Mathematical analysis of an Adaptive Biasing Potential method for diffusions

Charles-Edouard Bréhier

Joint work with
Michel Benaïm (Neuchâtel, Switzerland)

CNRS & Université Lyon 1, Institut Camille Jordan (France)

Stochastic Sampling and Accelerated Time Dynamics on Multidimensional Surfaces

Plan of the talk

Motivations: setting

Goal: estimating averages

$$\int \phi d\mu_\beta = \frac{1}{Z(\beta)} \int_{\mathbb{T}^d} \phi(x) e^{-\beta V(x)} dx,$$

where $V : \mathbb{T}^d \to \mathbb{R}$. From now on $\beta = 1$.

The probability distribution μ is ergodic for the overdamped Langevin dynamics:

$$dX_t^0 = -\nabla V(X_t^0)dt + \sqrt{2}dW_t.$$

Difficulty: slow convergence of temporal averages (metastability)

$$\frac{1}{t} \int_0^t \phi(X_s) ds \underset{t\to\infty}{\to} \int \phi d\mu.$$

Context

Use of an importance sampling strategy: change of the drift coefficient in the dynamics.

A nice review

B. M. Dickson.
Survey of adaptive biasing potentials: comparisons and outlook.
Current Opinion in Structural Biology, 2017.

Related to the following techniques: umbrella sampling

S. Marsili, A. Barducci, R. Chelli, P. Procacci, and V. Schettino.
Self-healing umbrella sampling: a non-equilibrium approach for quantitative free energy calculations.
The Journal of Physical Chemistry B, 2006.

G. Fort, B. Jourdain, T. Lelièvre, and G. Stoltz.
Self-healing umbrella sampling: convergence and efficiency.
Stat. Comput., 2017.

Context

Use of an importance sampling strategy: change of the drift coefficient in the dynamics.

A nice review

B. M. Dickson.
Survey of adaptive biasing potentials: comparisons and outlook.
Current Opinion in Structural Biology, 2017.

Related to the following techniques: metadynamics

A. Laio and M. Parrinello.
Escaping free-energy minima.
Proceedings of the National Academy of Sciences, 2002.

A. Barducci, G. Bussi, and M. Parrinello.
Well-tempered metadynamics: a smoothly converging and tunable free-energy method.
Physical review letters, 2008.

Context

Use of an importance sampling strategy: change of the drift coefficient in the dynamics.

A nice review

B. M. Dickson.
Survey of adaptive biasing potentials: comparisons and outlook.
Current Opinion in Structural Biology, 2017.

Related to the following techniques: Wang-Landau

F. Wang and D. Landau.
Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram.
Physical Review E, 2001.

F. Wang and D. Landau.
Efficient, multiple-range random walk algorithm to calculate the density of states.
Physical review letters, 2001.

G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz.
Convergence of the Wang-Landau algorithm.
Math. Comp., 2015.

Biasing methods for diffusions

- Adaptive Biasing Force (ABF)

$$dX_t = \big(-\nabla V(X_t) + F_t(X_t)\big)dt + \sqrt{2}dW_t.$$

- Adaptive Biasing Potential (ABP)

$$dX_t = \big(-\nabla V(X_t) + \nabla \mathcal{V}_t(X_t)\big)dt + \sqrt{2}dW_t.$$

References on ABF

J. Comer, J. C. Gumbart, J. Hénin, T. Lelièvre, A. Pohorille, and C. Chipot.
The adaptive biasing force method: Everything you always wanted to know but were afraid to ask.
The Journal of Physical Chemistry B, 2014.

T. Lelièvre, M. Rousset, and G. Stoltz.
Long-time convergence of an adaptive biasing force method.
Nonlinearity, 2008.

The bias F_t depends on the law $\mathcal{L}(X_t)$.

Biasing methods for diffusions

- Adaptive Biasing Force (ABF)

$$dX_t = \big(-\nabla V(X_t) + F_t(X_t)\big)dt + \sqrt{2}dW_t.$$

- Adaptive Biasing Potential (ABP)

$$dX_t = \big(-\nabla V(X_t) + \nabla \mathcal{V}_t(X_t)\big)dt + \sqrt{2}dW_t.$$

We study a continuous time version of the ABP method proposed in

B. Dickson, F. Legoll, T. Lelièvre, G. Stoltz, and P. Fleurat-Lessard.
Free energy calculations: An efficient adaptive biasing potential method.
J. Phys. Chem. B, 2010.

The bias $\mathcal{V}_t$ depends on the past trajectory of the process, X_r for $0 \leq r \leq t$.

Plan of the talk

Biasing the potential

Reaction coordinate: $\xi : x \in \mathbb{T}^d \mapsto x_1 \in \mathbb{T}$.

Biasing: for $A : \mathbb{T} \to \mathbb{R}$,

$$dX_t^A = -\nabla\big(V - A \circ \xi\big)(X_t^A)dt + \sqrt{2}dW_t.$$

Ergodic invariant law is modified $\propto e^{-V(x)+A(\xi(x))}$.

Why it is useful:

- weighted averages converge to $\int \phi d\mu$ (consistency)
- convergence is faster when A is well-chosen: free energy
- **can be made adaptive**

The free energy

$A_\star : \mathbb{T} \to \mathbb{R}$ is given by

$$e^{-A_\star(z)} = \int_{\mathbb{T}^{d-1}} Z(1)^{-1} e^{-V(z,x_2,\ldots,x_d)} dx_2 \ldots dx_d.$$

Acceleration of the sampling = removing free energy barriers.

Choosing $A = A_\star$: **flat histogram** for $\xi(X_t^{A_\star})$

$$\frac{1}{t}\int_0^t \delta_{\xi(X_s^{A_\star})} ds \underset{t\to\infty}{\to} dz.$$

Adaptive algorithm: $A_t \underset{t\to\infty}{\to} A_\infty \approx A_\star$.

The Adaptive Biasing Potential method

Two unknowns: X_t and A_t.

Dynamics for X_t:

$$dX_t = -\nabla\big(V - A_t \circ \xi\big)(X_t)dt + \sqrt{2}dW(t).$$

Computation of the bias A_t: convolution of a kernel function K

$$e^{-A_t(z)} = \int_{\mathbb{T}^d} K\big(z, \xi(x)\big)\overline{\mu}_t(dx),$$

with the weighted empirical averages

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-A_r \circ \xi(X_r)} \delta_{X_r} dr}{1 + \int_0^t e^{-A_r \circ \xi(X_r)} dr}.$$

Convergence results

$$dX_t = -\nabla\big(V - A_t \circ \xi\big)(X_t)dt + \sqrt{2}dW(t) \quad , \; e^{-A_t(z)} = \int K\big(z, \xi(x)\big)\overline{\mu}_t(dx)$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-A_r\circ\xi(X_r)}\delta_{X_r}dr}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr}.$$

Theorem (Benaïm-B.)

Almost surely,

$$\overline{\mu}_t \underset{t\to\infty}{\to} \mu \quad , \quad A_t \underset{t\to\infty}{\to} A_\infty,$$

with $\mu(dx) = e^{-V(x)}dx$ *and* $e^{-A_\infty(z)} = \int_{\mathbb{T}^d} K(z, \xi(\cdot))d\mu.$

Convergence results

$$dX_t = -\nabla\big(V - A_t \circ \xi\big)(X_t)dt + \sqrt{2}dW(t) \quad , \; e^{-A_t(z)} = \int K\big(z, \xi(x)\big)\overline{\mu}_t(dx)$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-A_r \circ \xi(X_r)} \delta_{X_r} dr}{1 + \int_0^t e^{-A_r \circ \xi(X_r)} dr}.$$

Theorem (Benaïm-B.)

Almost surely,

$$\overline{\mu}_t \underset{t \to \infty}{\to} \mu \quad , \quad A_t \underset{t \to \infty}{\to} A_\infty,$$

with $\mu(dx) = e^{-V(x)}dx$ *and* $e^{-A_\infty(z)} = \int_{\mathbb{T}^d} K(z, \xi(\cdot))d\mu$.

- consistent with non-adaptive versions ($A_t = A \; \forall \; t \geq 0$)
- $A_\infty \neq A_\star$ due to the kernel function.

Comments on the efficiency

- Convergence of histograms: almost surely

$$\frac{1}{t}\int_0^t \delta_{\xi(X_s)} ds \underset{t\to\infty}{\to} e^{A_\infty(z)-A_\star(z)} dz.$$

- Asymptotic variance: same as in non-adaptive version with $A = A_\infty$.

The choice of the kernel function K is important.

Assumption

- $K : \mathbb{T} \times \mathbb{T} \to (0,\infty)$ *is of class* $\mathcal{C}^\infty$ *and positive.*
- $\int K(z,\cdot)dz = 1.$

Examples: $K(z,\zeta) \propto e^{-\frac{|z-\zeta|^2}{2\delta}}$ vs. $K(z,\zeta) = e^{-A(z)}$.

Plan of the talk

1. Motivations

2. The Adaptive Biasing Potential method

3. Consistency and elements of proof
   - Self-interacting diffusions
   - Analysis of the ABP method

4. Extensions
   - Examples
   - Abstract framework

Role of the weighted empirical measures $\overline{\mu}_t$

$$dX_t = -\nabla\big(V - A_t \circ \xi\big)(X_t)dt + \sqrt{2}dW(t),$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-A_r\circ\xi(X_r)}\delta_{X_r}dr}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr},$$

$$\exp\big(-A_t(z)\big) = \int K\big(z, \xi(x)\big)\overline{\mu}_t(dx).$$

Considering $A_t = \mathbf{A}(\overline{\mu}_t)$, may be interpreted as a self-interacting diffusion, with unknowns X_t and $\overline{\mu}_t$

$$dX_t = -\nabla\big(V - \mathbf{A}(\overline{\mu}_t)\circ\xi\big)(X_t)dt + \sqrt{2}dW(t),$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-\mathbf{A}(\overline{\mu}_r)\circ\xi(X_r)}\delta_{X_r}dr}{1 + \int_0^t e^{-\mathbf{A}(\overline{\mu}_r)\circ\xi(X_r)}dr}$$

Self-interacting diffusions

Classical formulation:

$$dY_t = -\nabla V(Y_t, \nu_t)dt + \sqrt{2}dW_t$$

with

$$V(y,\nu) = \int V(y,\cdot)d\nu \quad , \; \nu_t = \frac{1}{1+t}\big(\nu_0 + \int_0^t \delta_{Y_r} dr\big).$$

M. Benaïm, M. Ledoux, and O. Raimond.
Self-interacting diffusions.
Probab. Theory Related Fields, 2002.

Main differences:

- type of the coupling
- $\overline{\mu}_t$ in ABP is a weighted empirical distribution.

Stochastic approximation: the ODE method

$$dY_t = -\nabla V(Y_t, \nu_t)dt + \sqrt{2}dW_t, \quad \nu_t = \frac{1}{1+t}\Big(\nu_0 + \int_0^t \delta_{Y_r} dr\Big).$$

The empirical distribution solves the random ODE

$$\frac{d\nu_t}{dt} = \frac{1}{1+t}\big(\delta_{Y_t} - \nu_t\big).$$

Asymptotic time-scale separation: slow-fast system ($t \to \infty$).

A. Benveniste, M. Métivier, and P. Priouret.
Adaptive algorithms and stochastic approximations.
Springer-Verlag, 1990.

H. J. Kushner and G. G. Yin.
Stochastic approximation and recursive algorithms and applications.
Springer-Verlag, 2003.

Stochastic approximation: the ODE method

$$dY_t = -\nabla V(Y_t, \nu_t)dt + \sqrt{2}dW_t, \quad \frac{d\nu_t}{dt} = \frac{1}{1+t}\big(\delta_{Y_t} - \nu_t\big).$$

Claim

The asymptotic behavior of ν_t is governed by the ODE

$$\frac{d\Gamma_t}{dt} = \Pi(\Gamma_t) - \Gamma_t,$$

where $\Pi(\nu)$ is the unique invariant distribution of

$$dY_t = -\nabla V(Y_t, \nu)dt + \sqrt{2}dW_t$$

Stochastic approximation: the ODE method

Claim

The asymptotic behavior of ν_t is governed by the ODE

$$\frac{d\Gamma_t}{dt} = \Pi(\Gamma_t) - \Gamma_t,$$

where $\Pi(\nu)$ is the unique invariant distribution of

$$dY_t = -\nabla V(Y_t, \nu)dt + \sqrt{2}dW_t$$

A precise statement: in terms of *asymptotic pseudo-trajectories* and of *chain-recurrent sets*.

M. Benaïm.
Dynamics of stochastic approximation algorithms.
In *Séminaire de Probabilités, XXXIII*, 1999.

Analysis of the ABP method (1)

System:

$$dX_t = -\nabla\big(V - A_t \circ \xi\big)(X_t)dt + \sqrt{2}dW(t),$$
$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-A_r\circ\xi(X_r)}\delta_{X_r}dr}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr}$$

The weighted empirical distribution $\overline{\mu}_t$ solves the ODE

$$\frac{d\overline{\mu}_t}{dt} = \frac{e^{-A_t\circ\xi(X_t)}}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr}\big(\delta_{X_t} - \overline{\mu}_t\big)$$

Analysis of the ABP method (1)

The weighted empirical distribution $\overline{\mu}_t$ solves the ODE

$$\frac{d\overline{\mu}_t}{dt} = \frac{e^{-A_t \circ \xi(X_t)}}{1 + \int_0^t e^{-A_r \circ \xi(X_r)} dr} \big(\delta_{X_t} - \overline{\mu}_t\big)$$

Observation

Weights are eliminated by the random change of time variable

$$s = \int_0^t e^{-A_r \circ \xi(X_r)} dr.$$

This may be performed thanks to stability bounds: almost surely

$$m \leq A_t(z) \leq M, \ \forall z \in \mathbb{T}, \ \forall t \geq 0.$$

Analysis of the ABP method (2)

In the new variables

$$Y_s = X_t \ , \ B_s = A_t \ , \ \overline{\nu}_s = \overline{\mu}_t,$$

the dynamics becomes

$$dY_s = -\nabla\big(V - B_s \circ \xi\big)(Y_s) e^{B_s(\xi(Y_s))} ds + \sqrt{2} e^{\frac{1}{2} B_s(\xi(Y_s))} d\tilde{W}(s),$$

$$\overline{\nu}_s = \frac{1}{1+s}\big(\overline{\mu}_0 + \int_0^s \delta_{Y_\sigma} d\sigma\big) \ , \quad e^{-\beta B_s} = \int_{\mathbb{T}} K(z, \xi(x)) \overline{\nu}_s(dx)$$

Now the weights appear in the dynamics.

Analysis of the ABP method (2)

In the new variables

$$Y_s = X_t \ , \ B_s = A_t \ , \ \overline{\nu}_s = \overline{\mu}_t,$$

the dynamics becomes

$$dY_s = -\nabla\big(V - B_s \circ \xi\big)(Y_s) e^{B_s(\xi(Y_s))} ds + \sqrt{2} e^{\frac{1}{2} B_s(\xi(Y_s))} d\tilde{W}(s),$$
$$\overline{\nu}_s = \frac{1}{1+s}\big(\overline{\mu}_0 + \int_0^s \delta_{Y_\sigma} d\sigma\big) \ , \quad e^{-\beta B_s} = \int_{\mathbb{T}} K(z,\xi(x)) \overline{\nu}_s(dx)$$

Now the weights appear in the dynamics.

Asymptotic behavior: almost surely

$$\overline{\mu}_t \underset{t\to\infty}{\to} \mu \quad \Longleftrightarrow \quad \overline{\nu}_s \underset{s\to\infty}{\to} \mu.$$

Analysis of the ABP method (3)

Strategy:

- identify $\nu \mapsto \Pi(\nu) =$ invariant distribution of

$$dY_s = -\nabla\big(V - \mathbf{B}(\nu)\circ\xi\big)(Y_s)e^{\mathbf{B}(\nu)(\xi(Y_s))}ds + \sqrt{2}e^{\frac{1}{2}\mathbf{B}(\nu)(\xi(Y_s))}d\tilde{W}(s),$$

  with $e^{-\mathbf{B}(\nu)(z)} = \int_{\mathbb{T}} K(z,\xi(x))\nu(dz)$?
- study the asymptotic behavior of the ODE

$$\frac{d\Gamma_s}{ds} = \Pi(\Gamma_s) - \Gamma_s,$$

Analysis of the ABP method (3)

Strategy:

- identify $\nu \mapsto \Pi(\nu) =$ invariant distribution of

$$dY_s = -\nabla\big(V - \mathbf{B}(\nu)\circ\xi\big)(Y_s)e^{\mathbf{B}(\nu)(\xi(Y_s))}ds + \sqrt{2}e^{\frac{1}{2}\mathbf{B}(\nu)(\xi(Y_s))}d\tilde{W}(s),$$

with $e^{-\mathbf{B}(\nu)(z)} = \int_{\mathbb{T}} K(z,\xi(x))\nu(dz)$?

- study the asymptotic behavior of the ODE

$$\frac{d\Gamma_s}{ds} = \Pi(\Gamma_s) - \Gamma_s,$$

Nice result: for all ν,

$$\Pi(\nu) = \mu.$$

Then $\Gamma_s = (1 - e^{-s}\mu) + e^{-s}\Gamma_0 \underset{s\to\infty}{\to} \mu$.

Hidden technical details (1)

Generator of the diffusion with fixed B:

$$\mathcal{L}^B = e^{B\circ\xi}\Big(-\langle\nabla(V - B\circ\xi),\nabla\rangle + \Delta\Big).$$

Invariant distribution:

$$\big(\mathcal{L}^B\big)^\star\mu = \Big(-\langle\nabla(V - B\circ\xi),\nabla\rangle + \Delta\Big)^\star\big(e^{B\circ\xi - V}dx\big) = 0.$$

Hidden technical details (1)

Generator of the diffusion with fixed B:

$$\mathcal{L}^B = e^{B\circ\xi}\Big(-\langle\nabla(V - B\circ\xi),\nabla\rangle + \Delta\Big).$$

Invariant distribution:

$$\big(\mathcal{L}^B\big)^\star\mu = \Big(-\langle\nabla(V - B\circ\xi),\nabla\rangle + \Delta\Big)^\star\big(e^{B\circ\xi - V}dx\big) = 0.$$

Poisson equation: for smooth $\varphi : \mathbb{T}^d \to \mathbb{R}$, $\exists!\ \Psi(\cdot,B)$ such that

$$\mathcal{L}^B\Psi(\cdot,B) = \varphi - \int\varphi d\mu, \quad \int\Psi(\cdot,B) = 0.$$

Dependence with respect to B: uniform bounds on $\Psi(\cdot,B)$ and all its derivatives, when controlling B and all its derivatives.

Hidden technical details (2)

Analysis of the error for the original system:

$$\overline{\mu}_t(\varphi) - \mu(\varphi) = \frac{\int_0^t e^{-A_r \circ \xi(X_r)} \big[\varphi(X_r) - \mu(\varphi)\big] dr}{1 + \int_0^t e^{-A_r \circ \xi(X_r)} dr} + \mathrm{o}(1)$$

$$= \frac{\int_0^t e^{-A_r \circ \xi(X_r)} \mathcal{L}^{A_r} \Psi(X_r, A_r) dr}{1 + \int_0^t e^{-A_r \circ \xi(X_r)} dr} + \mathrm{o}(1).$$

Hidden technical details (2)

Analysis of the error for the original system:

$$\overline{\mu}_t(\varphi) - \mu(\varphi) = \frac{\int_0^t e^{-A_r\circ\xi(X_r)}\big[\varphi(X_r) - \mu(\varphi)\big]dr}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr} + \mathrm{o}(1)$$

$$= \frac{\int_0^t e^{-A_r\circ\xi(X_r)}\mathcal{L}^{A_r}\Psi(X_r, A_r)dr}{1 + \int_0^t e^{-A_r\circ\xi(X_r)}dr} + \mathrm{o}(1).$$

Using Itô's formula:

$$\int_0^t e^{-A_r\circ\xi(X_r)}\mathcal{L}^{A_r}\Psi(X_r, A_r)dr = \int_0^t \Big(-\langle\nabla(V - A_r\circ\xi), \nabla\rangle + \Delta\Big)\Psi(X_r, A_r)dr$$

$$= \Psi(X_t, A_t) - \Psi(X_0, A_0) - \int_0^t \frac{\partial}{\partial r}\Psi(\cdot, B_r)(X_r)dr$$

$$+ \int_0^t \langle\nabla(V - A_r\circ\xi), \nabla\Psi(X_r, A_r)\rangle dW(r)$$

Plan of the talk

1. Motivations
2. The Adaptive Biasing Potential method
3. Consistency and elements of proof
   - Self-interacting diffusions
   - Analysis of the ABP method
4. Extensions
   - Examples
   - Abstract framework

Extensions

The analysis is not restricted to

$$dX_t^0 = -\nabla V(X_t^0)dt + \sqrt{2}dW_t,$$

- overdamped Langevin dynamics
- on the flat d-dimensional torus $\mathbb{T}^d$
- with reaction coordinate $\xi : x \in \mathbb{T}^d \mapsto x_1 \in \mathbb{T}$.

Extensions

The analysis is not restricted to

$$dX_t^0 = -\nabla V(X_t^0)dt + \sqrt{2}dW_t,$$

- overdamped Langevin dynamics $\rightarrow$ Langevin, extended dynamics
- on the flat d-dimensional torus $\mathbb{T}^d$ $\rightarrow$ dynamics on $\mathbb{R}^d$, on $L^2(0,1)$ (SPDE)
- with reaction coordinate $\xi : x \in \mathbb{T}^d \mapsto x_1 \in \mathbb{T}$.
  $\rightarrow$ ξ is smooth with values in a **compact** m-dimensional manifold $\mathbb{M}_m$, for instance $\mathbb{T}^m$ with arbitrary $m \geq 1$.

  The kernel function $K : \mathbb{M}_m \times \mathbb{M}_m \rightarrow (0,\infty)$ is still assumed smooth and positive.

Extension 1: Langevin dynamics

State space: $\mathcal{S} = \mathbb{T}^d \times \mathbb{R}^d$ or $\mathbb{R}^d \times \mathbb{R}^d$.
Reaction coordinate: $\xi(q,p) = \xi(q) \in \mathbb{M}_m$.

Dynamics:

$$\begin{cases} dq_p = p_t dt, \\ dp_t = -\nabla\big(V - A_t \circ \xi\big)(q_t)dt - \gamma p_t + \sqrt{2\gamma\beta^{-1}}dW(t), \end{cases}$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-\beta A_r \circ \xi(q_r)} \delta_{(q_r,p_r)} dr}{1 + \int_0^t e^{-\beta A_r \circ \xi(q_r)} dr} \ , \quad e^{-\beta A_t(z)} = \int K(z,\xi(q))\overline{\mu}_t(dqdp).$$

Consistency: almost surely $\overline{\mu}_t \underset{t\to\infty}{\to} \mu(dq) \otimes \mathcal{N}(0,\sigma^2)$.

Extension 2: extended dynamics

State space: $\mathcal{S} = \mathbb{T}^d \times \mathbb{M}_m$ or $\mathbb{R}^d \times \mathbb{M}_m$.
Reaction coordinate: $\xi(x,z) = z$.
Extended potential energy function: $U(x,z) = V(x) + \frac{1}{2\epsilon}\left|\xi(x) - z\right|^2$.

Dynamics:

$$\begin{cases} dX_t = -\nabla V(X_t)dt - \frac{1}{\epsilon}\nabla\xi(X_t).\big(\xi(X_t) - Z_t\big)dt + \sqrt{2\beta^{-1}}dW_t^x \\ dZ_t = -\frac{1}{\epsilon}\big(Z_t - \xi(X_t)\big)dt + \nabla A(Z_t)dt + \sqrt{2\beta^{-1}}dW_t^z. \end{cases}$$

$$\overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-\beta A_r(Z_r)}\delta_{(X_r,Z_r)}dr}{1 + \int_0^t e^{-\beta A_r(Z_r)}dr}\ , \quad e^{-\beta A_t(z)} = \int K(z,\zeta)\overline{\mu}_t(dxd\zeta).$$

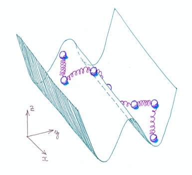
Consistency: almost surely $\overline{\mu}_t \underset{t\to\infty}{\to} \frac{1}{Z(\beta,\epsilon)} e^{-\beta V(x)} e^{-\frac{\beta(\xi(x)-z)^2}{2\epsilon}} dxdz$.

Extension 3: infinite dimensional dynamics (SPDE)

$$\begin{cases} du^0(t,x) = \frac{\partial^2 u^0(t,x)}{\partial x^2} dt - \nabla \mathcal{V}(u^0(t,x)) dt + \sqrt{2\beta^{-1}} dW(t,x) \\ u^0(t,0) = 0 = u^0(t,1). \end{cases}$$

Example: $\mathcal{V}(x) = \frac{x^4}{4} - \frac{x^2}{2}$ (Allen-Cahn equation).
Energy functional:

$$u \mapsto \int_0^1 \Big[\frac{1}{2}\Big|\frac{\partial u(x)}{\partial x}\Big|^2 + V\big(u(x)\big)\Big] dx,$$



Extension 3: infinite dimensional dynamics (SPDE)

State space: $H = L^2(0,1)$ separable infinite dimensional Hilbert space

Invariant distribution of the SPDE

$$du_t^0 = Lu_t^0 dt - DV(u_t^0)dt + \sqrt{2}dW(t)$$

is

$$\mu(du) = e^{-V(u)}\lambda(du)$$

with $V(u) = \int_0^1 \mathcal{V}(u(\cdot))$ and $\lambda(du) = \mathcal{N}(0, L^{-1})$ Gaussian distribution on H.

In this example: λ is the distribution of the Brownian Bridge.

Extension 3: infinite dimensional dynamics (SPDE)

Reaction coordinate: $\xi(u) = \int_0^1 u \mod M$.

Dynamics:

$$\begin{cases} du(t) = Lu(t)dt - D\big(V - A_t \circ \xi\big)(u(t))dt + \sqrt{2\beta^{-1}}dW(t) \\ \overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t e^{-\beta A_r \circ \xi(u_r)} \delta_{u_r} dr}{1 + \int_0^t e^{-\beta A_r \circ \xi(u_r)} dr} \\ e^{-\beta A_t(z)} = \int K\big(z, \xi(u)\big) \overline{\mu}_t(du) \end{cases}$$

Consistency: almost surely $\overline{\mu}_t(\varphi) \underset{t\to\infty}{\to} \mu(\varphi)$, for all smooth functions $\varphi : H \to \mathbb{R}$.

Abstract framework

Dynamics (with fixed bias A):

$$dX_t^A = \mathcal{D}(V, A)(X_t^A)dt + \sqrt{2\beta^{-1}}\Sigma dW(t)$$

on a state space $\mathcal{S}$.

Abstract framework

Dynamics (with fixed bias A):

$$dX_t^A = \mathcal{D}(V,A)(X_t^A)dt + \sqrt{2\beta^{-1}}\Sigma dW(t)$$

on a state space $\mathcal{S}$.

Potential energy function: $V : E \to \mathbb{R}$, smooth.
Bias: $A : \mathbb{M}_m \to \mathbb{R}$.
Reaction coordinates: $\xi : E \to \mathbb{M}$ and $\xi_{\mathcal{S}} : \mathcal{S} \to \mathbb{M}$.

Abstract framework

Invariant distribution:

$$\mu_\beta^A(dx) = \frac{1}{Z^A(\beta)} e^{-\beta \mathcal{E}(V,A)(x)} \lambda(dx),$$

with a reference measure λ on $\mathcal{S}$.
Compatibility condition: $\mathcal{E}(V, A) = \mathcal{E}(V, 0) - A \circ \xi_{\mathcal{S}}$.

Abstract framework

Invariant distribution:

$$\mu_\beta^A(dx) = \frac{1}{Z^A(\beta)} e^{-\beta \mathcal{E}(V,A)(x)} \lambda(dx),$$

with a reference measure λ on $\mathcal{S}$.
Compatibility condition: $\mathcal{E}(V,A) = \mathcal{E}(V,0) - A \circ \xi_{\mathcal{S}}$.

Free energy function: $e^{-\beta A_\star}$ is the Radon-Nikodym derivative:

$$\int_{\mathbb{M}_m} \phi(z) e^{-\beta A_\star(z)} \pi(dz) = \int_{\mathcal{S}} \phi\big(\xi_{\mathcal{S}}(x)\big) \mu_\beta(dx),$$

with respect to a reference probability distribution π on $\mathbb{M}_m$.

Abstract ABP

Adaptive Biasing Potential dynamics:

$$\begin{cases} dX_t = \mathcal{D}\big(V, A_t\big)(X_t)dt + \sqrt{2\beta^{-1}}\Sigma dW_t, \\ \overline{\mu}_t = \frac{\overline{\mu}_0 + \int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau))\delta_{X_\tau} d\tau}{1+\int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau))d\tau}, \\ F_t = \mathcal{N}\Big(\int_{\mathcal{S}} K\big(\cdot, \xi_{\mathcal{S}}(x)\big)\overline{\mu}_t(dx)\Big), \\ A_t = -\frac{1}{\beta}\log\big(\overline{F}_t\big), \end{cases}$$

with

- an additional variable F_t (weights)
- a normalization operator $\mathcal{N}$: e.g.

$$\mathcal{N}(f) = \overline{f} = \frac{f}{\int f d\pi}, \quad \mathcal{N}(f) = \frac{f}{\max f}, \quad \mathcal{N}(f) = \frac{f}{\min f}.$$

Conclusion and perspectives

Mathematical analysis of the asymptotic behavior of an ABP method:

- interpretation as a self-interacting diffusion, reinforcement with the past
- almost sure convergence results (consistency)
- a general framework

Some perspectives:

- more quantitative analysis?
- construction and analysis of ABF methods?

Conclusion and perspectives

Mathematical analysis of the asymptotic behavior of an ABP method:

- interpretation as a self-interacting diffusion, reinforcement with the past
- almost sure convergence results (consistency)
- a general framework

Some perspectives:

- more quantitative analysis?
- construction and analysis of ABF methods?

References:

M. Benaïm and C.-E. Bréhier.
Convergence of adaptive biasing potential methods for diffusions.
C. R. Math. Acad. Sci. Paris, 2016.

M. Benaïm and C.-E. Bréhier.
Convergence analysis of adaptive biasing potential methods for diffusions.
Arxiv preprint, 07 2017.

Conclusion and perspectives

Mathematical analysis of the asymptotic behavior of an ABP method:

- interpretation as a self-interacting diffusion, reinforcement with the past
- almost sure convergence results (consistency)
- a general framework

Some perspectives:

- more quantitative analysis?
- construction and analysis of ABF methods?

References:

M. Benaïm and C.-E. Bréhier.
Convergence of adaptive biasing potential methods for diffusions.
C. R. Math. Acad. Sci. Paris, 2016.

M. Benaïm and C.-E. Bréhier.
Convergence analysis of adaptive biasing potential methods for diffusions.
Arxiv preprint, 07 2017.

Thanks for your attention.