

生体分子におけるレアイベントサンプリング： Onsager-Machlup 作用を用いた計算手法

藤崎 弘士¹

¹ 日本医科大学 医学部 物理学教室

概要

本稿では、生体分子におけるレアイベントの意義、Onsager-Machlup 作用の定義と導出、それを用いたパスサンプリング手法、拡張アンサンブルや再重みつけ法との組み合わせなどについて解説する。また今後の展望についても述べる。

Rare event sampling for biomolecules: Computational approaches using the Onsager-Machlup action

Hiroshi Fujisaki¹

¹Department of Physics, Nippon Medical School

Abstract

Here we describe the meaning of “rare events” for computational biophysics, and explain the definition and derivation of the Onsager-Machlup (OM) action, path sampling methods using the OM action, and how to combine the OM action method with extended ensemble methods and reweighting methods. Finally we describe the future prospects on the computational use of the OM action.

1 はじめに:レアイベントとは何か

レアイベント (**rare event**) とは、ナイーブには「稀に起こる現象」ということだが、この用語を厳密に定義するのは難しい。ただし、次の2種類の意味で使われることが多い¹。

1. 生じる確率が非常に低い現象。時間的なことは考えずに定常的な分布があると仮定して、その裾野に対応する現象。具体的には、地震、テロリストの攻撃、タンパク質の準安定な状態など。

¹ただし、これはあくまで便宜的なものであり、この区別がしづらい場合も多い。また、1.のレアイベントは2.によって引き起こされることも多い（例えば、破壊現象が続くことで地震につながるなど）ので、もっと総合的に考えた方がよいかもしれない。

2. 反応のような山を越える現象。動的な観点を重視しており、その現象が動的に起こる確率それ自体は確率的には低い、その中でももっとも確からしく (most probable) 起こる現象。具体的には、分子系における化学反応、構造変化、相転移、(材料の) 破壊現象など。

1. に関しては数学的な定義は簡単であり、ある事象 i の起こる確率を P_i としたときに、これが指数関数的に小さいようなものである (例えば $P_i \sim 10^{-50}$)。2. に関しては、ある状態 i から別の状態 j に遷移する単位時間当たりの確率を k_{ij} としたときに、これを最大化するような現象と考えることができる。ただし、後者に関してはやや矛盾しているところがある。というのも、レアと言いつつ、most probable な現象のみに興味を持つとしているからである。これは実際、観測にかかるようなロバスタな現象に興味があるからということでもあるが、前者の観点から言うと、least probable な反応に興味をもってもいいかもしれない。しかし、単純に k_{ij} が小さい現象を探しても多分意味はなく、その場合は何らかの拘束条件のもとで考えなければならないだろう。ただし、これらのレアイベントについて一般的に語ろうとするとあまりにも範囲が広がってしまうので、以下では分子系におけるレアイベントのみに限定する。

分子系も、例えば2原子や3原子分子などを除けば、そのポテンシャル関数は非常に複雑であり、いくつもの安定状態や準安定状態が現れる。そこで、上で述べたようなレアイベントの問題意識が自然に現れる。1. に対応するのはエネルギーの高い準安定状態であり、2. に対応するのは、準安定な状態間を結ぶ経路もしくはパス (path) である。また実験的にも1. に関しては検出しにくいということから、2. に関してはその経路の途中の情報が実験的に得にくいということから、計算や理論に対する期待がかけられている。一般的に、分子の状態を計算的に調べるための手段は分子動力学 (molecular dynamics, MD) であるが、MD は一般的にレアイベントを調べるのには不向きである。というのも、MD では安定な状態の周りを軌道が経巡ることが殆どなので、準安定な状態や、経路を十分にサンプルすることが難しいからである。1. の意味でのレアイベントに関しては拡張アンサンブル法 [1] を使ったアプローチが有効であるが、本稿では2. の意味でのレアイベント、つまり、パスを求める問題について解説していこう。

2 パスサンプリングの問題

そこでパスを効率よくサンプルするための手法が数多く考え出された [2]。それは大きく2つの手法に大別される。時間の情報を捨てた経路の情報をサンプルする手法²と、時間の情報も含めて経路の情報をサンプルする手法である。前者としては nudged elastic band 法、ストリング (string) 法などが有名であり、化学反応や構造変化を調べる際にはよく用いられる [2]。特にストリング法は有限温度への拡張が容易であり、自由エネルギー面における最小自由エネルギー経路を求めるために使われる [3]。これらの手法は実装も比較的容易であり、並列計算と相性がよいので、これからも益々使われるようになるだろう。しかし、こういった手法で経路を求めても、時間の情報が抜けているので、ダイナミックな情報はそれ自体からは求まらないという問題がある³。

そこで動的なパス「そのもの」をサンプルする手法がないかということになるが、その代表的な手法が1999年ごろに David Chandler とその共同研究者たちによって提唱された transition path sampling (TPS) である [5, 6]。TPS はパスに対する重みに基づいた「パス空間での」モ

²これは一本のパスを求めることが主目的となるので、パスサーチの手法と呼ばれることも多い。

³ただし、ある状態から別の状態へのレートなどを求める場合は、ストリング法で求めた経路を使って、そこから並列な MD 計算を走らせることで求めることもできる (後述) [4]。

ンテカルロ・ムーブを導入することで、パス空間を大域的にサンプルする手法である。これは非常に一般的な手法であり、非平衡統計力学などを考える上でも出発点になる考え方だが、現実的な状況に適用しようとするとき様々な問題があり、現在でも改良が続けられている。TPS の一般的な review は [5, 6] にあるので、以降ではダイナミクスが overdamped Langevin 方程式によって表される場合の手法についてのみ述べる⁴。

3 Onsager-Machlup (OM) 作用とパスサンプリング

3.1 Onsager-Machlup 作用

以下では、1次元の overdamped Langevin 方程式

$$\frac{dx}{dt} = \frac{1}{\zeta} F(x) + \sqrt{2D} \eta(t) \quad (1)$$

に従うダイナミクスのみを考える。ここで、 $F(x)$ は系統的な力、 $\eta(t)$ はそれ以外のノイズを表し、 $\langle \eta(t)\eta(0) \rangle = \delta(t)$ を満たすものとする。また、 $\langle * \rangle$ はサンプル平均、 $\delta(t)$ はディラックのデルタ関数、 ζ は摩擦定数、 D は拡散定数であり、絶対温度 T のもとでは揺動散逸定理 (平衡状態がボルツマン因子になる) から $D = k_B T / \zeta$ の関係がある。式 (1) はランダム力学系 (random dynamical systems) のもっとも簡単な形である。

このダイナミクスによって生み出されるパス $x(t)$ の重み $P[x(t)]$ は、実は

$$P[x(t)] \propto e^{-S[x(t)]/2D} \quad (2)$$

と書くことができる。ここで現れる汎関数 $S[x(t)]$ は以下で定義される。

$$S[x(t)] = \frac{1}{2} \int_0^t dt \left(\frac{dx}{dt} - \frac{F(x)}{\zeta} \right)^2 \quad (3)$$

これが Onsager-Machlup 作用 (action) である。パスの重みの式から、パスとして実現されやすいのは、OM 作用が小さいパスであることが分かる。またもっとも確からしい (most probable) パスは、 $\delta S[x(t)] = 0$ という変分問題を解くことから出てくるパスである。こういった考えに基づいて、パスサーチやパスサンプリングのアルゴリズムを考えることができる [2]。以上は1次元の話であったが、これを形式的に多次元に拡張するのは容易である。

3.2 OM 作用の導出

3.2.1 シンプルな導出

この1次元系に対する OM 作用を導くにはいくつかの手法があるが、もっともシンプルなものは overdamped Langevin を離散化して考えるものである。つまり、式 (1) を時間間隔 Δt で離散化して、

$$x_{i+1} = x_i + \frac{F(x_i)}{\zeta} \Delta t + \sqrt{2D\Delta t} R_i \quad (4)$$

⁴ただし、underdamped Langevin 方程式に対しても似たような議論は可能である。

とする。ここで $x_i = x(i\Delta t)$ であり、 R_i は無相関で平均0、分散1のガウス分布から取り出される乱数である。これはもともと単純な Euler のスキームであり、数値計算でよく用いられるものである。

さて、このスキームを繰り返し用いることで、時系列 x_1, x_2, \dots, x_N を生成することができる。 R_i が揺らいでいることから、その時系列にはある分布が付随することになる。 R_i がガウス分布であり、 $p(R_i) \propto \exp\{-R_i^2/2\}$ という形をしているので、パスに対する重みは explicit には

$$p(R_1, R_2, \dots, R_N) \propto \exp\left\{-\sum_{i=1}^N R_i^2/2\right\} \quad (5)$$

となる。Euler のスキームを R_i について逆に解いて、上の式に入れると、

$$p(R_1, R_2, \dots, R_N) \propto \exp\left\{-\frac{\Delta t}{4D} \sum_{i=0}^{N-1} \left(\frac{x_{i+1} - x_i}{\Delta t} - \frac{F(x_i)}{\zeta}\right)^2\right\} \quad (6)$$

となる。これから離散化された OM 作用は

$$S[\{x_i\}] = \frac{\Delta t}{2} \sum_{i=0}^{N-1} \left(\frac{x_{i+1} - x_i}{\Delta t} - \frac{F(x_i)}{\zeta}\right)^2 \quad (7)$$

となる。この連続極限から OM 作用 (3) が出てくる [7]。ただし、数値計算などではもちろん離散化されたものを用いる。

3.2.2 Fokker-Planck 方程式の経路積分表示を用いた導出

別の導出法は、overdamped Langevin 方程式と以下の Fokker-Planck 方程式

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 P(x, t)}{\partial x^2} - \frac{\partial}{\partial x} \left(\frac{F(x)}{\zeta} P(x, t) \right) \quad (8)$$

が等価であることを使うものである [8]。ここで

$$P(x, t) = \exp(-\beta U/2) \phi(x, t) \quad (9)$$

という変換を行うと、 $\phi(x, t)$ に対する以下の方程式が得られる。

$$-\frac{\partial}{\partial t} \phi(x, t) = \left[-D \frac{\partial^2}{\partial x^2} + V_{\text{eff}}(x) \right] \phi(x, t), \quad (10)$$

$$V_{\text{eff}}(x) = \frac{1}{4D} \left[\left(\frac{F(x)}{\zeta} \right)^2 + \frac{2D}{\zeta} \frac{dF}{dx} \right]. \quad (11)$$

式(10)は虚時間でのシュレーディンガー方程式と等価である。シュレーディンガー方程式のプロパゲーターには経路積分表示が知られており、それは以下のように書ける。

$$P(x_N, t; x_0, 0) = \int_{x(0)=x_0}^{x(t)=x_N} Dx(s) \exp\{-S[x(s)]/2D\} \quad (12)$$

ここで $\int_{x(0)=x_0}^{x(t)=x_N} Dx(s)$ は $x(0) = x_0$ から $x(t) = x_N$ に至る経路全体の積分を表し、 $Dx(s)$ はその経路積分の測度である。上式は時刻 $t = 0$ で x_0 にいたとしたときに、 t 秒後に x_N に見出される確率（条件付き確率）としての意味を持つ。

ただし、この場合の（離散的な）作用は

$$S[\{x_i\}] = \frac{\Delta U}{\zeta} + \frac{\Delta t}{2} \sum_{i=0}^{N-1} \left[\left(\frac{x_{i+1} - x_i}{\Delta t} \right)^2 + \left(\frac{F(x_i)}{\zeta} \right)^2 + \frac{2D}{\zeta} \frac{dF(x)}{dx} \Big|_{x=x_i} \right] \quad (13)$$

という形になっており、シンプルな導出のときのもの (7) と形が異なる。しかし、通常は x_i が確率変数であると考えており、その状況においてはこれらの作用が異なることはない。この2つの作用の違いについては Adib の論文 [9] を参照。

3.3 OM 作用を用いたパスサンプリング

では OM 作用を使ったパスサンプリングの問題を考えよう。一番簡単には、素朴なモンテカルロ・ムーブを考えてみればよい。つまり、 $x_i \rightarrow x_i + \Delta x$ と動かしたときの作用の差 ΔS を計算し、それが負であればアクセプト、正の場合は $P_{\text{acc}} = \exp(-\Delta S/2D)$ の確率でアクセプトすればよい。ただし、これはシンプルではあるが、局所的なムーブなので、あまりパスが全体として動かないことが予想される。もっと大域的なムーブのほうが好ましい⁵。

そこで作用のある種のハミルトニアンと考えると、パス全体を動かすスキームを考えよう。仮想的に以下のダイナミクスを導入する。

$$\frac{dx_i}{d\tau} = -L_i \frac{\partial}{\partial x} \left(\beta_T \tilde{S}[\{x_i\}] \right) + \sqrt{2L_i} \eta(\tau) \quad (14)$$

ここで τ は仮想的な時間、 L_i は適当な定数である。 $\tilde{S}[\{x_i\}]$ は

$$\tilde{S}[\{x_i\}] \equiv \frac{S[\{x_i\}]}{2\beta_T D} \quad (15)$$

で定義されるものでエネルギーの次元をもつ。式 (14) は式 (1) の場合と同様に、

$$P[\{x_i\}] \propto \exp\{-\beta_T \tilde{S}[\{x_i\}]\} \quad (16)$$

を定常解としてもつ。よって、この仮想的なダイナミクスを長時間計算することでパス空間をサンプリングできるわけである。また、以下で定義する係数 u_{ik} を使って、この経路を

$$x_i = x_i^{(0)} + \sum_{k=2}^{N-1} u_{ik} q_k, \quad (17)$$

$$u_{ik} = \sqrt{\frac{2}{N-1}} \sin\left(\frac{\pi(i-1)(k-1)}{N-1}\right) \quad (18)$$

と線形変換することで、実空間 x_i でなく、フーリエ空間 q_k でのパスサンプリングを考えることもできる [10]。これも系によっては有効に働くだろう。

⁵ もちろん overdamped Langevin に対して、TPS のスキームを考えることもできる。その場合は shooting と shifting というパス空間でのムーブを実装すればよい。これは遷移する軌道から一点を取り出して、そこから MD を走らせて、各ベイスンに到達するかどうかを判別するアルゴリズムになる。詳細に関しては [5, 6] を参照。

ここで重要なことは、元はパスの重みという非平衡の量を考えていたのに対して、ここでは \bar{S} というハミルトニアン¹の平衡状態を計算すればよい、つまり、非平衡系の問題を平衡系の問題にマッピングできたということである。このマッピングの利点として、平衡系では効率的なサンプリング手法が数多く開発されているので、それらを用いることができる。例えば、平衡系の配置空間をサンプルするための拡張アンサンブル (extended ensemble) 法 [1] がその例である。また、配置空間のサンプリングにおいて、重みの小さい配置の情報を再重みつけ (reweighting) [7] から求めることもよく行われるが、パスの重みを使えば、パス空間でそれを行うこともできる。以下でそういった試みについて述べる。

3.4 OM 作用と拡張アンサンブル法

3.4.1 温度レプリカ交換法

拡張アンサンブル法 [1, 7] はエルゴード性の破れた系を扱うために考案された手法であり、いくつものコピー (レプリカ) を用いて同時に計算するところに特色がある。よって、並列計算に向いており、温度レプリカ交換 (temperature replica exchange) 法のような標準的な拡張アンサンブル法であれば、現在の大部分の MD コードには実装されている。

MD 計算の目的の一つは、ある温度 (室温であることが多い) での自由エネルギー地形の計算である。しかし、生体分子であれば、その地形が凸凹しており、また遷移状態のようなエネルギーの高いところには、室温のダイナミクスでは到達しにくい。よって、その遷移は非常にレアなもの (レアイベント) になりうる。しかし、高温であれば、熱揺らぎによって、低温のときの遷移状態を何度も訪れているかもしれない。ただし、そのときは自由エネルギー地形そのものが変わっている。温度レプリカ交換法は、大雑把に言うと、この高温 MD の性質を用いて自由エネルギー地形を探索し、それを所望の温度の自由エネルギー面に焼き直すという手法である。

ただし、そのためには異なる温度の配置のコピーを用意する必要がある。それぞれをまずは独立に MD 計算するが、ある頻度で、その温度 (隣接する温度) を交換する。交換をアクセプトする確率を

$$P_{\text{acc}} = \min(1, \exp\{-\Delta\}), \quad (19)$$

$$\Delta \equiv (\beta_i - \beta_j)[U(x_i) - U(x_j)] \quad (20)$$

で定めると、各温度でのアンサンブルがちゃんと各温度でのボルツマン因子 $e^{-\beta U(x)}$ になっているというのが温度レプリカ交換法のポイントである。よって、ある温度のアンサンブルが知りたければ、温度レプリカ交換法の計算が終わった後で、その温度を固定してアンサンブルを集めればよい [1, 7]。

これはスピン系で導入 [11] されてから、生体分子に関しても導入 [12, 13] されたが、非常に強力な方法であり、実装も難しくはないので非常に広く使われている。パス空間に関して、OM 作用をハミルトニアンとみなすことができるので、温度レプリカ交換法を併用することに原理的な問題はない。その結果として、エルゴード性が破れるようなパス空間においても有効なパスサンプリングができることが示されている [10]。ただし、温度レプリカ交換法の場合は、温度の比 T_{i+1}/T_i が $1+c/\sqrt{N}$ のような数で bound されてしまう。ここで N は系の自由度であり、 c は定数である。よって、 N が大きくなると温度の比を大きくとれない、よって、温度レプリカの数を非常に多くとらなければならないという問題が生じる。

3.4.2 ハミルトニアンレプリカ交換法

そこで温度以外のパラメータを交換することを考えてみよう。例えば、ハミルトニアンに入っているパラメータを交換するというアイデアがハミルトニアンレプリカ交換 (Hamiltonian replica exchange) 法である [14]。例えば、ポテンシャル関数が

$$U(x, \lambda) = U_0(x) + V(x, \lambda) \quad (21)$$

と与えられている場合、この λ を交換するわけであるが、そのときは

$$P_{\text{acc}} = \min(1, \exp\{-\Delta\}), \quad (22)$$

$$\Delta \equiv \beta[V(x_j, \lambda_i) + V(x_i, \lambda_j) - V(x_i, \lambda_i) - V(x_j, \lambda_j)] \quad (23)$$

のようにアクセプトの確率を定めればよい。この λ が大きいレプリカは大きな相互作用 (力) を持っていることが多いので、適切に V を設計すれば配置空間の所望の領域を効率的にサンプリングできることになる。ただし、その結果得られる配置のアンサンブルは相互作用によってバイアスを受けたものになってしまうので、計算が終わった後に $\lambda = 0$ のアンサンブルだけを取り出せばよい。

森次らはマルチスケールな配置空間のサンプリングにこのハミルトニアンレプリカ交換法を組み込んだ [15]。その基本的な考え方は以下の通りである。まず、知りたい自由度に関してのハミルトニアン H_{MM} を考える。これは例えば、タンパク質の全原子のモデルであり、これを MM 系と呼ぶ⁶。しかし、系が大きくなると、そのボルツマン因子 $e^{-\beta H_{\text{MM}}}$ をサンプリングすることが難しくなり、自由エネルギーの計算などもうまくいかなくなる。これは上で述べたレアイメントの問題が起り、準安定なペイスンの周りでしか配置が動かないからである。そこで、MM 系と似たもっと自由度の小さい系を考え、そのハミルトニアンを H_{CG} で表す。これを CG (coarse grained) 系と呼ぼう。その系の簡単さからレアイメントの問題は軽減されるので、CG 系の配置空間のサンプリングは容易である。

この系を MM 系と結合させる、つまり、全系のハミルトニアンを

$$H = H_{\text{MM}} + H_{\text{CG}} + \lambda H_{\text{MMCG}} \quad (24)$$

とする。この系で λ を交換するようなハミルトニアンレプリカ交換を行うわけである。これが森次らによって考えられた手法であり、**multiscale enhanced sampling (MSES)** と呼ばれる [15]。MSES は天然変性タンパク質 [16] やタンパク質タンパク質相互作用 [17] を調べるのに有効であることが示されているが、OM 作用を使ったパスサンプリングにこの手法を持ち込むことも可能である。その場合には OM 作用における MM 系と CG 系を導入し、それらを

$$S = S_{\text{MM}} + S_{\text{CG}} + \lambda S_{\text{MMCG}} \quad (25)$$

のように結合させればよい。モデルポリマーのパスサンプリングの問題に適用することで、この手法の有効性は確かめられている [18]。

⁶これは molecular mechanics の略であるが、一般的にはより fine な自由度を表す。

3.5 OM 作用と再重みつけ (reweighting)

上のレプリカを用いる手法は比較の実装するのが容易であり、並列計算機が潤沢に使えるような状況であれば有効である。しかし、その計算にはかなりのコア数が必要なので、計算機があまり使えない環境にいる場合や、系が非常に大きくなるような場合には、別の手法を考えておく必要がある。

その一つが再重みつけ (reweighting) を使う手法である。再重みつけを平衡系について説明すると以下のようになる [19]。エネルギーが $E_i^{(0)}$ で与えられている系をサンプルすれば、その結果として $e^{-\beta E_i^{(0)}}$ というボルツマン因子が得られる。ここで、このエネルギー地形をサンプルするのは容易であると仮定しよう。しかし、実際に知りたい系のエネルギー地形は $E_i = E_i^{(0)} + \Delta E_i$ と表せるとする。このエネルギー地形は凸凹しており、サンプリングがしにくい状況を考えてよう。しかし、そんな場合であっても、例えば G という物理量の平均を求める場合、以下のような計算をすればよい。

$$\langle G \rangle = \frac{\langle G_i e^{-\beta \Delta E_i} \rangle_0}{\langle e^{-\beta \Delta E_i} \rangle_0} \quad (26)$$

ここで 0 の添え字のついた平均は、 $E_i^{(0)}$ のエネルギーを使ってサンプリングすることを意味する。このとき $e^{-\beta \Delta E_i} / \langle e^{-\beta \Delta E_i} \rangle_0$ を再重みつけの因子と呼ぶ。この式は E_i を用いた統計平均の式から厳密に導くことができる [19]。

上で述べたように、バスの重みも OM 作用を使えば (overdamped Langevin でモデル化できる場合) ボルツマン因子のように書けるので、この再重みつけを使うことも可能である。Zuckerman と Woolf はこの考えを使うことで、レアイベントのレート計算の効率化が可能であることを示した [20]。その際に、時刻 $t=0$ でベイスン A 内の x_0 にいた軌道が、時刻 t でベイスン B に到達する条件付き確率 $P_B(t_n | x_0, t=0)$ (ここで時間を t_0, t_1, \dots, t_n と離散化した) を以下の経路積分で定義した。

$$P_B(t_n | x_0, t=0) = \int d\xi Q(\xi) h_B(x_n) \quad (27)$$

ここで $h_B(x_n)$ は指示関数であり、 x_n がベイスン B に入っているときは 1、それ以外は 0 となる関数、 $d\xi = dx_0 dx_1 \dots dx_n$ は経路積分の測度を表し、 $Q(\xi)$ は式 (2) で表されるバスの重みである。この計算をナイーブに行うと以下のようになる：通常のダイナミクスを計算し、時刻 t_n でベイスン B に入っているかどうか判別し、入っていたら 1 とする。この計算を多くの初期状態に対して繰り返し、統計平均をとる。これを t_n の関数としてみたときに、レアイベントであれば t_n に線形に依存する領域が出てくることが多いので、その傾きをもって反応レートを定義できる。しかし、レアイベントが生じるような系ではベイスン B に行く軌道が少ないので、この計算の効率が悪くなるかもしれない。

そこで、これを $D(\xi)$ という再重みつけ因子を使って以下のように計算する。

$$P_B(t_n | x_0, t=0) = \int d\xi D(\xi) \frac{Q(\xi)}{D(\xi)} h_B(x_n) \simeq \frac{1}{M} \sum_{\xi} \frac{Q(\xi)}{D(\xi)} h_B(x_n) \quad (28)$$

ここで $D(\xi)$ については具体的に触れないが、遷移状態を容易に越えるようなダイナミクス (つまり、力やポテンシャルを適当に修正する) から計算されるバスの重みとする。すると、元のダ

イナミクスでなく、修正されたダイナミクスを流して、上式の最右辺を計算することから条件付き確率を求めることができる。Zuckerman と Woolf は $D(\xi)$ をうまく選ぶことによって、1次元の系に対してバリアが高いときに数十倍から数百倍効率がよくなることを示した。

このような OM 作用の再重みつけによる計算の効率化は最近も行われている [21, 22] が、基本的には多次元系にいくと（実際、低次元系でも）うまくいかない。これは配置空間の再重みつけの場合と同じ問題が生じるからである：式 (28) の因子 Q/D に入ってくる作用の差は、次元が上がるか時間が長くなると大きくなり、 Q/D が極端に大きくなるか、小さくなってしまふのである。よって、収束が非常に悪くなる。作用の差を小さくするために、できるだけ元と同じダイナミクスで短時間計算すると、普通の MD をやるのと同じになってしまうので元の木阿弥である。再重みつけを多次元系の長時間ダイナミクスに使うにはやる事が多く残されている⁷。

4 まとめと展望

本稿では、生体分子のような複雑な系におけるレアイベントを調べる上では、通常の MD シミュレーションだけでなく、作用に基づく方法が必要であることを述べた。その際にはパスサンプリングが主要な問題として現れてくるので、それを効率よく行うためのいくつかの手法（レプリカ交換法、再重みつけ法など）について解説した。ここでは *overdamped Langevin* ダイナミクスに対応する OM 作用のみを取り上げたが、異なるダイナミクスに対応する他の作用もあり、それに関しては Elber らのレビューに詳しい [23]。また、作用は時間積分として定義されているが、これをパスに沿った線積分として定義する手法もあり、その場合は変数変換によって作用の形は煩雑になる [23]。

ただし、パスサンプリングを生体分子のような多次元の複雑な系に対して行うには、OM 作用を使うにしろ他の手法を使うにしろ、現在の計算機のパワーを用いても容易ではない。それは配置空間が凸凹した多次元空間なのに加えて、時間の情報も保たねばならないからである。そこで現在まず行われているのは、ある意味での最適経路 (**optimized path**) を一本求めて、そこから議論をスタートさせるということである。そのためには、OM 作用を最適化して一本パスを求めることが考えられる。これは Elber [23] や Faccioli [24] らによって進められているアプローチである。ただし、OM 作用の計算が煩雑だったり、離散化誤差の制御が難しいために、これを厳密に行うことは難しい。そこで、Vanden-Eijnden らによって考えられているのは、まず最小自由エネルギー経路 (**minimum free energy path, MFEP**) を求めて、それから動的な効果を取り込んで行こうというアプローチである [25]。これは OM 作用や TPS を使う手法よりは簡便なので、最近をよく使われるようになってきている。反応レートを計算するためには、MFEP を適当に区切って、その各領域で MD を走らせ、そこからマルコフ状態モデル (**Markov State Model**) やマイルストーン法 (**milestoning**) と呼ばれる手法を用いる [4]。ただし、これらの手法は様々な問題を抱えており、それらに関しては [26] を参照。

では OM 作用を用いた手法は最終的に使わなくてよいのだろうか？ これに関しては少なくとも 2 つの研究の方向性があると考えている。1 つ目は OM 作用を使ったパスの重みを時系列に関する尤度 (**likelihood**) と考えて、パラメータ推定するという方向性であり、これは生体分子に限らずさまざまな系のダイナミクスに対しても使えるので有効である [27, 28]。2 つ目の方向性は、非平衡定常状態 (**nonequilibrium steady states**) を計算するためのツールとして使うという

⁷パスに対する *simulated tempering* [1] を行う可能性も考えられる。

ことである。OM 作用から Jarzynski 等式や揺らぎの定理などが導出できることはよく知られている [7, 29] が、非平衡定常状態に対する議論も可能である [30]。細胞内では非平衡定常状態が実現されている場合が想定されるので、そのような状況下でいかに OM 作用を使うかはこれからの課題であろう。

志賀基之氏（原子力研究開発機構）と山下雄史氏（東京大学）には本稿に関して有益なコメントを頂いたことに感謝します。

参考文献

- [1] Y. Iba, *Int. J. Mod. Phys. C* **12**, 623-656 (2001).
- [2] 藤崎弘士, 日本医科大学基礎科学紀要, **40**, 83-98 (2010).
- [3] Y. Matsunaga, H. Fujisaki, T. Terada, T. Furuta, K. Moritsugu, and A. Kidera, *PLoS Comput. Biol.* **8**, e1002555 (2012).
- [4] 藤崎弘士, 分子シミュレーション研究会誌アンサンブル, **17**, 175-180 (2015).
- [5] P.G. Bolhuis and C. Dellago, *Eur. Phys. J. Special Topics* **224**, 2409-2427 (2015).
- [6] 藤崎弘士, 分子シミュレーション研究会誌アンサンブル, **16**, 8-15 (2014).
- [7] Daniel M. Zuckerman, *Statistical Physics of Biomolecules: An Introduction*, CRC Press (2010); 邦訳は, 藤崎弘士・藤崎百合訳, 生体分子の統計力学入門, 共立出版 (2014).
- [8] F.W. Wiegell, *Introduction to Path-Integral Methods in Physics and Polymer Science*, World Scientific, Singapore (1986).
- [9] A.B. Adib, *J. Phys. Chem. B* **112**, 5910 (2008).
- [10] H. Fujisaki, M. Shiga, and A. Kidera, *J. Chem. Phys.* **132**, 134101 (2010).
- [11] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.* **65**, 1604 (1996).
- [12] U.H.E. Hansmann, *Chem. Phys. Lett.* **281**, 140-150 (1997).
- [13] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141-151 (1999).
- [14] H. Fukunishi, O. Watanabe, and S. Takada, *J. Chem. Phys.* **116**, 9058 (2002).
- [15] K. Moritsugu, T. Terada, and A. Kidera, *J. Chem. Phys.* **133**, 224105 (2010).
- [16] K. Moritsugu, T. Terada, and A. Kidera, *J. Am. Chem. Soc.* **134**, 7094-7101 (2012).
- [17] K. Moritsugu, T. Terada, and A. Kidera, *PLoS Comput. Biol.* **10**, e1003901 (2014).
- [18] H. Fujisaki, M. Shiga, K. Moritsugu, and A. Kidera, *J. Chem. Phys.* **139**, 054117 (2013).
- [19] David Chandler, *Introduction to Modern Statistical Mechanics*, Oxford Univ. Press, New York (1987).

- [20] D.M. Zuckerman and T.B. Woolf, *Phys. Rev. E* **63**, 016702 (2000).
- [21] J. Shin, T. Ikonen, M.D. Khandkar, T. Ala-Nissila, and W. Sung, *J. Chem. Phys.* **133**, 184902 (2010).
- [22] G. Grazioli and I. Andricioaei, arXiv:1511.00044v1.
- [23] R. Elber, A. Cardenas, A. Ghosh, and H.A. Stern, *Adv. Chem. Phys.* **126**, 93 (2003).
- [24] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, *Phys. Rev. Lett.* **97**, 108101 (2006).
- [25] E. Vanden-Eijnden, *J. Comput. Chem.* **30**, 1737 (2009).
- [26] 藤崎弘士, 分子シミュレーション研究会誌アンサンプル, **18**, 39-44 (2016).
- [27] 藤崎弘士, 統計数理, **62**, 301-311 (2014).
- [28] J. Comer, J.C. Gumbart, J. Henin, T. Lelievre, A. Pohorille, and C. Chipot, *J. Phys. Chem. B* **119**, 1129-1151 (2015).
- [29] R.D. Astumian, *Am. J. Phys.* **74**, 683-688 (2006).
- [30] T. Taniguchi and E.G.D. Cohen, *J. Stat. Phys.* **130**, 633-667 (2008).