, \partial/\partial x^{n-m}$ and

$\partial/\partial y^1, \ldots, \partial/\partial y^m$, respectively. Abbreviating $z = (x, y)$ we can compute the local coordinate expression for the metric tensor by

$$g_{ij}(z) = \delta_{kl} \frac{\partial q^k}{\partial z^i} \frac{\partial q^l}{\partial z^j}\,, \quad i, j, k, l = 1, \ldots, n$$

where q^k are the components of the coordinate map $q(z)$. For the sake of convenience we may write $X^l = \partial\sigma/\partial x^l$ for vectors tangent to the manifold; then the metric tensor takes the following form

$$g(x, y) = \begin{pmatrix} G(x) + C(x, y) & A(x, y) \\ A(x, y)^T & \mathbf{1} \end{pmatrix}, \tag{II.5}$$

where the matrix $G(x) \in \mathbf{R}^{(n-m)\times(n-m)}$ is the metric tensor induced on Σ by the restriction of the Euclidean metric, that is $G_{ij} = \langle X^i, X^j \rangle$. The matrix $C(x, y) \in \mathbf{R}^{(n-m)\times(n-m)}$ has the entries

$$C_{ij} = 2y^l \left\langle \mathrm{d}n_l[X^i], X^j \right\rangle + y^k y^l \left\langle \mathrm{d}n_k[X^i], \mathrm{d}n_l[X^j] \right\rangle .$$

If Σ has codimension one, then the matrices G, C can be given a nice geometrical meaning: let M denote the matrix of the Weingarten map that is associated with the second fundamental form of the embedding; then it is true that $G + C = G(\mathbf{1} - M)^2$ (cf. [22, 23]). The elements of the off-diagonal matrix $A(x, y) \in \mathbf{R}^{(n-m)\times m}$ are

$$A_{ij} = y^l \left\langle n_i, \mathrm{d}n_l[X^j] \right\rangle .$$

Finally, the Lagrangian (II.2) has the local coordinate expression

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2} \left\langle (G(x) + C(x, y))\dot{x}, \dot{x} \right\rangle + \left\langle A(x, y)\dot{y}, \dot{x} \right\rangle + \frac{1}{2} \left\langle \dot{y}, \dot{y} \right\rangle - U(q(x, y))\,. \tag{II.6}$$

III. FORCE EXPRESSIONS

Using the previous arrangements we are able to compute the force along the reaction coordinate whenever it is close to the submanifold Σ. Let $z(t)$ be the coordinate expression for a curve $(\sigma(t), n(t))$ passing through Σ, and which is a solution of the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{z}^i} - \frac{\partial L}{\partial z^i} = 0\,. \tag{III.1}$$

For a given molecular potential this curve will only occasionally pass through Σ, or stay close to it. Suppose the trajectory hits Σ at $t = 0$, and stays in the tubular neighbourhood in the time interval $t \in [-\epsilon, \epsilon]$ for $\epsilon > 0$. As the curve is an integral curve of (III.1) we can compute the generalized force f_y at $t = 0$ by

$$f_y^i(x, 0, \dot{x}, \dot{y}) = \frac{d}{dt} \frac{\partial L}{\partial \dot{y}^i}\Big|_{y=0} = \frac{\partial L}{\partial y^i}\Big|_{y=0}\,.$$

Once we have computed f_y^i we can exploit (II.4), and switch to f_c^i by means of the relation $f_c^i = \|\mathrm{grad}\, c_i\|^{-1} f_y^i$. We can address the kinetic and the potential energy part separately, and we shall start with the molecular potential:

$$\frac{\partial U}{\partial y^i}\Big|_{y=0} = n_i(\sigma(x)) \cdot \mathrm{d}U(\sigma(x))\,. \tag{III.2}$$

The equation states that the potential energy contribution is the projection of the molecular force field onto the normal space which is basically the directional derivative of U along the vector n_i. Omitting the potential from now on we continue with the kinetic energy part: a direct calculation reveals

$$\frac{\partial L}{\partial y^i}\Big|_{y=0} = S_i^{jk}(\sigma(x))\, \dot{x}_j \dot{x}_k + \omega_i^l(\sigma(x))[X^k]\, \dot{y}_l \dot{x}_k\,, \tag{III.3}$$

where we introduced the short-hand notation $S_i^{jk} = \left\langle \mathrm{d}n_i[X^j], X^k \right\rangle$, and $\omega_i^l[\cdot] = \left\langle n^l, \mathrm{d}n_i[\cdot] \right\rangle$ for the partial derivatives of the metric tensor elements at $y = 0$. Both terms have a nice geometrical interpretation: S_i are the Weingarten maps, and ω_i^l are the connection coefficients of the normal connection, as we will explain now.

a. Extrinsic Curvature. We begin with the Weingarten maps: let X, Y, be two vector fields tangent to Σ; then the second fundamental form of the embedding $\Sigma \subset \mathbf{R}^n$ is defined by [24]

$$I\!I(X,Y) = V_\sigma \mathrm{d}X[Y] = n^l \left\langle n_l, \mathrm{d}X[Y] \right\rangle \,.$$

It follows by a simple polarization argument that $I\!I$ is symmetric in its arguments. Now if $X^k = \partial\sigma/\partial x^k \in \mathbf{R}^n$ is tangent to Σ, we can differentiate the relation $\left\langle n_i, X^k \right\rangle = 0$ along a vector X^j; it follows immediately that

$$I\!I(X^j, X^k) = n^l \left\langle X^k, S_l X^j \right\rangle$$

with the symmetric Weingarten maps $S_i = -H_\sigma \mathrm{d}n_i[\cdot]$, where we remind the reader that the summation convention is in force. The matrices of the Weingarten maps in the basis of the tangent vectors X^j, X^k have the entries

$$W_i^{jk} = G^{jl} \left\langle X_l, \mathrm{d}n_i[X^k] \right\rangle \,.$$

Bear in mind that $\dot\sigma = D_x\sigma \dot x$. Setting $J = D_x\sigma$ we find that $\dot x = G^{-1}J^T\dot\sigma$ as $G = J^TJ$ is invertible. Since $G^{-1}J^T$ is the Moore-Penrose pseudo-inverse of J, we uncover $R_i = H_\sigma J W_i G^{-1} J^T$ as the matrix of the Weingarten map in the original coordinate basis, namely

$$S_i^{jk}\, \dot x_j \dot x_k = R_i^{jk}\, \dot\sigma_j \dot\sigma_k \,, \quad R_i \in \mathbf{R}^{n\times n} \,,\; \dot\sigma \in T_\sigma\Sigma \,.$$

b. Normal connection. Commonly the normal connection of the embedding $\Sigma \subset \mathbf{R}^n$ is expressed by its connection 1-forms which coincide with our $\omega_i^l[\cdot]$ (cf. [25, 26]). Addressing these terms, we can distinguish two cases: first of all if Σ has codimension $m = 1$ in $\mathbf{R}^n$ then it is a well-known fact that $\mathrm{d}n[\cdot] \in T_\sigma\Sigma$ (Weingarten's equation [27]). Consequently the matrix $A \in \mathbf{R}^{(n-m)\times 1}$ in the expression (II.5) for the metric tensor is zero, and the connection term does not appear at all. However if Σ has codimension $m > 1$ then it is no longer true that $\mathrm{d}n_i[\cdot] \in T_\sigma\Sigma$, and ω_i^l which indicates the forcing of the normal vector field will contribute to the force. Recalling that $n = y^i n_i(\sigma(x)) \in \mathbf{R}^n$ we have the following identity to hold at $y = 0$

$$\omega_i^l[X^k]\, \dot y_l \dot x_k = \left\langle \dot n, \mathrm{d}n_i[\dot\sigma] \right\rangle \,, \quad \dot\sigma \in T_\sigma\Sigma \,,\; \dot n \in N_\sigma\Sigma \,.$$

Computing the mean force

Suppose we have computed the force at each hitting event. Let $I = \{0, t_1, \ldots, t_N\}$ denote the set of hitting times. For each hitting event $(\sigma_k, 0, \dot\sigma_k, \dot n_k) = (\sigma(t_k), n(t_k), \dot\sigma(t_k), \dot n(t_k))$ with $t_k \in I$ the force f_y^i reads

$$f_y^i(\sigma_k, 0, \dot\sigma_k, \dot n_k) = \langle n_i(\sigma_k), \mathrm{grad}\, U(\sigma_k)\rangle - \langle R_i(\sigma_k)\dot\sigma_k, \dot\sigma_k\rangle - \langle \dot n_k, \mathrm{d}n_i(\sigma_k)[\dot\sigma_k]\rangle \,.$$

The first two terms constitute the constraint force which depends only on the tangential velocities, and which would have appeared if we had imposed the additional condition $\dot n = 0$ on the Euler Lagrange equations [28].

We assume that the system (III.1) is ergodic with respect to the Gibbs density $\exp(-\beta E)$, where E is the total energy of the Lagrangian. Strictly speaking, we will only be able to assure ergodicity for the discrete Markov chain that is generated by the numerical hybrid Monte-Carlo scheme following in the next Section, and that approximates the flow of (III.1) with initial velocities distributed according to the Maxwell distribution [29]. From an algorithmic perspective we will be dealing with discrete trajectories only, and therefore as N goes to infinity the hits will be distributed according to the conditional Gibbs distribution on Σ. To set up the stage for the hybrid Monte-Carlo procedure now, we have to get rid of the velocities in the expression for the force. If we consult (II.2) while taking into account that $\dot n \in N_\sigma\Sigma$ whenever a trajectory passes through Σ, we see that the unnormalized density factorizes according to

$$\mu(\sigma, 0, \dot\sigma, \dot n) = \exp(-\beta E(\sigma, 0, \dot\sigma, \dot n)) = \exp(-\beta/2\, \langle \dot\sigma, \dot\sigma\rangle) \exp(-\beta/2\, \langle \dot n, \dot n\rangle) \exp(-\beta U(\sigma)) \,.$$

Hence we consider integrals of the type

$$\int \langle B(\sigma)\, v, v\rangle \, \exp\left(-\beta/2\, \langle v, v\rangle\right) \mathrm{d}v^1 \ldots \mathrm{d}v^n$$

with $B(\sigma) \in \mathbf{R}^{n\times n}$ being either the matrix of the Weingarten map or the matrix of the normal connection; its particular form will become apparent in the following. The remaining task is similar to computing the covariance

matrix of a Gaussian density. Since no cross-terms occur in the density, it is clear that only diagonal terms $B_{ii}v^iv^i$ will survive the averaging procedure. The curvature term entails

$$\begin{aligned}\theta_i^1 &= \int \langle R_i(\sigma)\,\dot\sigma,\dot\sigma\rangle\,\exp(-\beta/2\,\langle\dot\sigma,\dot\sigma\rangle)\,\mathrm{d}\dot\sigma\\ &= \int \langle S_i(\sigma(x))\,\dot x,\dot x\rangle\,\exp(-\beta/2\,\langle G(x)\dot x,\dot x\rangle)\,\sqrt{G(x)}\,\mathrm{d}\dot x^1\ldots\mathrm{d}\dot x^{n-m}\\ &= \int \left\langle G^{-1}(x)S_i(\sigma(x))\,\dot u,\dot u\right\rangle\,\exp(-\beta/2\,\langle\dot u,\dot u\rangle)\,\mathrm{d}\dot u^1\ldots\mathrm{d}\dot u^{n-m}\\ &= \beta^{-1}\mathrm{Tr}\,(W_i(\sigma))\,.\end{aligned}$$

The last expression is known to be the mean curvature of Σ in $\mathbf{R}^n$; since the trace is invariant under similarity transformations this is just as well computed by $\mathrm{Tr}\,(W_i) = \mathrm{div}\,n_i$. The connection vanishes completely as the first factor appears only linearly, and therefore has zero mean

$$\begin{aligned}\theta_i^2 &= \int \langle\dot n,\mathrm{d}n_i[\dot\sigma]\rangle\,\exp(-\beta/2\,\langle\dot n,\dot n\rangle)\,\mathrm{d}\dot n\\ &= \int \langle\dot y^l n_l(\sigma),\mathrm{d}n_i[\dot\sigma]\rangle\,\exp(-\beta/2\dot y^l\dot y_l)\,\mathrm{d}\dot y^1\ldots\mathrm{d}y^m\\ &= 0\,.\end{aligned}$$

Let us summarize the last few steps: denoting the velocity averaged force by a bar, $\bar f_y^i$ is given by

$$\bar f_y^i(\sigma) = \langle n_i(\sigma),\mathrm{grad}\,U(\sigma)\rangle - \beta^{-1}\mathrm{div}\,n_i(\sigma)\,. \tag{III.4}$$

Remark 1. *Carefully inspecting the last equation reveals a manifest physical interpretation for the curvature contribution: recall the thermodynamical definition of the Helmholtz free energy $F = U - TS$, where $T = 1/\beta$ is the physical temperature. The first part U is the energy contribution which is equivalent to our potential energy function U in the equations of motion. Thus comparing F to the formula (III.4) for the average force we claim that the mean curvature represents the entropy contribution to the free energy. This relation is in fact established for geodesic flows on compact manifolds of negative curvature, since this type of flow is known to be mixing and therefore ergodic (cf. [30] and the references therein).*

IV. HYBRID MONTE-CARLO

So far it is by no means obvious how a discretization of (III.1) should put forth a canonical density or Gibbs density rather than the microcanonical density. In order to calculate conditional averages in (I.2) we have to develop an algorithmic concept that allows for computing canonical averages preferably avoiding any artificially added heat bath dynamics. As the velocities in the force observable have been averaged out we only have to compute configurational averages. Our approach is divided into two steps: first of all we start with the constrained dynamics and show how the constrained Gibbs density in configuration space is generated by an appropriate hybrid Monte-Carlo scheme. Then we explain the relation between the constrained invariant density and the conditional density.

Constrained invariant measure

In order to simplify the following discussion concerning the Monte-Carlo approach we switch from the Lagrangian to the Hamiltonian view. To this end, we introduce the molecular Hamiltonian function which is the equivalent of the total energy E on the cotangent bundle. Let $L : T\mathbf{R}^n \to \mathbf{R}$ be the regular molecular Lagrangian (II.6). The Hamiltonian $H : T^*\mathbf{R}^n \to \mathbf{R}$ is obtained by the Legendre transform of L

$$H(z,w) = \langle\dot z,w\rangle - L(z,\dot z) \quad\text{with}\quad \mathbb{F}L : T\mathbf{R}^n \to T^*\mathbf{R}^n,\ w_i = \frac{\partial L}{\partial\dot z^i}\,, \tag{IV.1}$$

which can be equivalently written in the form $H = E\circ\mathbb{F}L^{-1}$. By definition of the fibre derivative $\mathbb{F}L$ the momenta $w = (u,v)$ are the conjugate momenta to $z = (x,y)$ which means that

$$\omega = \mathrm{d}x^i\wedge\mathrm{d}u_i + \mathrm{d}y^j\wedge\mathrm{d}v_j$$

is the standard symplectic form in local coordinates. Here we use ω to label the symplectic form rather than the connection 1-form as before. The constrained symplectic form is the restriction of ω to the constrained phase space $T^*\Sigma$, that is $\omega_\Sigma = \omega|_{T^*\Sigma}$,

$$\omega_\Sigma = \mathrm{d}x^i \wedge \mathrm{d}u_i \,.$$

This tells us how the canonical invariant density $\mu \propto \exp(-\beta E \circ \mathbb{F}L^{-1})$, and the corresponding invariant measure of the constrained dynamics looks like, namely

$$\nu_\Sigma(\mathrm{d}x, \mathrm{d}u) = Z_\Sigma^{-1} \exp(-\beta/2 \left\langle G(x)^{-1}u, u\right\rangle) \exp(-\beta U(\sigma(x)) \, \mathrm{d}x^1 \ldots \mathrm{d}u^{n-m} \tag{IV.2}$$

with $u_i = \partial L/\partial \dot{x}^i$ denoting the conjugate momentum to x, and Z_Σ being a normalization constant. In the following we will denote by μ_Σ the corresponding smooth density, such that $\nu_\Sigma(\mathrm{d}x, \mathrm{d}u) = \mu_\Sigma(x,u)\lambda_\Sigma$. The Liouville measure $\lambda_\Sigma = \mathrm{d}x^1 \ldots \mathrm{d}u^{n-m}$ thus defined is associated with the constrained Liouville volume form

$$\Lambda_\Sigma = \frac{(-1)^{d(d-1)/2}}{d!} \omega_\Sigma \wedge \ldots \wedge \omega_\Sigma \tag{IV.3}$$

that is obtained from the symplectic form ω_Σ by taking the d-fold exterior product, where $d = n - m$ is the dimension of the constraint submanifold $\Sigma \subset \mathbf{R}^n$.

Constrained hybrid Monte-Carlo

Apparently we cannot separate the density into merely momentum and position dependent parts, and so we write $\mu_\Sigma(x,u) = Q(x)P_x(u)$ for the density, indicating that the kinetic energy depends on the position coordinates as well. In order to see how the constrained HMC scheme works with μ_Σ, we let the discrete one-parameter semigroup $\varphi^{k\tau}$ denote a symmetric and symplectic flow map acting on the constrained subspace. Then for a single integration step $(\tilde{x}, \tilde{u}) = \varphi^\tau(x,u)$ the HMC acceptance function is defined by

$$a(x,\tilde{x}) = \min\left\{1, \frac{\mu_\Sigma(\tilde{x}, -\tilde{u})\mathrm{d}\tilde{x}^1 \ldots \mathrm{d}\tilde{u}^d}{\mu_\Sigma(x,u)\mathrm{d}x^1 \ldots \mathrm{d}u^d}\right\} = \min\left\{1, \frac{Q(\tilde{x})P_{\tilde{x}}(\tilde{u})}{Q(x)P_x(u)}\right\} . \tag{IV.4}$$

The definition of $a(x,\tilde{x})$ relies on the symmetry of $P_x(u)$ and the Liouville measure with respect to momentum inversion $u \mapsto -u$ as well as on the invariance of the Liouville measure under φ^τ (see Remark 2 below). Note that if φ^τ were the exact Hamiltonian flow, then clearly $a(x,\tilde{x}) = 1$. But as in general the discrete flow is not exactly energy preserving, the acceptance function takes values less than one. In this sense the acceptance and rejection procedure also controls the numerical error, and we have to show that propagating according to the HMC transition function which for deterministic flows is equal to the acceptance function $a(x,\tilde{x})$ leaves the Gibbs density μ_Σ invariant. This is by no means trivially fulfilled, since our Hamiltonian system has no unique invariant measure. The following statement holds:

Proposition 1. *Let $x_* \in \mathbf{R}^d$ be an accepted position value after a single HMC run involving an integration and an acceptance step. We assume that the initial momentum u is distributed according to $\mu_\Sigma(x,\cdot) = Q(x)P_x(\cdot)$. Furthermore let $\vartheta(x_*)$ denote the marginal distribution of the position variables after one HMC step. Then for any position dependent observable $\mathcal{O} \in L^1(\nu_\Sigma)$ the identity $\mathbf{E}_\vartheta[\mathcal{O}(x)] = \mathbf{E}_Q[m(x)\mathcal{O}(x)]$ holds for the expectation values, where $m(x)$ depends only on the Riemannian structure of the constraint surface Σ.*

Proof. The proof follows the outline of [31]. For convenience, we drop the subscript Σ. Suppose the initial position x follows the Gibbs distribution $\mu(\cdot,u)$. Then for a given x we draw a momentum vector from $P_x(u)$, and propagate a single time-step according to $(\tilde{x},\tilde{u}) = \varphi^\tau(x,u)$. Using $\tilde{\lambda} = \lambda \circ \varphi^{-\tau}$ to label the pulled-back constrained Liouville measure, and performing the acceptance and rejection procedure we obtain for the expectation value

$$\begin{aligned}\mathbf{E}_\vartheta[\mathcal{O}] \;=\; \int \mathcal{O}(x_*) \Bigg(& \mu(\varphi^{-\tau}(x_*,\tilde{u})) \min\left\{1, \frac{\mu(x_*,\tilde{u})}{\mu(\varphi^{-\tau}(x_*,\tilde{u}))}\right\} \det(\mathbf{D}\varphi^{-\tau}) \\ & + \mu(x_*,\tilde{u}) \left(1 - \min\left\{1, \frac{\mu(\varphi^{-\tau}(x_*,\tilde{u}))}{\mu(x_*,\tilde{u})}\right\}\right)\Bigg) \tilde{\lambda}\,,\end{aligned}$$

where the first term of the right-hand side of the equation comes from the acceptance, the second one stems from the rejection step. Furthermore in the second term we made use of the symmetry of φ^τ. On condition that φ^τ is

symplectic, $\det(\mathbf{D}\varphi^{-\tau})$ equals one. By integrating out the momenta we easily find that

$$\begin{aligned}\mathbf{E}_\vartheta[\mathcal{O}] &= \int \mathcal{O}(x_*)\Big(\mu(x_*,\tilde{u}) + \min\left\{\mu(\varphi^{-\tau}(x_*,\tilde{u})),\mu(x_*,\tilde{u})\right\} - \min\left\{\mu(x_*,\tilde{u}),\mu(\varphi^{-\tau}(x_*,\tilde{u}))\right\}\Big)\tilde{\lambda} \\ &= \int \mathcal{O}(x_*)\,\mu(x_*,\tilde{u})\tilde{\lambda} \\ &= \int \mathcal{O}(x)Q(x)\sqrt{\det G(x)}\,\mathrm{d}x^1\ldots\mathrm{d}x^d\,.\end{aligned}$$

The square root in the very last expression originates from the kinetic energy in the constrained measure (IV.2). Hence we have $m = \sqrt{\det G}$ which is indeed the correct marginal distribution in configuration space. □

So, why does HMC give a symplectic and time-reversible mapping, too? Certainly, the HMC inherits symplecticity from the discrete flow φ^τ. Time-reversibility can be verified by checking *detailed balance*:

$$\mu(x,u)a(x,\tilde{x}) = \mu(x,u)\min\left\{1,\,\frac{\mu(\tilde{x},\tilde{u})}{\mu(x,u)}\right\} = \min\left\{\mu(\tilde{x},\tilde{u}),\mu(x,u)\right\}\,.$$

The last equation is symmetric with respect to the initial and propagated variables. Hence the detailed balance condition is satisfied, for $\mu(x,u)a(x,\tilde{x}) = \mu(\tilde{x},\tilde{u})a(\tilde{x},x)$.

Remark 2. *The invariance of the Liouville measure is a consequence of the symplecticity of the flow map φ^τ. Bear in mind that the Liouville form is related to the symplectic form by (IV.3), where the right hand side contains a number of exterior products. Since the pull-back of an exterior product is the exterior product of the pull-back, i.e., $(\alpha\wedge\beta)\circ\varphi^{-\tau} = (\alpha\circ\varphi^{-\tau})\wedge(\beta\circ\varphi^{-\tau})$, we can conclude from symplecticity of the flow map, $\omega\circ\varphi^{-\tau} = \omega$ that the constrained Liouville form is conserved. Therefore $\tilde{\lambda} = \lambda$ holds true throughout the proof.*

Building averages from constrained simulation

We shall try to establish a relation between the constrained probability density μ_Σ living on the constrained phase space bundle $T^*\Sigma$ and the conditional density μ that is defined on $T^*\mathbf{R}^n|\Sigma = \Sigma\times\mathbf{R}^n$. Let E be the total energy of the molecular Lagrangian (II.6), and let $H = E\circ\mathbb{F}L^{-1}$ denote the corresponding Hamiltonian. In coordinates the Gibbs density $\mu\propto\exp(-\beta H)$ restricted to Σ is given by

$$\mu(x,w) = Z^{-1}\exp(-\beta/2\left\langle g(x,0)^{-1}w,w\right\rangle)\exp(-\beta U(\sigma(x))$$

with $w=(u,v)\in\mathbf{R}^n$ denoting the conjugate momentum to $z=(x,y)$ on Σ. We can exploit that the inverse metric tensor at $y=0$ takes the simple form $g^{-1} = G^{-1}\otimes\mathbf{1}$, such that the conditional density reads

$$\mu(x,u,v) = Z^{-1}\exp(-\beta/2\left\langle G(x)^{-1}u,u\right\rangle)\exp(-\beta/2\left\langle v,v\right\rangle)\exp(-\beta U(\sigma(x)) \tag{IV.5}$$

with the normal momenta $v_i = \partial L/\partial\dot{y}^i$. A direct consequence is

Lemma 1. *Let $\mu_\Sigma : T^*\Sigma\to\mathbf{R}$ be the canonical invariant density of the constrained dynamics subject to $c=c'$ and $\dot{c}=0$, and let $\mu:\Sigma\times\mathbf{R}^n\to\mathbf{R}$ denote the density on the submanifold Σ. Then up to a constant μ_Σ is simply the restriction of μ to the constrained phase space bundle, $\mu_\Sigma = \mu|_{T^*\Sigma}$.*

Proof. The assertion can be verified by a direct calculation in coordinates: clearly we can neglect the normalization above, that is, we fix $Z=1$ in (IV.5). Consulting the equations (II.5) and (II.6) it follows immediately that $v_i = \partial L/\partial\dot{y}^i = \dot{y}^i$ whenever a curve hits Σ. Hence we can impose the constraint $\dot{c}=0$ just by setting $v=\dot{y}=0$ in the density. Due to the relation (II.4) and the full-rank assumption on the Jacobian $\mathbf{D}c(\sigma(x))$ this is completely equivalent. Consequently $\mu_\Sigma(x,u) = \mu(x,u,0)$ which proves the Lemma. □

This relation will turn out to be very useful for the HMC algorithm because it tells us how to generate constrained random momenta from a full phase space density, for instead of restricing the probability distributions we can just as well restrict the corresponding random variables. Also note that up to a constant factor the marginal distributions of $\mu_\Sigma(x,u)$ and $\mu(x,u,v)$ in configuration space are identical to each other which can be seen by integrating out the normal momenta in (IV.5). Therefore it is a straight consequence that we can compute the conditional average of a position dependent observable simply by taking the average with respect to the constrained marginal density as it pops out of the HMC scheme. This is different from the common blue moon method which is built on the Fixman Theorem [32], and that requires appropriate reweighting of the distributions. We shall comment on this point at the end of this paper.

V. ALGORITHMIC REALIZATION

In order to implement the proposed constraint algorithm, we switch back to a representation of the equations of motion that is more convenient for practical purposes; we shall also drop the mass scaling assumption for the sake of practicability. We define the molecular Lagrangian as

$$L(q,v) = \frac{1}{2}\langle Mv, v\rangle - U(q)\,.$$

giving rise to the usual Newtonian equations of motion. Introducing a family of Lagrangians $L_\lambda = L - \lambda^i c_i(q)$ we can deduce the constrained Euler-Lagrange equations

$$\ddot{q} = -\mathrm{grad}\, U(q) - \lambda^T \mathbf{D}c(q) \quad \text{subject to} \quad c(q) = 0\,, \tag{V.1}$$

that can be discretized by means of any symplectic integration scheme [33, 34]. This is standard, and we have moreover shown that the discretized equations of motion together with the hybrid Monte-Carlo procedure will generate the constrained canonical distribution. The Hamiltonian associated with L is easily found to be

$$H(q,p) = \frac{1}{2}\left\langle M^{-1}p, p\right\rangle + U(q) \quad \text{with} \quad \mathbb{F}L:\ p = Mv\,.$$

As the ordinary canonical density is defined by $\mu \propto \exp(-\beta H)$ we can use Lemma 1 which states a simple relation between the constrained and the conditional density in order to express the constrained density in Cartesian coordinates. To this end let $q \in \Sigma$. Since the range of the Jacobian $\mathbf{D}c(q)$ is the normal space $N_q\Sigma$, the point-wise projection onto the constraint tangent space $T_q\Sigma$ is

$$H_q = \mathbf{1} - A^T(AA^T)^{-1}A\,, \quad A(q) = \mathbf{D}c(q) \in R^{m\times n}$$

such that the corresponding projection onto the cotangent plane $T_q^*\Sigma$ is provided by $H_q^* = MH_qM^{-1}$. Consequently for each $q \in \Sigma$ we obtain the constrained density by constraining the momenta according to [10]

$$\mu_\Sigma(q,p) = Z^{-1}\exp(-\beta/2\left\langle H_q^* M^{-1} H_q^* p, p\right\rangle)\exp(-\beta U(q))\,.$$

The easiest way to draw momenta from the constrained canonical distribution is to generate a vector P due to the unconstrained canonical distribution $\mu(q,\cdot)$, and then apply the projection onto the constrained cotangent plane which yields a vector P_Σ that is properly distributed according to $\mu_\Sigma(q,\cdot)$. Note that $H_q^*M^{-1}H_q^*$ can be considered as the shape matrix for the constrained momentum distribution in Cartesian coordinates. It is clearly rank-deficient, because the projector H_q^* has rank $n-m$ only, but P_Σ is a vector in $\mathbf{R}^n$. In this way the projection maintains the full dimensionality for the HMC algorithm, and we can stick to the equations of motion (V.1).

We want to replace the ensemble average with respect to the configurational part of μ_Σ by an appropriate time average. This requires that the constrained HMC dynamics is ergodic with regard to this probability distribution. This will be guaranteed by irreducibility and positive recurrence of the underlying Markov chain which will be the case for all practical purposes [35, 36]. Thus once we have generated a discrete trajectory $\{q_1, \ldots, q_N\}$ we can compute the conditional mean force simply by

$$\frac{\partial F}{\partial c_i'} = N^{-1}\sum_{k=1}^{N}\left(\langle n_i(q_k), \mathrm{grad}\, U(q_k)\rangle - \beta^{-1}\mathrm{div}\, n_i(q_k)\right)\|\mathrm{grad}\, c_i(q_k)\|^{-1}\,.$$

If this is done for successively fixed values of c', then the corresponding free energy landscape as a function of c' is recovered after numerical integration of the averaged vector field $\partial F/\partial c'$.

A note on different free energy concepts

We have stressed that there are incommensurable notions of free energy in the literature. The definition which is considered here is used to compute dividing surfaces and transition rates between metastable sets [37, 38]. Nevertheless it may seem a little obscure that many of the references in this paper refer to a concept of free energy that is different from the one displayed here. The reason however is, that it occurs more frequently in the classical MD literature,

since it is the traditional definition. Those authors use a free energy which is naturally connected to the marginal distribution of the reaction coordinate, and which typically reads

$$\exp(-\beta G(c')) = \int_{\mathbf{R}^n \times \mathbf{R}^n} \exp(-\beta H(q,p))\delta(c(q) - c')\,\mathrm{d}q\mathrm{d}p\,.$$

The definition has the unlovely property that it is not gauge invariant under transformations of the reaction coordinate, $h(c(q)) = h(c')$, with a smooth and monotone function $h : \mathbf{R}^m \to \mathbf{R}^m$. This can be seen if the Dirac notation above is brought into the equivalent form

$$\exp(-\beta G(c')) = \int_{\Sigma \times \mathbf{R}^n} \exp(-\beta H)\mathrm{vol}(\mathbf{D}c)^{-1}\mathrm{d}\sigma\,,$$

where $\mathrm{vol}(\mathbf{D}c) = \sqrt{\det(\mathbf{D}c\mathbf{D}c^T)}$ denotes the matrix volume of the rectangular full-rank Jacobian matrix. Assuming again that the c_i form an orthogonal set of reaction coordinates one can perform pretty much the same exercise as in (I.2) to see that $G(c')$ is the potential of a mean force which is the average of a quantity

$$g_c^i = f_c^i + \beta^{-1}\left\langle n_i, \mathrm{grad}\,\|\mathrm{grad}\,c_i\|^{-1}\right\rangle\,.$$

It is not hard to see that the additional term stems from the derivative of the inverse matrix volume. Unfortunately we see no straightforward interpretation of g_c as for the generalized force $f_c = -\partial H/\partial c$. Let us further remark that the orthogonality condition on the reaction coordinates is required only to keep the notation simple; however it is rather a matter of notational convenience than it is necessary to state the results.

Moreover the associated conditional density carries the same gauge dependency due to the matrix volume in the denominator, which in fact is the origin of the weight that appears in the blue moon sampling procedure. This can be easily seen on the level of the local coordinates expressions for the marginal distributions of μ_Σ and μ in the position variables, where the constraint makes no difference at all. However the average has to be taken with respect to a conditional density $\tilde{\mu}$ which is determined by

$$\tilde{\mu}|_{T^*\Sigma} = \mathrm{vol}(\mathbf{D}c)^{-1}\mu|_{T^*\Sigma} = \mathrm{vol}(\mathbf{D}c)^{-1}\mu_\Sigma\,.$$

Since constrained dynamics only generates μ_Σ, one has to insert the inverse matrix volume as a weight when computing averages with respect to $\tilde{\mu}$ as it is done in the classical blue moon ensemble method [15].

---

[1] B. Roux. The calculation of the potential of mean force using computer simulations. *Comp. Phys. Comm.*, 91:275–282, 1995.

[2] J.G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3:300–313, 1935.

[3] M. Sprik and G. Ciccotti. Free energy from constrained molecular dynamics. *J. Chem. Phys.*, 109(18):7737–7744, 1998.

[4] E. Darve, M.A. Wilson, and A. Pohorille. Calculating free energies using a scaled-force molecular dynamics algorithm. *Mol. Sim.*, 28(1-2):113–144, 2002.

[5] T. Mülders, P. Krüger, W. Swegat, and L. Schlitter. Free energy as the potential of mean force. *J. Chem. Phys.*, 104(12):4869–4870, 1996.

[6] 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.

[7] W. E and E. Vanden-Eijnden. Metastability, conformation dynamics, and transition pathways in complex systems. In S. Attinger and P. Koumoutsakos, editors, *Multiscale, Modelling, and Simulation*, pages 35–68. Springer, Berlin, 2004.

[8] J.E. Marsden, R. Montgomery, and T. S. Ratiu. Reduction, symmetry and phases in mechanics. *Memoirs AMS*, 88(436):1–110, 1990.

[9] A. Mielke. *Hamiltonian and Lagrangian Flows on Center Manifolds*, volume 1489 of *Lecture Notes in Mathematics*. Springer, Berlin, 1991.

[10] C. Hartmann and C. Schütte. A geometric approach to free energy calculations. *Comm. Math. Sci.*, 3(1):1–20, 2005.

[11] H. Rubin and P. Ungar. Motion under a strong constraining force. *Comm. Pure Appl. Math.*, 10:65–87, 1957.

[12] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, New York, 1989.

[13] F.A. Bornemann. *Homogenization of Singularly Perturbed Mechanical Systems*, volume 1687 of *Lecture Notes in Mathematics*. Springer, Berlin, 1998.

[14] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 1984.

[15] E.A. Carter, G. Ciccotti, J.T. Hynes, and R. Kapral. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.*, 156(5):472–477, 1989.

[16] W.G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695–1697, 1985.

[17] B.L. Holian and W.G. Hoover. Numerical test of the Liouville equation. *Phys. Rev. A*, 34(5):4229–4239, 1986.
[18] G.J. Martyna, M.L. Klein, and M.E. Tuckerman. Nosé-Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635–2643, 1992.
[19] H.C. Andersen. Molecular dynamics simulations at constant temperature and/or pressure. *J. Chem. Phys.*, 71(4):2384–2393, 1980.
[20] S.D. Bond, B.J. Leimkuhler, and B.B. Laird. The Nosé-Poincaré method for constant temperature molecular dynamics. *J. Comp. Phys.*, 151:114–134, 1999.
[21] R. Abraham, J.E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Springer, New York, 1988.
[22] R. Herbst and I. Froese. Realizing holonomic constraints in classical and quantum mechanics. *Comm. Math. Phys.*, 220:489–535, 2001.
[23] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics und Symmetry*. Springer, New York, 1999.
[24] M.P. do Carmo. *Riemannian Geometry*. Birkhäuser, Boston, 1992.
[25] M. Spivak. *Differential Geometry*, volume 4. Publish or Perish, Boston, 1975.
[26] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry I*. Wiley, New York, 1963.
[27] M. Spivak. *Differential Geometry*, volume 3. Publish or Perish, Boston, 1975.
[28] R. Abraham and J.E. Marsden. *Foundations of Mechanics*. Benjamin/Cummings, Massachusetts, 1978.
[29] C. Schütte, W. Huisinga, and P. Deuflhard. Transfer operator approach to conformational dynamics in biomolecular systems. In B. Fiedler, editor, *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, pages 191–223. Springer, Berlin, 2001.
[30] P.-D. Liu and M. Qian. *Smooth Ergodic Theory of Random Dynamical Systems*, volume 1606 of *Lecture Notes in Mathematics*. Springer, New York, 1995.
[31] J.S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, New York, 2001.
[32] M. Fixman. Classical statistical mechanics of constraints: a theorem and applications to polymers. *PNAS*, 71:3050–3053, 1974.
[33] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer, Berlin, 2002.
[34] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 9:357–514, 2001.
[35] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, New York, 1999.
[36] C. Schütte. *Conformational Dynamics: Modelling, Theory, Algorithm, and Application to Biomolecules*. Habilitation Thesis, Fachbereich Mathematik und Informatik, Freie Universität Berlin, 1998.
[37] H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3:107–115, 1935.
[38] E.P. Wigner. Calculation of the rate of elementary association reactions. *J. Chem. Phys.*, 5:720–725, 1937.