



**HAL**  
open science

# CONVERGENCE ANALYSIS OF ADAPTIVE BIASING POTENTIAL METHODS FOR DIFFUSION PROCESSES

Michel Benaïm, Charles-Edouard Bréhier

► **To cite this version:**

Michel Benaïm, Charles-Edouard Bréhier. CONVERGENCE ANALYSIS OF ADAPTIVE BIASING POTENTIAL METHODS FOR DIFFUSION PROCESSES. *Communications in Mathematical Sciences*, 2019, 10.4310/CMS.2019.v17.n1.a4 . hal-01562639

**HAL Id: hal-01562639**

**<https://hal.science/hal-01562639>**

Submitted on 16 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CONVERGENCE ANALYSIS OF ADAPTIVE BIASING POTENTIAL METHODS FOR DIFFUSION PROCESSES

MICHEL BENAÏM AND CHARLES-EDOUARD BRÉHIER

ABSTRACT. This article is concerned with the mathematical analysis of a family of adaptive importance sampling algorithms applied to diffusion processes. These methods, referred to as Adaptive Biasing Potential methods, are designed to efficiently sample the invariant distribution of the diffusion process, thanks to the approximation of the associated free energy function (relative to a reaction coordinate). The bias which is introduced in the dynamics is computed adaptively; it depends on the past of the trajectory of the process through some time-averages.

We give a detailed and general construction of such methods. We prove the consistency of the approach (almost sure convergence of well-chosen weighted empirical probability distribution). We justify the efficiency thanks to several qualitative and quantitative additional arguments. To prove these results, we revisit and extend tools from stochastic approximation applied to self-interacting diffusions, in an original context.

## 1. INTRODUCTION

In many applications in physics, biology, chemistry, etc... there is a huge interest in the two following problems. First, in sampling probability distributions (denoted by  $\mu$ ), *i.e.* in constructing families of independent random variables with distribution  $\mu$ . Second, in computing averages  $\int \varphi d\mu$  of real-valued functions  $\varphi$ . These questions lead to challenging computational issues, when the dimension of the support of  $\mu$  is large (possibly infinite) – for instance, when  $\mu$  is the equilibrium distribution of a large system of particles, which is the typical situation in the field of molecular dynamics. The scientific literature contains many examples, as well as many approaches to construct efficient approximation procedures. We do not intend to provide an extensive review; some relevant examples, which are connected to the methodology studied in this article, will be provided below.

Many methods are based on a Markov Chain Monte Carlo (MCMC) strategy, *i.e.* on sampling a stochastic process which is ergodic with respect to  $\mu$ . The standard example is given by the overdamped Langevin dynamics on  $\mathbb{R}^d$ ,

$$dx_t = -\nabla V(x_t)dt + \sqrt{2\beta^{-1}}dW_t,$$

associated – under appropriate growth and regularity assumptions on the function  $V$  – with the Boltzmann-Gibbs probability distribution

$$\mu(dx) = \mu_\beta(dx) = \frac{e^{-\beta V(x)}}{Z(\beta)}dx.$$

The choice of ergodic processes associated with a probability distribution  $\mu$  is not unique. In this article, we will only focus on diffusion processes, in continuous time. These processes are in most cases solutions of Stochastic Differential Equations (SDEs), in finite dimension. We will also study the case of Stochastic Partial Differential Equations (SPDEs), in infinite dimension. For a practical implementation of the methods, an additional time discretization of the continuous-time dynamics is required. However, in this article, we do not discuss this question, and we only perform the analysis at the continuous-time level.

---

*Key words and phrases.* adaptive biasing, self-interacting diffusions, free energy computation.

For the class of problems we are interested in, one of the limitations of standard MCMC approaches is the multimodality of  $\mu$ : in this situation, the support of  $\mu$  contains several disjoint regions which all have large measure, and are separated by regions with low probability. When  $\mu$  is multimodal, the associated ergodic dynamics are metastable. A direct simulation is not able to efficiently and accurately sample the rare transitions between metastable states, hence the need for advanced Monte-Carlo methods.

Many strategies have been proposed, analyzed and applied, to overcome the issues raised above. The associated variance reduction approaches may be divided into two main families. On the one hand, Importance Sampling strategies are based on changing the reference probability measure. The realization of the rare events limiting standard approaches is enhanced by appropriate reweighting of  $\mu$ . In our context, such strategies require to simulate modified processes, which are constructed by biasing the dynamics. On the other hand, Splitting strategies use interacting replicas, with mutation and selection procedures, without modifying the process dynamics.

The methods studied in this article are an example of Adaptive Biasing methods. They are based on the Importance Sampling strategy and the use of the so-called Free Energy function (which will be introduced below). When going into the details of the schemes and of the applications, there are many different versions; they all aim at flattening the free energy landscape, and to make free energy barriers disappear. We refer to the monograph [33] for an extensive review of such methods, and to [34, Section 4] for a survey on mathematical techniques. To name a few of the versions, we mention the following examples of adaptive biasing methods: the adaptive biasing force [18], [26] [15]; the Wang-Landau algorithm [37], [38]; metadynamics [30], [1]; the self-healing umbrella sampling method [36]. For related mathematical results, see for instance [14], [27], [31], [32] (adaptive biasing force); [23], [24] (Wang-Landau), [25] (self-healing umbrella sampling). This list is not exhaustive. We also refer to the recent survey paper [20] (and to references therein) for discussions and comparison of these methods.

Our aim in this article is to give a mathematical analysis of a family of methods, independently of a comparison with the other methods mentioned above: the Adaptive Biasing Potential methods, related to [19]. We emphasize on one feature of the method studied here: the process dynamics is modified with quantities computed as time-averages over a single realization of the dynamics.

Let us mention the type of mathematical properties such algorithms are required to satisfy (exactly or in an approximate sense) concerning estimators of averages  $\int \varphi d\mu$ . On the one hand, the consistency is the long-time convergence to the average, in a strong sense (almost sure or  $L^p$ ), or in a weak sense (convergence of expected value). On the other hand, the efficiency is generally studied in terms of the asymptotic mean-square error; as we will also see in this article, it may also be considered from the point of view of the long-time behavior of occupation measures of the process.

A preliminary analysis of the Adaptive Biasing Potential methods considered in this article, is provided in [3], in a simplified framework. We go beyond in this present article, and give detailed arguments and substantial generalizations. Below we first present the methods in a simplified framework, see Section 1.1, following [3]: we explain the strategy and the type of results. We then present the general framework of the article, and the associated results, in Sections 1.2 and 1.3.

**1.1. Adaptive Biasing Potential method in a simplified framework.** We introduce the purpose and results of the present article in a simplified framework, following [3]. This section is pedagogical: the ideas are introduced independently of the abstract notation which will allow us to consider many examples of diffusion processes. We are interested in sampling probability distributions on the flat  $d$ -dimensional torus  $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$ , of the following form:

$$\mu(dx) = \mu_\beta(dx) = \frac{\exp(-\beta V(x))}{Z(\beta)},$$

where  $V : \mathbb{T}^d \rightarrow \mathbb{R}$  is a smooth potential energy function, and  $\beta \in (0, \infty)$  is referred to as the inverse temperature. Finally,  $dx$  is the Lebesgue measure on  $\mathbb{T}^d$ , and  $Z(\beta) = \int_{\mathbb{T}^d} e^{-\beta V(x)} dx$  is the normalizing constant.

A natural choice of associated ergodic process is given by the overdamped Langevin dynamics (or Brownian dynamics)

$$(1) \quad dX_t^0 = -\nabla V(X_t^0)dt + \sqrt{2\beta^{-1}}dW_t, \quad X_0^0 = x_0,$$

where  $(W_t)_{t \geq 0}$  is a standard Wiener process on  $\mathbb{T}^d$ .

It is a standard result that the empirical distribution  $\bar{\mu}_t^0 = \frac{1}{t} \int_0^t \delta_{X_r^0} dr$  converges (in distribution), almost surely, towards  $\mu_\beta$ , when time  $t$  goes to infinity. However, the convergence may be slow due to the metastability of the process  $X^0$ .

To accelerate convergence to equilibrium, other stochastic processes need to be used. In this article, the modification of the dynamics is an adaptive change of the potential energy function: the function  $V$  is replaced with a time-dependent function  $V_t$  – hence the terminology *Adaptive Biasing Potential* (ABP) method:

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

Compared with other methods mentioned above, one of the specificities of the method considered in this article, is the structure of the time-dependent potential energy function  $V_t$ : it is constructed as  $V_t = V - A_t \circ \xi$ , where  $\xi : (x_1, \dots, x_d) \in \mathbb{T}^d \mapsto (x_1, \dots, x_m) \in \mathbb{T}^m$  is an auxiliary function, referred to as the reaction coordinate, and  $A_t : \mathbb{T}^m \rightarrow \mathbb{R}$  is an approximation (in the regime  $t \rightarrow \infty$ ) of the so-called Free Energy function. In general, one assumes  $m \in \{1, \dots, d-1\}$ ; however, in applications, the dimension  $m$  is chosen much smaller than  $d$ , and typically  $m \in \{1, 2, 3\}$ . In other words, the difference  $V_t - V$  only depends on  $m < d$  components of the process. It thus remains to explain how the function  $A_t$  is constructed adaptively: it is done in terms of the values  $(X_r)_{0 \leq r < t}$  of the process  $X$  up to time  $t$ .

Precisely, the dynamics of the ABP method, in the simplified framework considered in the current section (for the generalized version, see Equation (10)), is given by the following system:

$$(2) \quad \begin{cases} dX_t = -\nabla(V - A_t \circ \xi)(X_t)dt + \sqrt{2\beta^{-1}}dW(t) \\ \bar{\mu}_t = \frac{\bar{\mu}_0 + \int_0^t \exp(-\beta A_r \circ \xi(X_r)) \delta_{X_r} dr}{1 + \int_0^t \exp(-\beta A_r \circ \xi(X_r)) dr} \\ \exp(-\beta A_t(z)) = \int_{\mathbb{T}^d} K(z, \xi(x)) \bar{\mu}_t(dx), \quad \forall z \in \mathbb{T}^m, \end{cases}$$

where a smooth kernel function  $K : \mathbb{T}^m \times \mathbb{T}^m \rightarrow (0, +\infty)$ , such that  $\int_{\mathbb{T}^m} K(z, \zeta) dz = 1, \forall \zeta \in \mathbb{T}^m$ , is introduced. It serves to define a smooth real-valued function  $A_t$  defined on  $\mathbb{T}^m$ , in terms of the probability distribution  $\bar{\mu}_t$  on  $\mathbb{T}^d$ .

The unknowns in (2) are the stochastic processes  $t \mapsto X_t \in \mathbb{T}^d$ ,  $t \mapsto \bar{\mu}_t \in \mathcal{P}(\mathbb{T}^d)$  (the set of Borel probability distributions on  $\mathbb{T}^d$ , endowed with the usual topology of weak convergence of probability distributions), and  $t \mapsto A_t \in \mathcal{C}^\infty(\mathbb{T}^m)$  (the set of infinitely differentiable functions on  $\mathbb{T}^m$ ). Initial conditions  $X_{t=0} = x_0$  and  $\bar{\mu}_{t=0} = \bar{\mu}_0$  are prescribed.

We emphasize on a key observation. The third equation in (2) introduces a coupling between the evolutions of the diffusion process  $X_t$  and of the probability distribution  $\bar{\mu}_t$ . Thus the system defines a type of self-interacting diffusion process. However, a comparison with [4] and subsequent articles [5], [6] and [7], reveals a different form of coupling. The aim of this article is to study the new arguments which are required for the study of the system (2).

The most important quantity in (2) is the random, time-dependent, probability distribution  $\bar{\mu}_t$ . Observe that its construction uses two successive operations: first, a weighted occupation measure  $\mu_t = \bar{\mu}_0 + \int_0^t \exp(-\beta A_r \circ \xi(X_r)) \delta_{X_r} dr$  is computed; second, this measure is normalized as

a probability distribution,  $\bar{\mu}_t = \frac{\mu_t}{\int_{x \in \mathbb{T}^d} \mu_t(dx)}$ . The weights  $\exp(-\beta A_r \circ \xi(X_r))$  in the definition of  $\bar{\mu}_t$  are chosen so as to obtain the following consistency result.

**Theorem 1.1.** *Almost surely,  $\bar{\mu}_t \xrightarrow[t \rightarrow \infty]{} \mu_\beta$ , in  $\mathcal{P}(\mathbb{T}^d)$ .*

*Moreover, define the function  $A_\infty$ , such that  $\exp(-\beta A_\infty(\cdot)) = \int K(\cdot, \xi(x)) \mu_\beta(dx)$ . Then, almost surely,  $A_t \xrightarrow[t \rightarrow +\infty]{} A_\infty$ , in  $\mathcal{C}^k(\mathbb{T}^m)$ ,  $\forall k \in \mathbb{N}$ .*

The proof of this result is described in our previous work [3]. In the present article, in a more general context, we provide the technical details.

Let us briefly explain the role played by the weights in this result. On the one hand, their presence ensures the convergence of  $\bar{\mu}_t$  to  $\mu_\beta$ , *i.e.* the consistency. This observation is not surprising, indeed it is a standard feature of importance sampling approaches. On the other hand, the convergence of  $A_t$  to  $A_\infty$  comes from the way the evolutions of  $X_t$  and  $\bar{\mu}_t$  are coupled, in the third equation of (2). The convergence of  $A_t$  reveals the efficiency of the method: indeed,  $A_\infty$  is an approximation of the so-called Free Energy function  $A_\star$ , defined by

$$\exp(-\beta A_\star(z)) = \int_{\mathbb{T}^{d-m}} \frac{\exp(-\beta V(z, x_{m+1}, \dots, x_d))}{Z(\beta)} dx_{m+1} \dots dx_d.$$

Note that, by construction,  $\exp(-\beta A_\star(z)) dz$  is a probability distribution on  $\mathbb{T}^m$ , which is the image of  $\mu_\beta$  by the reaction coordinate  $\xi$ . The Free Energy function depends in general on the temperature parameter  $\beta$ , but to simplify notation we simply write  $A_\star(z)$ . As will be explained below, biasing the dynamics (in a non-adaptive way) using the function  $A_\star$  (which is not known in practice) would be optimal. The adaptive method can thus naturally be seen as a stochastic approximation algorithm, with the optimal parameter being learned on-the-fly. See [2],[8],[21],[29] for standard references.

The observations made above, concerning the system (2) and the consistency result, Theorem 1.1, are not specific to the simple example considered in this section; they will also hold true in the general framework which is described next, in Section 1.2.

**1.2. General framework.** We now describe the framework considered in this article, and how the content of Section 1.1 can be generalized in various directions.

First, contrary to (1), the state space of the dynamics is not necessarily compact; in addition, it may be infinite dimensional.

The main generalization then concerns the type of diffusion processes which are considered: instead of the overdamped Langevin dynamics (1), the framework will encompass the following examples: (hypoelliptic) Langevin dynamics – the unknowns being positions and momenta, instead of positions only; extended dynamics – where an auxiliary variable is associated with the mapping  $\xi$ , see [35]; Stochastic Partial Differential Equations (SPDEs) – which are infinite dimensional diffusion processes. Considering diffusions on smooth manifolds could also be possible in the abstract framework developed below; however to simplify the presentation we do not treat such examples.

Abstract notation and analysis allows us to treat jointly these cases; however the SPDE example is in fact presented separately, in Section 7, to simplify the exposition. Indeed several arguments need to be adapted, whereas it is simple to consider the other examples in the same framework.

Another generalization concerns the function  $\xi$ , referred to as the reaction coordinate mapping. It is allowed to have more complicated structure than in the simplified example above. The definition of the associated free energy is made precise in terms of a Radon-Nikodym derivative of the image of the invariant distribution  $\mu_\beta$  by  $\xi$ , with respect to a reference measure. The key properties which are required are that  $\xi$  takes values in a compact set, and that  $\xi$  is smooth.

Note that the general framework and the associated abstract notation are constructed to emphasize on the main assumptions on the models and on the algorithm, which are required for the

well-posedness and the consistency of the approach. The list of examples we propose and treat is not intended to be exhaustive.

**1.3. Organization of the paper.** In Section 2, the abstract framework is introduced, with emphasis on the following objects: the diffusion process dynamics, Section 2.1 – with examples provided in Section 2.2 – the invariant probability distribution  $\mu$ , Section 2.3, and the free energy function, Section 2.4.

The construction of generalized versions of the Adaptive Biasing Potential method, given by (2), is provided in Section 3. In particular, well-posedness results and important estimates are stated precisely there.

Section 4 contains the main results of this article, in finite dimensional cases, concerning the long-time behavior of the method. On the one hand, the consistency of the approach, *i.e.* the almost sure convergence (in distribution) of  $\bar{\mu}_t$  to  $\mu$ , is given in Theorem 4.1 and Corollary 4.2. On the other hand, the efficiency is analyzed first in terms of the convergence of the approximation  $A_t$  of the free energy function, Corollary 4.3, and of occupation measures, Corollary 4.4; second, in terms of the asymptotic mean-square error, Proposition 4.5.

The proof of consistency (Theorem 4.1) is detailed in Section 5. We first describe (see Section 5.1) how to obtain a system which is similar to the self-interacting diffusion processes studied in [4], as explained in [3]. This system is obtained thanks to a (random) change of time variable, such that the weights in the definition of  $\bar{\mu}_t$  (second line in (2)) are eliminated. We then explain how the consistency of the approach may be seen as a straightforward consequence of the ODE method from stochastic approximation (see [2], [8], [21], [29]), thanks to an asymptotic time scale separation into slow (occupation measure) and fast (diffusion process) evolutions. We do not provide all the details for this approach; indeed, a direct analysis, without the change of variable, is possible. The analysis is performed, with details, in Section 5.2. From a technical point of view, the main specific task for the present article is the study of a related Poisson equation.

Section 6 is devoted to the analysis of the asymptotic mean-square error. We recover in our particular framework a rather standard fact in stochastic approximation: asymptotically the adaptive system variance is the same as for a non-adaptive system where the bias is chosen as the limit of the adaptive bias  $A_t$ .

Finally, Section 7 is devoted to an analysis in infinite dimension, where the diffusion processes are solutions of some SPDEs. The abstract framework encompasses this situation. However the analysis is presented separately: indeed several arguments need substantial modifications.

## 2. FRAMEWORK

**2.1. Dynamics and abstract notation.** The discussion in the present section is general. Examples are provided in Section 2.2.

**2.1.1. Unbiased dynamics.** The unbiased (or original) dynamics is a diffusion process  $(X_t^0)_{t \geq 0}$ , with values on a state space denoted by  $\mathcal{S}$ . The process is solution of a Stochastic Differential Equation (SDE), when the dimension of  $\mathcal{S}$  is finite; or of a Stochastic Partial Differential Equation (SPDE), when the dimension of  $\mathcal{S}$  is infinite. The SDE or the SPDE is written in the following form

$$(3) \quad dX_t^0 = \mathcal{D}(V)(X_t^0)dt + \sqrt{2\beta^{-1}\Sigma}dW_t \quad , \quad X_0^0 = x_0,$$

where  $(W_t)_{t \in \mathbb{R}_+}$  is a standard Wiener process on  $\mathcal{S}$ , and  $\Sigma$  is a linear mapping which is specified in each example below.

In (3), the initial condition  $x_0 \in \mathcal{S}$  is arbitrary, and is assumed to be deterministic for simplicity. The convergence results may be extended to a random initial condition (independent of the Wiener noise) by a standard conditioning argument. The value of  $x_0$  plays no role in the analysis below.

The drift coefficient  $\mathcal{D}(V)$  in (3) depends on the potential energy function  $V : E_d \rightarrow \mathbb{R}$ , defined on a set  $E_d$ , where  $E_d = \mathbb{T}^d$  (periodic, compact case) or  $E_d = \mathbb{R}^d$  (non compact case). Note that in general  $\mathcal{S} \neq E_d$ . The functions  $V$  and  $\mathcal{D}(V)$  are assumed of class  $\mathcal{C}^\infty$ ; in the non compact case  $E_d = \mathbb{R}^d$ , growth conditions are required, they will be described for each example.

Finally,  $\beta \in (0, \infty)$  is a positive, fixed, parameter, referred to as the inverse temperature.

**Remark 2.1.** *Instead of assuming that  $V$  is of class  $\mathcal{C}^\infty$ , we may assume that  $V$  is of class  $\mathcal{C}^n$  for some sufficiently large integer  $n$  (at least  $n \geq 2$ ). The results stated in this article remain valid (if properly adapted) in this setting. To simplify the analysis and avoid arguments and statements depending on the parameter  $n$ , we have chosen to deal with  $\mathcal{C}^\infty$  regularity.*

2.1.2. *Non-adaptively biased dynamics.* Having introduced the unbiased dynamics (3), we now describe the family of biased dynamics which we consider in this article, first in a non-adaptive context. The drift coefficient  $\mathcal{D}(V)$  in (3) is modified, being replaced by  $\mathcal{D}(V, A)$ ; more precisely, the function  $A$  depends only on a small number of degrees of freedom of the system. In the current section, the bias is non-adaptive: the function  $A$  is deterministic and does not depend on time.

We now make precise how  $\mathcal{D}(V, A)$  is defined: it depends on the mapping  $A \circ \xi$ , where

- $\xi : E_d \rightarrow \mathbb{M}_m$  is a fixed smooth function, where  $\mathbb{M}_m$  is a  $m$ -dimensional smooth, compact, manifold, and  $m \in \{1, \dots, d-1\}$ .
- $A : \mathbb{M}_m \rightarrow \mathbb{R}$  is a smooth function.

The mapping  $\xi$  is called the reaction coordinate; variables  $z = \xi(x)$  are often called collective variables. The name “reaction coordinate” comes from molecular dynamics applications, where methods such as the one studied in this article are popular computational approaches. The functions  $\xi$  and  $A \circ \xi$  are defined on  $E_d$ , like the potential energy function  $V$ .

Finally, an extension

$$\xi_{\mathcal{S}} : \mathcal{S} \rightarrow \mathbb{R}$$

of the reaction coordinate  $\xi$ , is also defined on the state space  $\mathcal{S}$  – the procedure depends on the example of diffusion process.

In the non-compact case  $E_d = \mathbb{R}^d$ , all the derivatives of  $\xi$  are assumed to be bounded.

As explained in the introduction, in practice one chooses  $m$  much smaller than  $d$ , and typically for concrete applications  $m \in \{1, 2, 3\}$

Note that the compactness assumption on  $\mathbb{M}_m$  is crucial in this article. In particular, it allows us to establish some stability estimates and the well-posedness of the ABP system. In some cases, it might be possible to remove this restriction (and consider for instance  $\mathbb{M}_m = \mathbb{R}^m$ ), by proving appropriate estimates; we leave this non trivial technical issue for future works. The assumption that  $\mathbb{M}_m$  is a smooth manifold is requested to define potential energy functions  $V - A \circ \xi$  with nice regularity properties.

To simplify the discussion, from now on  $\mathbb{M}_m = \mathbb{T}^m$  is the flat  $m$ -dimensional torus. However, we use the abstract notation and conditions to suggest possible straightforward generalizations.

We are now in position to define the biased dynamics  $(X_t^A)_{t \geq 0}$ , for any given  $A : \mathbb{M}_m \rightarrow \mathbb{R}$  of class  $\mathcal{C}^\infty$ :

$$(4) \quad dX_t^A = \mathcal{D}(V, A)(X_t^A)dt + \sqrt{2\beta^{-1}}\Sigma dW_t \quad , \quad X_0^A = x_0.$$

Consistently, we will have  $\mathcal{D}(V, 0) = \mathcal{D}(V)$ : in the absence of bias, the biased dynamics (4) is simply the unbiased dynamics (3).

**2.2. Examples of diffusions processes.** In this section, we now make explicit what are the main examples of diffusion processes we study in this article. First, we consider (elliptic) Brownian dynamics and (hypoelliptic) Langevin dynamics. The third example is popular for applications; it is called the extended dynamics, since an additional variable is considered.

We postpone the study of a fourth example, given by infinite dimensional diffusion processes (SPDE), to Section 7. Indeed, several arguments need to be modified.

From now on, except in Section 7, the state space  $\mathcal{S}$  is finite dimensional.

### 2.2.1. Brownian dynamics.

- **State space:**  $\mathcal{S} = E_d$ .
- **Reaction coordinate:**  $\xi_{\mathcal{S}} = \xi$ .
- **Drift coefficient:**  $\mathcal{D}(V, A) = \mathcal{D}(V - A \circ \xi) = -\nabla(V - A \circ \xi)$ . **Diffusion operator:**  $\Sigma = I$ , where  $I$  denotes the identity matrix.

In the Brownian case, the dynamics (4) is written as

$$dx_t^A = -\nabla(V - A \circ \xi)(x_t^A)dt + \sqrt{2\beta^{-1}}dW_t.$$

In the non compact case,  $E_d = \mathbb{R}^d$ , the following assumption on the potential energy function  $V$  is granted.

**Assumption 2.2.** When  $E_d = \mathbb{R}^d$ , assume that there exist  $\alpha_V \in (0, \infty)$  and  $C_V \in \mathbb{R}$ , such that for all  $x \in E_d$ ,

$$\langle x, \nabla V(x) \rangle \geq \alpha_V |x|^2 - C_V.$$

Assumption 2.2 is satisfied for instance for smooth potential functions  $V$  which behave like  $|\cdot|^k$  at infinity, for some  $k \geq 2$ .

**Remark 2.3.** Assume  $E_d = \mathbb{R}^d$ . Let  $J$  be a  $d \times d$  skew-symmetric matrix. The results of this article also apply when choosing  $\mathcal{D}(V, A) = -(I + J)\nabla(V - A \circ \xi)$ . When  $J \neq 0$ , this corresponds to non-reversible dynamics (with respect to the associated invariant distribution). In our study, the reversibility plays no role, however, the knowledge of a bijective mapping  $\mu_\beta \mapsto \mathcal{D}(V)$  is crucial.

### 2.2.2. Langevin dynamics.

- **State space:**  $\mathcal{S} = E_d \times \mathbb{R}^d$ . Elements of  $\mathcal{S}$  are denoted by  $(q, p)$ .
- **Reaction coordinate:**  $\xi_{\mathcal{S}}(q, p) = \xi(q)$ .
- **Drift coefficient:**  $\mathcal{D}(V, A)(q, p) = \begin{pmatrix} -\nabla(V - A \circ \xi)(q) - \gamma p \\ \gamma p \end{pmatrix}$ . **Diffusion operator:**  $\Sigma = \sqrt{\gamma} \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$ , for  $\gamma \in (0, \infty)$  a damping parameter.

Observe that in this example, the state space  $\mathcal{S}$  is never compact, even if  $E_d$  is.

In the Langevin case, the dynamics (4) is written

$$\begin{cases} dq_t^A = p_t^A dt & , \quad q_0^A = q_0, \\ dp_t^A = -\nabla(V - A \circ \xi)(q_t^A)dt - \gamma p_t^A dt + \sqrt{2\beta^{-1}\gamma}d\tilde{W}_t & , \quad p_0^A = p_0, \end{cases}$$

where  $(\tilde{W}_t)_{t \geq 0}$  is a standard Wiener process on  $\mathbb{R}^d$ . In applications, the variable  $q$  represents positions of particles, whereas the variable  $p$  represents their momenta.

The value of the damping parameter  $\gamma$  plays no role in the analysis below. We recall that in the limit  $\gamma \rightarrow \infty$ , one recovers (up to a rescaling of the time variable) the Brownian dynamics of Section 2.2.1, which is thus often referred to as the overdamped Langevin dynamics. Recall also the analysis of these two cases is different, since the Langevin diffusion is hypoelliptic, whereas the Brownian dynamics is elliptic.



In the non-compact case,  $E_d = \mathbb{R}^d$ , the following assumption on the potential energy function  $V$  is granted.

**Assumption 2.4.** When  $E_d = \mathbb{R}^d$ , assume that  $V$  is bounded from below: there exists  $V_- \in \mathbb{R}$  such that  $V(q) \geq V_-$  for all  $q \in \mathbb{R}^d$ .

Moreover, assume that there exist  $A_V, B_V \in (0, \infty)$  and  $C_V \in \mathbb{R}$  such that for all  $q \in \mathbb{R}^d$

$$\langle q, \nabla V(q) \rangle \geq A_V V(q) + B_V |q|^2 + C_V.$$

2.2.3. *Extended dynamics.* This example is a modification of the Brownian dynamics from Section 2.2.1. It is also straightforward to propose a similar modification of the Langevin dynamics of Section 2.2.2; the details are left to the reader.

- **State space:**  $\mathcal{S} = E_d \times \mathbb{M}_m$ . Elements of  $\mathcal{S}$  are denoted by  $(x, z)$ .
- **Reaction coordinate:**  $\xi_{\mathcal{S}}(x, z) = z$ .
- **Drift coefficient:**  $\mathcal{D}(V, A)(x, z) = \begin{pmatrix} -\nabla_x U_A(x, z) \\ -\nabla_z U_A(x, z) \end{pmatrix}$  where  $U_A(x, z) = U(x, z) - A(z)$ ,  $U(x, z) = V(x) + \frac{1}{2\epsilon} V_{\text{ext}}(\xi(x), z)$  is the extended potential energy function. It depends on a smooth function  $V_{\text{ext}} : \mathbb{M}_m \times \mathbb{M}_m \rightarrow \mathbb{R}$ , and on  $\epsilon \in (0, \infty)$ . **Diffusion operator:**  $\Sigma$  is the identity.

In the case  $\mathbb{M}_m = \mathbb{T}^m$  considered here, one may choose  $V_{\text{ext}}(\xi(x), z) = (\xi(x) - z)^2$ . Then, in the (Brownian) extended case, the dynamics (4) is written as

$$\begin{cases} dX_t^A = -\nabla V(X_t^A)dt - \frac{1}{\epsilon} \langle \nabla \xi(X_t^A), \xi(X_t^A) - Z_t^A \rangle dt + \sqrt{2\beta^{-1}} dW_t^x & , \quad X_0^A = x_0, \\ dZ_t^A = -\frac{1}{\epsilon} (Z_t^A - \xi(X_t^A))dt + \nabla A(Z_t^A)dt + \sqrt{2\beta^{-1}} dW_t^z & , \quad Z_0^A = z_0, \end{cases}$$

for some arbitrary initial condition  $z_0 \in \mathbb{M}_m$ , and where  $(W_t^x)_{t \geq 0}$  and  $(W_t^z)_{t \geq 0}$  are independent standard Wiener processes, on  $E_d$  and on  $\mathbb{M}_m$  respectively. The dynamics is thus obtained by considering the Brownian dynamics on  $E_d \times \mathbb{M}_m$ , with potential energy function  $U$ .

Observe that in this case  $\mathcal{D}(V, A)$  is not a function of  $V - A \circ \xi$ , contrary to the other examples. Indeed, in the extended dynamics, the variable  $z = \xi_{\mathcal{S}}(x, z)$  is considered as an additional, non-physical, variable; in addition,  $A$  is a function of  $z$  only, not of  $\xi(x)$ .

We will explain below why the extended dynamics is relevant, in the limit  $\epsilon \rightarrow 0$ , for the problem of sampling the initial distribution, and thus why it may be sufficient to deal with the extended dynamics case. Since  $\xi_{\mathcal{S}}(x, z) = z$  in this example, it is not restrictive to consider reaction coordinates of the form  $\xi(x_1, \dots, x_d) = (x_1, \dots, x_m)$ , up to considering the extended dynamics.

In the non compact case,  $E_d = \mathbb{R}^d$ , it is assumed that  $V$  satisfies Assumption 2.2 (or Assumption 2.4 if one starts from the Langevin dynamics). Then the extended potential energy function  $U$  also satisfies a similar condition (recall that  $\mathbb{M}_m$  is compact) on the extended state space  $E_d \times \mathbb{M}_m$ .

**2.3. Invariant probability distributions of the diffusion processes.** In all the examples presented above in Section 2.2, the diffusion processes,  $(X_t^0)_{t \geq 0}$  and  $(X_t^A)_{t \geq 0}$ , given by (3) and (4), are ergodic. The associated unique invariant distributions, defined on  $\mathcal{S}$  (equipped with the Borel  $\sigma$ -field), are denoted by  $\mu_{\beta}^0$  and  $\mu_{\beta}^A$ . Since the notation is consistent when  $A = 0$ , we only deal with  $\mu_{\beta}^A$ , with arbitrary  $A$ , in the remainder of this section.

The ergodicity in our context is understood in the following sense:

- there exists a unique invariant probability distribution for the Markov process  $X^A$  defined by (4), which is equal to  $\mu_{\beta}^A$ ;

- for any initial condition  $x_0 \in \mathcal{S}$ , almost surely,

$$\frac{1}{t} \int_0^t \delta_{X_\tau^A} d\tau \xrightarrow{t \rightarrow \infty} \mu_\beta^A,$$

where the convergence is interpreted in the sense of convergence of probability distributions on  $\mathcal{S}$ .

The invariant distribution  $\mu_\beta^A$  is expressed explicitly in terms of the following tools:

- a reference Borel,  $\sigma$ -finite, measure  $\lambda$  on  $\mathcal{S}$ , which does not depend on  $V$  and  $A$ ;
- a Total Energy function  $\mathcal{E}(V, A) : \mathcal{S} \rightarrow \mathbb{R}$ .

The expression of  $\mu_\beta^A$  is then provided by:

$$(5) \quad \mu_\beta^A(dx) = \frac{\exp(-\beta \mathcal{E}(V, A)(x))}{Z^A(\beta)} \lambda(dx),$$

where  $Z^A(\beta) = \int_{\mathcal{S}} \exp(-\beta \mathcal{E}(V, A)(x)) \lambda(dx)$  is a normalizing constant.

Even if the expression is explicit, in general computing averages  $\mu_\beta^A(\varphi) = \int_{\mathcal{S}} \varphi d\mu_\beta^A$  is challenging, for instance in large dimensional spaces; and especially when  $A = 0$  and  $\mu_\beta^0$  is multimodal, which is the main concern in this article. This is why non-adaptive or adaptive importance sampling strategies are useful: they propose choices of  $A$  which reduce the computational cost of the sampling.

Let us make precise the reference measure  $\lambda$  and the mapping  $(V, A) \mapsto \mathcal{E}(V, A)$  for the diffusion processes of Section 2.2. First, the total energy function satisfies the important equality

$$(6) \quad \mathcal{E}(V, A) = \mathcal{E}(V, 0) - A \circ \xi_{\mathcal{S}}.$$

Note that this condition depends on the extended reaction coordinate  $\xi_{\mathcal{S}} : \mathcal{S} \rightarrow \mathbb{M}_m$ , not directly on  $\xi : E_d \rightarrow \mathbb{M}_m$ .

It thus remains to specify  $\mathcal{E}(V) = \mathcal{E}(V, 0)$ .

- Brownian dynamics. The reference measure  $\lambda$  is the Lebesgue measure on  $\mathcal{S}$ . The total energy function is  $\mathcal{E}(V) = V$ .
- Langevin dynamics. The reference measure  $\lambda$  is the Lebesgue measure on  $\mathcal{S}$ . The total energy function is the Hamiltonian function,  $\mathcal{E}(V)(q, p) = H(q, p) = V(q) + \frac{|p|^2}{2}$ . The total energy is thus the sum of potential and kinetic energies.
- Extended dynamics. The reference measure  $\lambda$  is the Lebesgue measure on  $\mathcal{S}$  – more precisely, the product of Lebesgue measures on  $E_d$  and  $\mathbb{M}_m = \mathbb{T}^m$ . The total energy function is  $\mathcal{E}(V)(x, z) = U(x, z) = V(x) + \frac{1}{2\epsilon} (\xi(x) - z)^2$ .

The fact that, in all the examples, the dynamics (3) is ergodic with respect to  $\mu_\beta^0$  is a well-known result (with Assumptions 2.2 and 2.4 granted when necessary). Then, the fact that the dynamics (4) is ergodic with respect to  $\mu_\beta^A$  is straightforward. Indeed, for the Brownian and Langevin dynamics examples, note that  $V - A \circ \xi$  satisfies Assumption 2.2 or 2.4 when  $V$  does – recall that  $\mathbb{M}_m$  is compact, and that the derivatives of  $\xi$  are assumed to be bounded. In the case of the extended dynamics, similar arguments, concerning the extended potential energy function  $U_A$ , are sufficient to prove the ergodicity. More generally, note the following property, which will be used later in the adaptive context. The notation  $\partial^k$  represents the derivative of order  $k$ .

**Property 2.5.** *Let  $m, M, M^{(1)}, M^{(2)}, \dots \in \mathbb{R}$  denote real numbers, and*

$$\mathcal{A} \subset \left\{ A \in C^\infty(\mathbb{M}_m, \mathbb{R}) ; \min A \geq m, \max A \leq M, \max |\partial^k A| \leq M^{(k)}, \forall k \geq 1 \right\}.$$

*Then*

- if  $V$  satisfies Assumption 2.2, there exists  $\alpha_{V,\mathcal{A}} \in (0, \infty)$  and  $C_{V,\mathcal{A}} \in (0, \infty)$  such that for every  $A \in \mathcal{A}$  and every  $x \in E_d = \mathbb{R}^d$ ,

$$\langle x, \nabla(V - A \circ \xi)(x) \rangle \geq \alpha_{V,\mathcal{A}}|x|^2 - C_{V,\mathcal{A}}.$$

- if  $V$  satisfies Assumption 2.4, there exists  $A_{V,\mathcal{A}}, B_{V,\mathcal{A}} \in (0, \infty)$  and  $C_{V,\mathcal{A}} \in (0, \infty)$  such that for every  $A \in \mathcal{A}$  and every  $q \in E_d = \mathbb{R}^d$ ,

$$\langle q, \nabla(V - A \circ \xi)(q) \rangle \geq A_{V,\mathcal{A}}V(q) + B_{V,\mathcal{A}}|q|^2 + C_{V,\mathcal{A}}.$$

- For every  $p \geq 1$ ,

$$(7) \quad \sup_{A \in \mathcal{A}} \int_{\mathcal{S}} |x|^p \mu_{\beta}^A(dx) < \infty.$$

The distribution of interest, in practice, is  $\mu_{\beta} = \mu_{\beta}^0$ ; however, sampling the process  $X^A$  provides an approximation of  $\mu_{\beta}^A$ . The following expression provides a way to compute an average  $\mu_{\beta}(\varphi) = \int_{\mathcal{S}} \varphi(x) \mu_{\beta}(dx)$  in terms of averages with respect to  $\mu_{\beta}^A$ : for bounded and continuous functions  $\varphi : \mathcal{S} \rightarrow \mathbb{R}$ ,

$$(8) \quad \mu_{\beta}(\varphi) = \frac{\mu_{\beta}^A(\varphi \exp(-\beta A \circ \xi_{\mathcal{S}}))}{\mu_{\beta}^A(\exp(-\beta A \circ \xi_{\mathcal{S}}))}.$$

Using (6), averages with respect to  $\mu_{\beta}$  may therefore be computed using ergodic averages along the biased dynamics (4): indeed,

$$(9) \quad \bar{\mu}_t^A(\varphi) = \frac{1 + \int_0^t e^{-\beta A(\xi_{\mathcal{S}}(X_r^A))} \varphi(X_r^A) dr}{1 + \int_0^t e^{-\beta A(\xi_{\mathcal{S}}(X_r^A))} dr} \xrightarrow{t \rightarrow +\infty} \frac{\mu_{\beta}^A(\varphi \exp(-\beta A \circ \xi_{\mathcal{S}}))}{\mu_{\beta}^A(\exp(-\beta A \circ \xi_{\mathcal{S}}))} = \mu_{\beta}(\varphi).$$

This expression serves as the guideline for the construction of the Adaptive Biasing Potential methods (2), and (10) in the general case: the empirical distributions  $\bar{\mu}_t$  are weighted, to ensure consistency. For well chosen functions  $A$ , the convergence is expected to be faster than when  $A = 0$ . In Section 2.4 below, we identify such a function  $A$ , the so-called Free Energy function.

**2.4. The Free Energy function.** In this section, we introduce one of the key quantities in our study: the Free Energy function  $A_{\star} : \mathbb{M}_m \rightarrow \mathbb{R}$ . Note that in general this function depends on the parameter  $\beta$ ; however, in this article, we consider  $\beta$  as a fixed parameter and do not give any result which depends on the value of  $\beta$ . Therefore, to simplify notation, we do not explicitly mention the dependence with respect to  $\beta$  of the Free Energy function.

We explain why this function is a quantity of interest for the computational problem we are interested in, and why it is expected that choosing  $A = A_{\star}$  in the biased dynamics (4) leads to efficient sampling. This property is indeed the guideline of the Adaptive Biasing Potential approach of this article: we construct an adaptive version which is both consistent and designed such that  $A_t$  converges to an approximation  $A_{\infty}$  of  $A_{\star}$  when  $t \rightarrow \infty$ .

The definition of the Free Energy function depends on the choice of a reference probability distribution  $\pi$  on  $\mathbb{M}_m$ . In this article, since  $\mathbb{M}_m = \mathbb{T}^m$ , it is natural to choose the Lebesgue measure, but abstract notation suggests other possible choices, see Remark 2.7 below.

For every  $\beta \in (0, \infty)$ , and every smooth  $A : \mathbb{M}_m \rightarrow \mathbb{R}$ , let  $\pi_{\beta}^A$  denote the image by  $\xi_{\mathcal{S}} : \mathcal{S} \rightarrow \mathbb{M}_m$  of the probability distribution  $\mu_{\beta}^A$  on  $\mathcal{S}$ : recall that this means that for any bounded, continuous function  $\phi : \mathbb{M}_m \rightarrow \mathbb{R}$ ,

$$\int_{\mathbb{M}_m} \phi(z) \pi_{\beta}^A(dz) = \int_{\mathcal{S}} \phi(\xi_{\mathcal{S}}(x)) \mu_{\beta}^A(dx).$$

The following assumption is required.

**Assumption 2.6.** Assume that  $\pi_\beta^0$  and  $\pi$  are equivalent:  $\pi_\beta^0$  (resp.  $\pi$ ) is absolutely continuous with respect to  $\pi$  (resp.  $\pi_\beta^0$ ).

When Assumption 2.6 holds true, then  $\pi_\beta^A$  is equivalent to  $\pi$ , for all smooth functions  $A : \mathbb{M}_m \rightarrow \mathbb{R}$ . Thanks to the smoothness conditions on  $\xi$ , and to growth conditions on  $V$ , Assumption 2.6 is satisfied in all the examples presented above, when  $\pi$  is the Lebesgue measure on  $\mathbb{M}_m = \mathbb{T}^m$ .

**Remark 2.7.** Another natural choice, in the periodic case  $E_d = \mathbb{T}^d$ , for finite dimensional dynamics, is as follows:  $\pi$  is defined as the image by  $\xi : E_d \rightarrow \mathbb{M}_m$  of the Lebesgue measure on  $E_d$ . With this definition,  $\pi$  depends on  $\xi$ . In the non-compact case, for instance one may define (for instance)  $\pi$  as the image by  $\xi$  of the standard Gaussian distribution on  $\mathbb{R}^d$ .

With these examples, Assumption 2.6 is satisfied by construction of  $\pi$ .

We are now in position to define the free energy function  $A_\star$ .

**Definition 2.8.** The free energy function  $A_\star : \mathbb{M}_m \rightarrow \mathbb{R}$  is defined by the following property:  $\exp(-\beta A_\star(\cdot))$  is the Radon-Nikodym derivative of  $\pi_\beta^0$  with respect to  $\pi$ .

This means that for every bounded measurable function  $\phi : \mathbb{M}_m \rightarrow \mathbb{R}$ ,

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

Observe that, thanks to Assumption 2.6,  $A_\star$  takes values in  $(-\infty, \infty)$ . Moreover,  $e^{-\beta A_\star(z)} \pi(dz)$  is by construction a probability distribution on  $\mathbb{M}_m$ , thus no normalizing constant appears on the left-hand side.

It is then straightforward to check that the Radon-Nikodym derivative of  $\pi_\beta^A$  with respect to  $\pi$  is equal to  $\exp(-\beta(A_\star - A))$ , thanks to the condition (6).

The function  $A_\star$  may be interpreted as an effective potential energy function, for the unbiased dynamics, depending on the variable  $z = \xi_{\mathcal{S}}(x)$  only. Indeed, note that for any sufficiently smooth, bounded, function  $\phi : \mathbb{M}_m \rightarrow \mathbb{R}$ , by ergodicity of the unbiased dynamics (3), with respect to  $\mu_\beta$ , almost surely

$$\frac{1}{t} \int_0^t \phi(\xi_{\mathcal{S}}(X_t^0)) dt \xrightarrow{t \rightarrow \infty} \int_{\mathcal{S}} \phi \circ \xi_{\mathcal{S}} d\mu_\beta^0 = \int_{\mathbb{M}_m} \phi d\pi_\beta^0 = \int_{\mathbb{M}_m} \phi(z) e^{-\beta A_\star(z)} d\pi(z).$$

Similarly, when considering the biased dynamics,

$$\frac{1}{t} \int_0^t \phi(\xi_{\mathcal{S}}(X_t^A)) dt \xrightarrow{t \rightarrow \infty} \int_{\mathbb{M}_m} \phi(z) e^{-\beta(A_\star(z) - A(z))} d\pi(z).$$

We now give an interpretation of the qualitative properties of the free energy function  $A_\star$ . Assume that  $\pi$  is the Lebesgue measure on  $\mathbb{M}_m$ , and that  $A_\star$  admits several local minima: then the distribution  $\pi_\beta^0$  is multimodal, and the convergence to equilibrium, when using the unbiased dynamics, is slow. Indeed, the process must visit regions near all the local minima of  $A_\star$ , and hopping events between these metastable states are rare. Thus  $A_\star$  encodes the metastability of the dynamics with respect to the variable  $z = \xi(z) \in \mathbb{M}_m$ .

On the contrary, if the biased dynamics with  $A = A_\star$  is used, the associated ergodicity result indicates that convergence is expected to be faster – at least if the convergence in the other variables is not slow due to metastability. Indeed, the repartition of the values of  $\xi_{\mathcal{S}}(X_t^A)$  tends to be uniform when  $t \rightarrow \infty$ ; this is the flat-histogram property which is the guideline of advanced Monte Carlo strategy mentioned in Section 1.

Note also that, in many applications (for instance in molecular dynamics), computing free energy differences, *i.e.*  $A_\star(z_1) - A_\star(z_2)$ , may be the ultimate goal of the simulation, instead of computing averages  $\int \varphi d\mu_\beta$ . The Adaptive Biasing Potential methods of this article can also be seen as efficient Free Energy computation algorithms.

Since in general the free energy function is not known, the associated biased dynamics with  $A = A_\star$  cannot be simulated in practice; the guideline of the adaptive version proposed and analyzed below is to (approximately) reproduce the nice flat-histogram property for variable  $z = \xi(x)$  in the asymptotic regime  $t \rightarrow \infty$ , without a priori knowing the free energy function  $A_\star$ ; in fact it is also estimated in this procedure.

### 3. ABP: CONSTRUCTION AND WELL-POSEDNESS

In this section, the construction of the Adaptive Biasing Potential (ABP) system is performed in the general framework of Section 2. The rigorous construction of the process, and the statement of appropriate assumptions, is one of the contributions of this paper. The ABP system is built starting from the unbiased dynamics (3), with an adaptive bias  $A = A_t$  (random and depending on time  $t$ ) introduced in the biased dynamics (4). The construction is a generalization of (2), considered in Section 1.1 in a simplified setting.

In an abstract framework, the coupling of the evolutions of the diffusion process  $X_t$  and of the bias  $A_t$  requires the introduction of two auxiliary tools, which are used to define mappings  $\bar{\mu} \mapsto \mathcal{K}(\bar{\mu})(\cdot)$  (see (11)) and  $F \mapsto \bar{F}$  below:

- a kernel function  $K : \mathbb{M}_m \times \mathbb{M}_m \rightarrow (0, \infty)$ ;
- a normalization operator  $\mathcal{N} : \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) \rightarrow \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ , on the set of continuous functions on  $\mathbb{M}_m$  with values in  $(0, \infty)$ .

The ABP system in its general formulation is written as follows:

$$(10) \quad \begin{cases} dX_t = \mathcal{D}(V, A_t)(X_t)dt + \sqrt{2\beta^{-1}\Sigma}dW_t, \\ \bar{\mu}_t = \frac{\bar{\mu}_0 + \int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau))\delta_{X_\tau}d\tau}{1 + \int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau))d\tau}, \\ F_t = \mathcal{N}(\mathcal{K}(\bar{\mu}_t)), \\ A_t = -\frac{1}{\beta} \log(\bar{F}_t), \end{cases}$$

where there are four unknown processes:  $(X_t)_{t \geq 0}$  (with values in  $\mathcal{S}$ ),  $(\bar{\mu}_t)_{t \geq 0}$  (with values in  $\mathcal{P}(\mathcal{S})$  the set of probability distributions on  $\mathcal{S}$ ),  $(F_t)_{t \geq 0}$  (with values in  $\mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ ), and  $(A_t)_{t \geq 0}$  (with values in  $\mathcal{C}^\infty(\mathbb{M}_m)$ ). Note that the initial conditions  $F_0 = \mathcal{N}(\mathcal{K}(\bar{\mu}_0))$  and  $A_0 = -\frac{1}{\beta} \log(\bar{F}_0)$  are prescribed by the initial condition  $\bar{\mu}_0$ ; we also set  $X_0 = x_0$ .

Observe that it is not necessary to consider the four unknowns in (10). Indeed, as will be explained below,  $F_t$  and  $\bar{F}_t = \exp(-\beta A_t)$  only differ by a multiplicative constant (depending on  $t$ ), which is determined only by the choice of the normalization operator  $\mathcal{N}$ . Moreover, it would be possible to consider only the processes  $(X_t)_{t \geq 0}$  and  $(A_t)_{t \geq 0}$  to define the dynamics of the ABP system; however, we wish to emphasize on the role of the probability distribution  $\bar{\mu}_t$ , this is why it is included explicitly in (10).

Important observations concerning the system (10) are in order.

The diffusion process is biased, following (4), and the bias  $A_t$  at time  $t$  is defined in terms of the values  $(X_r)_{0 \leq r \leq t}$  of the diffusion process up to time  $t$ . As a consequence, the diffusion process in (10) can be considered as a self-interacting diffusion on  $\mathcal{S}$ . However, the standard framework of self-interacting processes does not encompass the system (10), and we thus need to adapt and generalize the arguments concerning well-posedness and convergence in our setting.

The function  $A_t$  is constructed in order to be an approximation, in the regime  $t \rightarrow \infty$ , of the Free Energy function  $A_*$ , introduced in Section 2.4; indeed, knowing  $A_*$  would lead to an optimal non-adaptive biased dynamics. The adaptive system is designed to approximate both adaptively and efficiently  $A_*$ .

As already mentioned in the introduction (see Theorem 1.1), the central object in the analysis is the probability distribution  $\bar{\mu}_t$ . Indeed, we will prove that it converges almost surely to  $\mu_\beta$ , see Theorem 4.1. Note that  $\bar{\mu}_t$  is defined as a weighted empirical distribution, with weights  $F_\tau(\xi_S(X_\tau))$ ; this choice is motivated by (8) (in the non-adaptive setting).

Below, we state assumptions on the kernel and on the normalization operator, which play a key role first for the well-posedness of the algorithms, second for the analysis of its asymptotic behavior. Assumptions on the model, from Section 2, are granted in the sequel.

**3.1. Kernel.** The kernel function  $K : (z, \zeta) \in \mathbb{M}_m \times \mathbb{M}_m \mapsto K(z, \zeta) \in (0, +\infty)$  is a continuous, positive, smooth function. In the following, this function is often referred to as the regularization kernel. Precisely, the following assumption is granted.

**Assumption 3.1.** *Assume that  $K$  is positive, of class  $\mathcal{C}^\infty$  on  $\mathbb{M}_m \times \mathbb{M}_m$ , and that  $\int_{\mathbb{M}_m} K(z, \zeta) \pi(dz) = 1$  for every  $\zeta \in \mathbb{M}_m$ .*

Since  $\mathbb{M}_m$  is compact, one has  $m(K) = \min_{z, \zeta \in \mathbb{M}_m} K(z, \zeta) > 0$ , and, for  $r \in \{0, 1, \dots\}$ ,  $M^{(r)}(K) = \sup_{z, \zeta \in \mathbb{M}_m} |\partial_z^r K(z, \zeta)| < +\infty$ . Moreover,  $\sup_{z \in \mathbb{M}_m} \sup_{\zeta_1, \zeta_2} \frac{K(z, \zeta_1) - K(z, \zeta_2)}{d(\zeta_1, \zeta_2)} < +\infty$  (Lipschitz continuity in the second variable, uniformly in the first variable).

The mapping  $\mathcal{K} : \bar{\mu} \in \mathcal{P}(\mathcal{S}) \mapsto \mathcal{K}(\bar{\mu}) \in \mathcal{C}^\infty(\mathbb{M}_m)$ , is then defined as follows:

$$(11) \quad \mathcal{K}(\bar{\mu})(z) = \int_{\mathcal{S}} K(z, \xi_S(x)) \bar{\mu}(dx).$$

Note that  $\int_{\mathcal{S}} \mathcal{K}(\bar{\mu})(z) \pi(dz) = 1$ , and that the mapping  $\mathcal{K}(\bar{\mu})$  is of class  $\mathcal{C}^\infty$ , thanks to Assumption 3.1.

Note also that (11) also makes sense if the probability distribution  $\bar{\mu}$  is replaced with a positive, finite, measure  $\mu$ .

The role played by the kernel  $K$  in the adaptive dynamics (10) is to define a smooth function  $A_t$  depending on the probability distribution  $\bar{\mu}_t$ , to be used as a bias in the dynamics like in (4).

One may consider the following example of kernel  $K$ , in the case  $\mathbb{M}_m = \mathbb{T}^m$ . Let  $k : \mathbb{R}^m \rightarrow (0, \infty)$  be an even function of class  $\mathcal{C}^\infty$ , with bounded derivatives, such that  $\int_{\mathbb{M}_m} k(z) \pi(dz) = 1$ . For  $\epsilon \in (0, 1)$ , let  $K(z, \zeta) = \frac{1}{\epsilon} k(\frac{z-\zeta}{\epsilon})$ . In the regime  $\epsilon \rightarrow 0$ , such kernels  $K = K_\epsilon$  are smooth mollifiers. If the function  $k$  is chosen with compact support, the positivity condition on  $K$  is satisfied by choosing  $K(z, \zeta) = \frac{\alpha}{\epsilon} k(\frac{z-\zeta}{\epsilon}) + 1 - \alpha$ , with  $\alpha \in (0, 1)$ .

It may also be useful to consider kernel functions which are not homogeneous, *i.e.*  $K(z, \zeta)$  does not depend only on  $z - \zeta$ . For instance, set  $K(z, \zeta) = \sum_{n=1}^N K_n(z, \zeta) \theta_n(\zeta)$ , where  $N \in \mathbb{N}$ ,  $K_1, \dots, K_N$  are kernel functions satisfying Assumption 3.1, and  $\theta_1, \dots, \theta_N$  are smooth functions  $\mathbb{M}_m \rightarrow (0, \infty)$ , such that  $\sum_{n=1}^N \theta_n(\zeta) = 1$  for all  $\zeta \in \mathbb{M}_m$ . Such examples are useful to build a bias which takes into account local properties.

Note that a symmetry assumption for the kernel –  $K(z, \zeta) = K(\zeta, z)$  – is not required to prove the consistency of the approach. For instance, assume that  $K(z, \zeta) = \tilde{K}(z)$  does not depend on  $\zeta$ ; in this case, one checks that  $\mathcal{K}(\bar{\mu}_t) = \mathcal{K}(\bar{\mu}_0) = \tilde{K}(\cdot)$  does not depend on  $t$ , and thus  $A_t = A_0$ : the adaptive system (10) reduces for this choice of kernel to the non-adaptive biased dynamics (4). Based on this observation, it is clear that the kernel  $K$  is the object which governs the coupling of the evolutions of  $X$  and  $A$  in the adaptive dynamics (10), and that its choice may be crucial in practice to define an efficient algorithm. In the sequel, we consider that a kernel function  $K$ ,

satisfying Assumption 3.1, is given, and do not study quantitatively the dependence with respect to  $K$  of the asymptotic results.

**3.2. Normalization.** The aim of this section is to introduce normalization operators, denoted by  $\mathcal{N} : \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) \rightarrow \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$  on the set of continuous functions from  $\mathbb{M}_m$  to  $(0, \infty)$ . The compactness of  $\mathbb{M}_m$  plays a crucial role again. We provide below several natural families of normalization operators. However, the presentation remains abstract to emphasize on the key assumptions which will lead to the stability estimates provided below.

We will use the following convention:  $f$  denotes an arbitrary element in  $\mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ , whereas  $F = \mathcal{N}(f)$  (capital letter) denotes its normalized version.

The most important example, for which a specific notation is introduced, is when normalization is meant to construct probability distributions  $\bar{f}d\pi$  which are equivalent to the reference measure  $\pi$  on  $\mathbb{M}_m$ :

$$\bar{f}(z) = \frac{f(z)}{\int_{\mathbb{M}_m} f(\zeta)\pi(d\zeta)}.$$

In the ABP system (10),  $\exp(-\beta A_t)$  is thus the density (with respect to  $\pi$ ) of a probability distribution on  $\mathbb{M}_m$ , for every  $t \geq 0$ .

More generally, the normalization operator  $\mathcal{N}$  is defined by

$$\mathcal{N}(f) = \frac{f}{\mathfrak{n}(f)},$$

where  $\mathfrak{n} : \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) \rightarrow (0, \infty)$  is a function which satisfies the technical (but easy to check in practice) conditions presented below.

**Assumption 3.2.** *Assume that  $\mathfrak{n} : \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) \rightarrow (0, \infty)$  satisfies the following conditions.*

- *There exists a sequence  $(\mathfrak{n}^{(k)})_{k \in \mathbb{N}}$ , such that, for every  $k \in \mathbb{N}$ ,  $\mathfrak{n}^{(k)} : \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) \rightarrow (0, \infty)$  is continuously differentiable, and for every  $f \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ ,*

$$\mathfrak{n}^{(k)}(f) \xrightarrow[k \rightarrow \infty]{} \mathfrak{n}(f);$$

*moreover the convergence is assumed to be uniform on sets of the form*

$$\{f \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty)) ; \min f \geq m, \max f \leq M\},$$

*for every  $0 < m \leq M < \infty$ .*

- *There exists  $\gamma_{\mathfrak{n}} \in (0, \infty)$  such that for all  $f \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$  and  $k \in \mathbb{N}^*$*

$$\frac{1}{\gamma_{\mathfrak{n}}} \min f \leq \mathfrak{n}^{(k)}(f) \leq \gamma_{\mathfrak{n}} \max f.$$

- *For all  $f \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ ,  $\alpha \in (0, \infty)$  and  $k \in \mathbb{N}^*$*

$$\mathfrak{n}^{(k)}(\alpha f) = \alpha \mathfrak{n}^{(k)}(f).$$

- *There exists  $C_{\mathfrak{n}} \in (0, \infty)$  such that for all  $f_1, f_2 \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$  and  $k \in \mathbb{N}^*$*

$$|\mathfrak{n}^{(k)}(f_1) - \mathfrak{n}^{(k)}(f_2)| \leq C_{\mathfrak{n}} \max |f_1 - f_2|.$$

Only the continuous differentiability condition is relaxed when considering the limit  $k \rightarrow \infty$ :  $\mathfrak{n}$  is not required to satisfy this condition. The three other conditions are satisfied when  $\mathfrak{n}^{(k)}$  is replaced with  $\mathfrak{n}$ .

Let us provide some important consequences of the definition of  $\mathcal{N}$  in terms of an operator  $\mathfrak{n}$  satisfying Assumption 3.2. First, note that  $\mathcal{N} \circ \mathcal{N} = \mathcal{N}$ : the normalization operator is a projection.

Moreover,  $F = \mathcal{N}(f)$  and  $f$  are equal up to a multiplicative constant; more generally, for two different normalization operators  $\mathcal{N}_1$  and  $\mathcal{N}_2$ , and any function  $f$ , the normalized versions  $F_1 = \mathcal{N}_1(f)$  and  $F_2 = \mathcal{N}_2(f)$  are equal up to a multiplicative constant. In particular,  $F = \mathcal{N}(\bar{F})$ , and thus in the ABP system (10), the weights  $F_\tau(\xi_S(X_\tau))$  are not necessarily equal to  $\exp(-\beta A_\tau(\xi_S(X_\tau)))$  like in (2) from the introduction; however it is important to have a fixed normalization operator, since by the second condition in Assumption 3.2 the ratio between these quantities remains bounded from below and from above by positive constants.

We conclude this section with additional examples of normalization operators.

- Let  $q \in [1, \infty)$ , and define

$$n_q(f) = \left( \int_{\mathbb{M}_m} f(z)^q \pi(dz) \right)^{\frac{1}{q}}.$$

In the case  $q = 1$ , we recover the example introduced above:  $\mathcal{N}_1(f) = \frac{f}{n_1(f)} = \bar{f}$ .

- Let  $z_0 \in \mathbb{M}_m$ , then define

$$n^{z_0}(f) = f(z_0) = \int_{\mathbb{M}_m} f(z) \delta_{z_0}(dz).$$

- Let also

$$n_{\min}(f) = \min_{z \in \mathbb{M}_m} f(z) \quad , \quad n_{\max}(f) = \max_{z \in \mathbb{M}_m} f(z).$$

For these examples, the relaxation of the continuous differentiability condition in Assumption 3.2 is essential: continuously differentiable approximations are given by

$$n_{\min}(f) = \lim_{q \rightarrow +\infty} \frac{1}{n_q(1/f)} \quad , \quad n_{\max}(f) = \lim_{q \rightarrow +\infty} n_q(f).$$

**3.3. Well-posedness.** The aims of this section are, first, to provide a well-posedness statement for the ABP system (10); second, to provide important stability properties for the processes  $X = (X_t)_{t \geq 0}$  and  $F = (F_t)_{t \geq 0}$ . We only provide a sketch of the proof of the existence-uniqueness of the solution, since it follows from a standard application of the Picard iteration scheme. One of the key arguments is Lemma 3.5 below.

**Theorem 3.3** (Well-posedness of (10)). *Grant assumptions of Section 2 concerning the model, and assumptions of Section 3 concerning the algorithm.*

- There exists a unique process  $t \in [0, \infty) \mapsto (X_t, \bar{\mu}_t, F_t, A_t)$ , taking values in the space  $\mathcal{C}(\mathbb{R}^+, \mathcal{S} \times \mathcal{P}(\mathcal{S}) \times \mathcal{C}^0(\mathbb{M}_m, (0, \infty))^2)$ , which is solution of the ABP system (10).
- For all  $n \geq 2$ ,  $\sup_{t \geq 0} \mathbb{E}|X_t|^n < +\infty$ .
- There exist  $m \in (0, \infty)$  and  $(M^{(r)})_{r \in \{0, 1, \dots\}} \in (0, \infty)$  such that for all  $t \in \mathbb{R}^+$   $F_t \in \mathcal{F}$  and  $A_t \in \mathcal{A}$ , almost surely, where

$$(12) \quad \begin{cases} \mathcal{F} = \{F \in \mathcal{C}^\infty(\mathbb{M}_m); \min F \geq m, \max |\partial^k F| \leq M^{(k)}, k \geq 0\}, \\ \mathcal{A} = \left\{ A = \frac{-1}{\beta} \log(\bar{F}); F \in \mathcal{F} \right\}. \end{cases}$$

The parameters  $m$  and  $M^{(k)}$  are not random; they depend only on the kernel  $K$ , on the normalization operator  $n$ , and on the initial condition  $\bar{\mu}_0$ . The notation  $\partial^k$  represents the derivative of order  $k$ .

Theorem 3.3 is a justification that the problem we are considering is well-posed. The asymptotic regime  $t \rightarrow \infty$  is studied later on. In particular, ergodicity properties of the unbiased and biased dynamics are not required.



**Remark 3.4.** In the proof of Theorem 3.3, the way the drift coefficient  $\mathcal{D}(V, A)$  depends on  $A$ , needs to be taken into account. We refrain from stating a general result, instead we consider only the examples of Section 2.2 – where  $\mathcal{D}(V, A)$  depends on the gradient  $\nabla A$  of  $A$ .

The proof of the third part of Theorem 3.3 (stability estimates) is based on the following result.

**Lemma 3.5.** Let  $m = \frac{\min(\min h_0, m(K))}{\gamma_n \max(\max h_0, M^{(0)}(K))}$  and  $M^{(k)} = \frac{\max(\max h_0, m^{(k)}(K))\gamma_n}{\max(\min h_0, m(K))}$ , for  $k \in \{0, 1, \dots\}$ , where  $h_0 = \mathcal{K}(\bar{\mu}_0)$ ,  $m(K)$ ,  $M^{(k)}(K)$  are given by Assumption 3.1, and  $\gamma_n$  is given by Assumption 3.2. Let  $\tau \mapsto x_\tau \in \mathcal{S}$  and  $\tau \mapsto F_\tau \in \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$  be continuous mappings, such that  $n(F_\tau) = 1$  for all  $\tau \in \mathbb{R}^+$ . Define

$$\mu_t = \bar{\mu}_0 + \int_0^t F_\tau(\xi_{\mathcal{S}}(x_\tau))\delta_{x_\tau} d\tau \quad , \quad h_t = \mathcal{K}(\mu_t) \quad , \quad H_t = \mathcal{N}(h_t).$$

Then, for all  $t \in \mathbb{R}^+$ ,  $k \in \mathbb{N}$ ,  $z \in \mathbb{M}_m$ ,

$$m \leq H_t(z) \leq M^{(0)} \quad , \quad |\partial^k H_t(z)| \leq M^{(k)}.$$

The Assumptions 3.1 and 3.2 are used explicitly in the proof; moreover, the compactness of  $\mathbb{M}_m$  is also a crucial property. The proof of Lemma 3.5 is elementary, however we provide some details since it is important in the sequel.

*Proof.* We only prove the estimates on  $\min H_t$  and  $\max H_t$ , since the proof of the estimates on the derivatives is similar. Note that

$$h_t(z) = h_0(z) + \int_0^t K(z, \xi_{\mathcal{S}}(x_\tau))F_\tau(\xi(x_\tau))d\tau,$$

where  $h_0 = \mathcal{K}(\bar{\mu}_0)$ , resp.  $K$ , are positive and continuous on  $\mathbb{M}_m$ , resp.  $\mathbb{M}_m \times \mathbb{M}_m$ . Thus for all  $t \in \mathbb{R}^+$

$$\begin{aligned} \min_{z \in \mathbb{M}_m} h_t(z) &\geq \min(\min h_0, m(K)) \left(1 + \int_0^t F_\tau(\xi_{\mathcal{S}}(x_\tau))d\tau\right) \\ \max_{z \in \mathbb{M}_m} h_t(z) &\leq \max(\max h_0, M^{(0)}(K)) \left(1 + \int_0^t F_\tau(\xi_{\mathcal{S}}(x_\tau))d\tau\right). \end{aligned}$$

Then the claim follows since  $H_t = \frac{h_t}{n(h_t)}$ , and using the second condition in Assumption 3.2.  $\square$

We are now in position to define the Picard iteration scheme, which gives existence and uniqueness of the solution of the system (10). First, note that only the variables  $X_t$  and  $F_t$  are considered in (10), and that it is not necessary to deal explicitly with the other ones. The choice of considering  $X_t$  and  $F_t$  is motivated by Lemma 3.5.

We are able to treat arbitrary time intervals  $[0, T]$ , by an appropriate choice of norm to construct a contraction. For an arbitrary  $T \in (0, \infty)$ , define the mapping  $\Psi^T$  as follows: let  $(X, F) \in L^2(\Omega, \mathcal{C}([0, T], \mathcal{S})) \times L^2(\Omega, \mathcal{C}([0, T], \mathcal{F}))$ , and define  $\Psi^T(X, F) = (Z, H)$  with

$$\begin{aligned} Z_t &= x + \sqrt{2\beta^{-1}}W_t + \int_0^t \mathcal{D}(V, A_\tau)(X_\tau)d\tau, \quad A_\tau = -\frac{1}{\beta} \log(\bar{F}_\tau), \\ \mu_t &= \bar{\mu}_0 + \int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau))\delta_{X_\tau} d\tau \quad , \quad H_t = \mathcal{N}(\mathcal{K}(h_t)). \end{aligned}$$

Thanks to Lemma 3.5,  $H_t \in \mathcal{F}$  for all  $t \in [0, T]$ , more precisely  $H \in L^2(\Omega, \mathcal{C}([0, T], \mathcal{F}))$ ; the property  $Z \in L^2(\Omega, \mathcal{C}([0, T], \mathcal{S}))$  comes from the uniform estimates over  $A \in \mathcal{A}$ , from Property 2.5 in non compact cases  $E_d = \mathbb{R}^d$ , and from standard computations.

For  $\alpha \in [0, \infty)$ , introduce the distance  $d_{\alpha, T}$  on  $L^2(\Omega, \mathcal{C}([0, T], \mathcal{S})) \times L^2(\Omega, \mathcal{C}([0, T], \mathcal{F}))$  as follows:

$$d_{\alpha, T}((x^1, f^1), (x^2, f^2)) = \left\| \sup_{0 \leq t \leq T} e^{-\alpha t} d_{\mathcal{S}}(x_t^1, x_t^2) \right\|_{L^2(\Omega)} + \left\| \sup_{0 \leq t \leq T} e^{-\alpha t} d_{\mathcal{F}}(f_t^1, f_t^2) \right\|_{L^2(\Omega)},$$

where  $d_{\mathcal{S}}$  (resp.  $d_{\mathcal{F}}$ ) is the natural distance on  $\mathcal{S}$  (resp.  $\mathcal{F}$ , induced by the usual distance on  $\mathcal{C}^\infty(\mathbb{M}_m, \mathbb{R})$ ). Note that the distances  $d_{\alpha, T}$  and  $d_{0, T}$  are equivalent, for all  $\alpha \in (0, \infty)$ .

Standard computations then yield the following result. The proof is omitted.

**Proposition 3.6.** *There exists  $C_{\text{sol}} \in (0, \infty)$ , such that for all  $\alpha \in (0, \infty)$ , for all  $T \in (0, \infty)$ , the mapping  $\Psi^T$  is Lipschitz continuous, with respect to the distance  $d_{\alpha, T}$ , with Lipschitz constant less than  $\frac{C_{\text{sol}}}{\alpha}$ :*

$$d_{\alpha, T}(\Psi^T(X^1, F^1), \Psi^T(X^2, F^2)) \leq \frac{C_{\text{sol}}}{\alpha} d_{\alpha, T}((X^1, F^1), (X^2, F^2)).$$

Choosing  $\alpha > C_{\text{sol}}$  gives a contraction in the Banach space  $L^2(\Omega, \mathcal{C}([0, T], \mathcal{S})) \times L^2(\Omega, \mathcal{C}([0, T], \mathcal{F}))$ , and thus proves the existence of a unique solution to the fixed point equation on  $[0, T]$ . Since  $T \in (0, \infty)$  is arbitrary, in fact (10) admits a unique global solution, well-defined on  $\mathbb{R}^+$ .

This construction proves the first part of Theorem 3.3, as well as the third part (combined with Lemma 3.5). The second part is straightforward in the case of Brownian or extended dynamics in the periodic case  $E_d = \mathbb{R}^d$ . In other cases, it is necessary to rely on Assumptions 2.2 and 2.4, and more precisely on Proposition 2.5, since  $A_t \in \mathcal{A} = \left\{ -\frac{1}{\beta} \log(\bar{F}) ; F \in \mathcal{F} \right\}$  thanks to Lemma 3.5. We omit these standard computations.

These arguments conclude the proof of Theorem 3.3.

## 4. CONVERGENCE RESULTS

This section contains the main results of this article, concerning the asymptotic behavior, when  $t \rightarrow \infty$ , of the solution of the ABP system (10). We first study consistency of the approach, and then its efficiency. The most important result dealing with consistency is Theorem 4.1: it states almost sure convergence of averages  $\bar{\mu}_t(\varphi)$  to  $\mu_\beta(\varphi)$  (where  $\mu_\beta = \mu_\beta^0$ , see (5)).

Section 4.4 is devoted to an interpretation of the ABP system (10) as an Adaptive Biasing Force method, and to the interpretation of the consistency results presented here in this context.

In the remainder of this section, all the Assumptions from Section 2, on the model, and of Section 3, on the algorithm, are granted. In particular, Theorem 3.3 ensures that the ABP system (10) is well defined. Moreover, the state space  $\mathcal{S}$  is finite dimensional.

### 4.1. Consistency of ABP.

4.1.1. *Convergence of weighted empirical averages.* The main result of this article concerns the consistency of the approach, for estimating averages  $\mu_\beta(\varphi)$  using weighted empirical averages  $\bar{\mu}_t(\varphi)$  (defined by (10)).

**Theorem 4.1.** *Let  $\varphi \in \mathcal{C}^\infty(\mathcal{S}, \mathbb{R})$  be a bounded function, with bounded derivatives of any order. Then, almost surely,*

$$\bar{\mu}_t(\varphi) \xrightarrow[t \rightarrow \infty]{} \mu_\beta(\varphi).$$

This result is a generalization in the adaptive case of (9). The proof of Theorem 4.1 requires the introduction of auxiliary tools, and is provided in Section 5. Several straightforward consequences of Theorem 4.1 are stated and proved in the next sections.

4.1.2. *Consequences of Theorem 4.1.* The first consequence of Theorem 4.1 is the almost sure convergence of the probability distribution  $\bar{\mu}_t$  to  $\mu_\beta$ , in the sense of the weak convergence of probability distributions.

**Corollary 4.2.** *We have the almost sure convergence*

$$\bar{\mu}_t \xrightarrow[t \rightarrow \infty]{} \mu_\beta.$$

*Precisely, almost surely,  $\bar{\mu}_t(\varphi) \xrightarrow[t \rightarrow \infty]{} \mu_\beta(\varphi)$  for every bounded continuous function  $\varphi : \mathcal{S} \rightarrow \mathbb{R}$ .*

*Proof.* We first state an auxiliary result: for every  $\varphi : \mathcal{S} \rightarrow \mathbb{R}$ , bounded and Lipschitz continuous function, almost surely

$$\bar{\mu}_t(\varphi) \xrightarrow[t \rightarrow \infty]{} \mu_\beta(\varphi), \text{ almost surely.}$$

The proof is standard, applying Theorem 4.1 for an approximating sequence  $\varphi_\epsilon = \rho_\epsilon \star \varphi$ , defined by convolution with smooth functions  $\rho_\epsilon(\cdot) = \frac{1}{\epsilon} \rho_1(\frac{\cdot}{\epsilon})$ , where  $\rho$  is of class  $\mathcal{C}^\infty$ , with compact support, and  $\int_{\mathcal{S}} \rho d\lambda = 1$ . Indeed,  $\mathcal{S}$  is finite dimensional.

Let  $\text{BL}(\mathcal{S}, \mathbb{R}) = \{\varphi : \mathcal{S} \rightarrow \mathbb{R} ; \varphi \text{ bounded and Lipschitz continuous}\}$ . Then there exists a sequence of functions  $(\varphi_n)_{n \geq 0}$  defined from  $\mathcal{S}$  to  $\mathbb{R}$ , bounded and Lipschitz continuous, such that

$$\bar{\mu}_t \xrightarrow[t \rightarrow \infty]{} \mu_\beta \iff d(\bar{\mu}_t, \mu_\beta) \xrightarrow[t \rightarrow \infty]{} 0,$$

where

$$d(\mu^1, \mu^2) = \sum_{n=0}^{\infty} \frac{1}{2^n} \min(1, |\int_{\mathcal{S}} \varphi_n d\mu^1 - \int_{\mathcal{S}} \varphi_n d\mu^2|).$$

Thanks to the convergence result above, almost surely, for every  $n \geq 0$ ,  $\bar{\mu}_t(\varphi_n) \xrightarrow[t \rightarrow \infty]{} \mu_\beta(\varphi_n)$ , and thus  $d(\bar{\mu}_t, \mu_\beta) \xrightarrow[t \rightarrow \infty]{} 0$  almost surely.

This concludes the proof of Corollary 4.2. □

The following result deals with the almost sure convergence of the functions  $\bar{F}_t$  and  $A_t$ . Note that contrary to Theorem 4.1 and Corollary 4.2, the limits  $\bar{F}_\infty$  and  $A_\infty$  depend on the parameters of the algorithm, precisely on the kernel function  $K$ . Note that these almost sure limits are not random.

The convergence of  $A_t$  to  $A_\infty$ , which is close to the Free Energy function  $A_\star$  for well-chosen kernel functions, is one of the nice features of the ABP method, in particular when one is interested in computing free energy differences.

**Corollary 4.3.** *Define, for all  $z \in \mathbb{M}_m$ ,*

$$\begin{cases} \bar{F}_\infty(z) = \mu_\beta(K(z, \cdot)), \\ A_\infty(z) = -\frac{1}{\beta} \log(\bar{F}_\infty(z)). \end{cases}$$

*Then, almost surely, for every  $\ell \in \{0, 1, \dots\}$ , uniformly on  $\mathbb{M}_m$ ,*

$$\begin{cases} \partial^\ell \bar{F}_t \xrightarrow[t \rightarrow \infty]{} \partial^\ell \bar{F}_\infty, \\ \partial^\ell A_t \xrightarrow[t \rightarrow \infty]{} \partial^\ell A_\infty. \end{cases}$$

*Proof.* The result is a consequence of the regularity properties of the kernel mapping  $K$ , of Ascoli's theorem, and of Theorem 4.1.

Let  $\mathcal{K} : \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{C}^\infty(\mathbb{M}_m)$  be the mapping defined by (11).

Let  $(z_n)_{n \in \mathbb{N}}$  denote a dense sequence in  $\mathbb{M}_m$ , and define, for all  $\mu^1, \mu^2 \in \mathcal{P}(\mathcal{S})$ ,

$$d_\infty(\mu^1, \mu^2) = \sum_{\ell, n=0}^{\infty} \frac{1}{2^{\ell+n}} \min(1, \left| \int_{\mathcal{S}} \partial_z^\ell K(z_n, \xi_{\mathcal{S}}(\cdot)) d\mu^1 - \int_{\mathcal{S}} \partial_z^\ell K(z_n, \xi_{\mathcal{S}}(\cdot)) d\mu^2 \right|).$$

Then for any sequence  $(\mu^k)_{k \in \mathbb{N}}$  and any  $\mu$  in  $\mathcal{P}(\mathcal{S})$ ,

- if  $\mu^k \xrightarrow[k \rightarrow \infty]{} \mu$  in  $\mathcal{P}(\mathcal{S})$ , then  $d_\infty(\mu^k, \mu) \xrightarrow[k \rightarrow \infty]{} 0$ ;
- if  $d_\infty(\mu^k, \mu) \xrightarrow[k \rightarrow \infty]{} 0$ , then for every  $\ell \in \{0, 1, \dots\}$ ,

$$\partial^\ell \mathcal{K}(\mu^k) \xrightarrow[k \rightarrow \infty]{} \partial^\ell \mathcal{K}(\mu),$$

uniformly on  $\mathbb{M}_m$ , thanks to Ascoli's theorem and the bound  $\|\partial_z^{k+1} K\|_\infty \leq M^{(k+1)}(K)$ .

Thanks to Theorem 4.1, it is straightforward to conclude that almost surely

$$d_\infty(\bar{\mu}_t, \mu_\beta) \xrightarrow[t \rightarrow \infty]{} 0.$$

These arguments yield the convergence of  $\bar{F}_t$ . The convergence of  $A_t = -\frac{1}{\beta} \log(\bar{F}_t)$  is then obtained thanks to the almost sure lower bound from Theorem 3.3,

$$\min_{\mathbb{M}_m} \bar{F}_t \geq m > 0.$$

□

**4.2. Applications to the diffusion processes of Section 2.2.** The aim of this section is to specify, for each of the examples of diffusion processes from Section 2.2:

- the convergence result of Theorem 4.1, for well chosen test functions  $\varphi$ ;
- the expression of the limit  $\bar{F}_\infty = e^{-\beta A_\infty}$ , in terms of the kernel  $K$  and of the free energy function  $A_\star$ .

We introduce the probability distribution  $\mu_\beta^{\text{ref}}(dx) = \frac{e^{-\beta V(x)}}{\int_{E_d} e^{-\beta V(y)} dy} dx$  on  $E_d$ . Observe that in all the examples  $\mu_\beta^{\text{ref}}$  is the marginal of the distribution  $\mu_\beta$  with respect to its  $E_d$ -valued component (in the Brownian dynamics case,  $\mu_\beta^{\text{ref}} = \mu_\beta$ ). In fact, in these cases, one is in fact more interested in  $\mu_\beta^{\text{ref}}$  than in  $\mu_\beta$ .

We also denote by  $A_\star^{\text{ref}}$  the Free Energy function associated with the reaction coordinate  $\xi$  and the probability distribution  $\mu_\beta^{\text{ref}}$ : by definition,  $e^{-\beta A_\star}$  is the Radon-Nikodym derivative of the image of  $\mu_\beta^{\text{ref}}$  by  $\xi$ , with respect to the probability distribution  $\pi$  on  $\mathbb{M}_m$ .

Assume that the kernel  $K = K_\delta$  depends on  $\delta > 0$ , and is such that the probability distribution  $K_\delta(z, \zeta) \pi(dz) \pi(d\zeta)$  converges when  $\delta \rightarrow 0$ , to  $\delta_z(d\zeta) \pi(dz)$ . Then, when  $\delta \rightarrow 0$  (and also  $\epsilon \rightarrow 0$ , in the extended dynamics case), the expressions below prove that  $A_\infty$  is an approximation of  $A_\star^{\text{ref}}$ . We do not provide quantitative estimates.

4.2.1. *Brownian dynamics (Section 2.2.1).*

- Computation of averages: for every  $\varphi \in C^\infty(E_d, \mathbb{R})$ , bounded and with bounded derivatives, almost surely

$$\int \varphi d\mu_\beta^{\text{ref}} = \lim_{t \rightarrow \infty} \frac{1 + \int_0^t F_\tau(\xi(X_\tau)) \varphi(X_\tau) d\tau}{1 + \int_0^t F_\tau(\xi(X_\tau)) d\tau}.$$

- Free Energy function:

$$e^{-\beta A_\infty(\cdot)} = \bar{F}_\infty(\cdot) = \int_{E_d} K_\delta(\cdot, \xi(x)) \mu_\beta^{\text{ref}}(dx) = \int_{\mathbb{M}_m} K_\delta(\cdot, \zeta) e^{-\beta A_\star^{\text{ref}}(\zeta)} \pi(d\zeta).$$

In particular, Theorem 1.1, stated in Section 1.1 and taken from [3], is a consequence of Corollaries 4.2 and 4.3, in the simplified context.

4.2.2. *Langevin dynamics (Section 2.2.2)*. We use the notation  $X_t = (q_t, p_t)$ .

- Computation of averages: for every  $\varphi \in C^\infty(E_d, \mathbb{R})$ , bounded and with bounded derivatives, almost surely

$$\int \varphi d\mu_\beta^{\text{ref}} = \lim_{t \rightarrow \infty} \frac{1 + \int_0^t F_\tau(\xi(q_\tau))\varphi(q_\tau)d\tau}{1 + \int_0^t F_\tau(\xi(q_\tau))d\tau}.$$

- Free Energy function:

$$e^{-\beta A_\infty(\cdot)} = \overline{F}_\infty(\cdot) = \int_{E_d} K_\delta(\cdot, \xi(q))\mu_\beta^{\text{ref}}(dq) = \int_{\mathbb{M}_m} K_\delta(\cdot, \zeta)e^{-\beta A_\star^{\text{ref}}(\zeta)}\pi(d\zeta).$$

Observe that the free energy function  $A_\infty$  is the same for the Brownian and the Langevin dynamics. This identity is in fact obtained since  $\xi_S(q, p) = \xi(q)$  only depends on  $q \in E_d$ .

4.2.3. *Extended dynamics (Section 2.2.3)*. We use the notation  $(X_t, Z_t)$ . Recall that  $\xi_S(x, z) = z$  in this case.

- Computation of averages: for every  $\varphi \in C^\infty(E_d, \mathbb{R})$ , bounded and with bounded derivatives, almost surely

$$\int \varphi d\mu_\beta^{\text{ref}} = \lim_{t \rightarrow \infty} \frac{1 + \int_0^t F_\tau(Z_\tau)\varphi(X_\tau)d\tau}{1 + \int_0^t F_\tau(Z_\tau)d\tau}.$$

- Free Energy function:

$$\begin{aligned} e^{-\beta A_\infty(\cdot)} = \overline{F}_\infty(\cdot) &= \int_{E_d \times \mathbb{M}_m} K(\cdot, z)\mu_\beta(dx dz) \\ &= \int_{E_d \times \mathbb{M}_m} K(\cdot, z)K_\epsilon^{\text{ext}}(z, \xi(x))\mu_\beta^{\text{ref}}(dx)\pi(dz) \\ &= \int_{\mathbb{M}_m} \left( \int_{\mathbb{M}_m} K(\cdot, z)K_\epsilon^{\text{ext}}(z, \zeta)\pi(dz) \right) e^{-\beta A_\star^{\text{ref}}(\zeta)}\pi(d\zeta), \end{aligned}$$

where we have introduced the auxiliary kernel  $K_\epsilon^{\text{ext}} : \mathbb{M}_m \times \mathbb{M}_m \rightarrow (0, \infty)$ , such that  $\mu_\beta(dx dz) = K_\epsilon^{\text{ext}}(z, \xi(x))\mu_\beta^{\text{ref}}(dx)\pi(dz)$ : up to a multiplicative constant,  $K_\epsilon^{\text{ext}}(z, \zeta) = \exp(-\frac{1}{2\epsilon}|z - \zeta|^2)$ . Note that the expression of  $A_\infty$  is not the same as in the previous examples, due to the additional term in the definition of the extended potential energy function on  $E_d \times \mathbb{M}_m$ . However, when  $\epsilon \rightarrow 0$ ,  $A_\infty$  converges to  $A_\star^{\text{ref}}(\cdot, \beta)$ : this observation justifies the use of the extended dynamics in the context of free energy computations.

4.3. **Efficiency.** We now state and prove a series of results concerning the efficiency of the approach, first in a qualitative way, second with a more quantitative approach. Corollary 4.4 deals with the convergence of the non-weighted empirical distribution  $\overline{\rho}_t$ , defined by (13); it is a straightforward consequence of Corollary 4.2. Proposition 4.5 deals with the mean-square error, and identifies an asymptotic variance. Since the proof of Proposition 4.5 requires tools introduced in Section 5, we postpone its proof to Section 6.

In terms of the behavior of the occupation measure and of the asymptotic variance, the results stated below may be interpreted as follows: in the asymptotic regime  $t \rightarrow \infty$ , the Adaptive Biasing Potential method (10) performs in the same way as the non-adaptive Biasing Potential method (4), with the bias  $A = A_\infty$ .

Note that these results are asymptotic, when  $t \rightarrow \infty$ ; it would also be interesting to study quantitatively the rates of convergence, with respect to time  $t$ , for each of the results. This question is left for future works.

4.3.1. *Convergence of non-weighted empirical distributions.* In this section, we focus on the convergence of non-weighted empirical averages  $\bar{\rho}_t(\varphi)$ , where  $\bar{\rho}_t$  is the probability distribution on  $\mathcal{S}$  defined by

$$(13) \quad \bar{\rho}_t = \frac{\bar{\mu}_0 + \int_0^t \delta_{X_\tau} d\tau}{1+t}.$$

We refer to  $\bar{\rho}_t$  as the non-weighted empirical distribution, or as the occupation measure, associated with the diffusion process  $(X_t)_{t \geq 0}$  defined by (10). We have the following result.

**Corollary 4.4.** *Let  $\varphi \in C^\infty(\mathcal{S}, \mathbb{R})$  be a bounded function, with bounded derivatives of any order. Then*

$$(14) \quad \bar{\rho}_t(\varphi) \xrightarrow[t \rightarrow +\infty]{} \mu_\beta^{A_\infty}(\varphi), \text{ almost surely,}$$

where  $A_\infty = \lim_{t \rightarrow \infty} A_t$  (see Corollary 4.3), and  $\mu_\beta^{A_\infty}$  is given by (5).

Moreover, almost surely,  $\bar{\rho}_t \xrightarrow[t \rightarrow \infty]{} \mu_\beta^{A_\infty}$ .

The arguments below justify that Corollary 4.4 can be interpreted, qualitatively, as an efficiency property of the ABP method.

First, observe that considering the biased dynamics  $(X_t^A)_{t \geq 0}$  given by (4), and setting

$$\bar{\rho}_t^A = \frac{\bar{\mu}_0 + \int_0^t \delta_{X_\tau^A} d\tau}{1+t},$$

then almost surely  $\bar{\rho}_t^A(\varphi) \xrightarrow[t \rightarrow \infty]{} \mu_\beta^A(\varphi)$ . The limit in (14), when the adaptive dynamics is used, is the same as when using the non-adaptive dynamics (4), with  $A = A_\infty$ .

Second, observe that the image by the mapping  $\xi_S : \mathcal{S} \rightarrow \mathbb{M}_m$  of the probability distribution  $\mu_\beta^A$  has density with respect to  $\pi$  proportional to

$$\exp(\beta A - \beta A_\star(\cdot)).$$

This density is constant, equal to 1, when  $A = A_\star$ : this means that in the asymptotic limit  $t \rightarrow \infty$ , the values of  $\xi_S(X_t^{A_\star})$  are distributed according to the reference probability distribution  $\pi$ . On the contrary, when  $A = 0$ , the values of  $\xi_S(X_t^0)$  are distributed according to  $\pi_\beta^0 = e^{-\beta A_\star} d\pi$ .

Assume that  $\pi$  is the uniform distribution on  $\mathbb{M}_m = \mathbb{T}^m$ ; assume also that all the metastability of the system is encoded by the reaction coordinate  $\xi$ . If  $A_\star$  has several local minima, then  $\pi_\beta^0$  is a multimodal distribution, and the diffusion process  $(X_t^0)_{t \geq 0}$  is metastable, and does not efficiently sample all the state space. Thus the convergence of  $\bar{\rho}_t^0$  to  $\mu_\beta^0$  is expected to be slower than the convergence of  $\bar{\rho}_t^{A_\star}$  to  $\mu_\beta^{A_\star}$ . Indeed, the exploration of the metastable states tends to be uniform, when  $t \rightarrow \infty$ , when observed through the reaction coordinate mapping.

Since  $A_\infty$  is an approximation of the Free Energy function  $A_\star$ , for well-chosen kernel functions  $K$ , efficiency of the ABP method is justified by the observations above.

We now provide the proof of Corollary 4.4, with elementary arguments. The proof of the almost sure convergence of the probability distributions is obtained as in the proof of Corollary 4.2, therefore we only focus on the convergence of averages  $\bar{\rho}_t(\varphi)$ .

*Proof of Corollary 4.4.* Introduce the auxiliary measure

$$(15) \quad \rho_t = \frac{\bar{\mu}_0 + \int_0^t \delta_{X_\tau} d\tau}{1 + \int_0^t F_\tau \circ \xi_S(X_\tau) d\tau} = \frac{(1+t)\bar{\rho}_t}{1 + \int_0^t F_\tau \circ \xi_S(X_\tau) d\tau}.$$

Since the measures  $\rho_t$  and  $\bar{\rho}_t$  only differ by a multiplicative (normalization) constant, one has the identity  $\bar{\rho}_t = \frac{\rho_t}{\rho_t(1)}$ . Then, note that

$$\begin{aligned}\rho_t(\varphi) &= \frac{\bar{\mu}_0(\varphi) + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) \frac{\varphi(X_\tau)}{F_\tau \circ \xi_{\mathcal{S}}(X_\tau)} d\tau}{1 + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) d\tau} \\ &= \frac{\bar{\mu}_0(\varphi) + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) \frac{\varphi(X_\tau)}{F_\infty \circ \xi_{\mathcal{S}}(X_\tau)} d\tau}{1 + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) d\tau} \\ &\quad + \frac{1}{1 + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) d\tau} \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) \varphi(X_\tau) \left( \frac{1}{F_\tau \circ \xi_{\mathcal{S}}(X_\tau)} - \frac{1}{F_\infty \circ \xi_{\mathcal{S}}(X_\tau)} \right) d\tau \\ &= \bar{\mu}_t \left( \frac{\varphi}{F_\infty \circ \xi_{\mathcal{S}}} \right) + o(1),\end{aligned}$$

using Cesaro's Lemma. Indeed, one has the almost sure lower bound  $1 + \int_0^t F_\tau \circ \xi_{\mathcal{S}}(X_\tau) d\tau \geq 1 + mt$ ; moreover thanks to Corollary 4.3,  $F_t = \mathcal{N}(\bar{F}_t) \xrightarrow[t \rightarrow +\infty]{} F_\infty = \mathcal{N}(\bar{F}_\infty)$ , uniformly on  $\mathbb{M}_m$ , almost surely.

Moreover, the function  $\frac{\varphi}{F_\infty \circ \xi_{\mathcal{S}}}$  is bounded and of class  $\mathcal{C}^\infty$ , with bounded derivatives (using  $\min F_\infty \geq m > 0$  thanks to Theorem 3.3). Applying Theorem 4.1, almost surely

$$\begin{aligned}\bar{\rho}_t(\varphi) &= \frac{\rho_t(\varphi)}{\rho_t(1)} \xrightarrow[t \rightarrow +\infty]{} \frac{\mu_\beta(\varphi/F_\infty(\xi_{\mathcal{S}}(\cdot)))}{\mu_\beta(1/F_\infty(\xi_{\mathcal{S}}(\cdot)))} = \frac{\mu_\beta(\varphi/\bar{F}_\infty(\xi_{\mathcal{S}}(\cdot)))}{\mu_\beta(1/\bar{F}_\infty(\xi_{\mathcal{S}}(\cdot)))} \\ &= \frac{\int_{\mathcal{S}} \varphi(x) \exp(-\beta(\mathcal{E}(V)(x) - A_\infty(\xi_{\mathcal{S}}(x)))) \lambda(dx)}{\int_{\mathcal{S}} \exp(-\beta(\mathcal{E}(V)(x) - A_\infty(\xi_{\mathcal{S}}(x)))) \lambda(dx)} \\ &= \frac{\int_{\mathcal{S}} \varphi(x) \exp(-\beta(\mathcal{E}(V, A_\infty)(x))) \lambda(dx)}{\int_{\mathcal{S}} \exp(-\beta(\mathcal{E}(V, A_\infty)(x))) \lambda(dx)} \\ &= \mu_\beta^{A_\infty}(\varphi),\end{aligned}$$

thanks to the identity (6), and to (5). This concludes the proof.  $\square$

4.3.2. *Asymptotic mean-square error.* This section is devoted to a more quantitative approach, concerning the behavior when  $t \rightarrow \infty$  of the mean-square error

$$\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2,$$

for functions  $\varphi \in \mathcal{C}^\infty(\mathcal{S}, \mathbb{R})$ , bounded and with bounded derivatives.

In order to compare the performance of the adaptive and non-adaptive versions of the biasing potential approach, introduce the following quantity

$$\mathbf{V}_\infty(\varphi, A) = \limsup_{t \rightarrow \infty} t \mathbb{E}|\bar{\mu}_t^A(\varphi) - \mu_\beta(\varphi)|^2 \in [0, \infty],$$

where  $A : \mathbb{M}_m \rightarrow \mathbb{R}$  is fixed,  $\bar{\mu}_t^A(\varphi)$  is the estimator of  $\mu_\beta(\varphi)$  defined by the left-hand side of (9), for every  $t \geq 0$ , using the biased dynamics (4).

In Section 6, it will be proved that in fact

$$\mathbf{V}_\infty(\varphi, A) = \lim_{t \rightarrow \infty} t \mathbb{E}|\bar{\mu}_t^A(\varphi) - \mu_\beta(\varphi)|^2 \in (0, \infty)$$

is a non-degenerate limit.

The following result, concerning the asymptotic mean-square error of the estimator  $\bar{\mu}_t(\varphi)$  of  $\mu_\beta(\varphi)$ , constructed using the adaptively biased dynamics (10).

**Proposition 4.5.** *Let  $\varphi \in C^\infty(\mathcal{S}, \mathbb{R})$  be a bounded function, with bounded derivatives of any order. Then*

$$t\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 \xrightarrow{t \rightarrow \infty} \mathbf{V}_\infty(\varphi, A_\infty),$$

where  $A_\infty = \lim_{t \rightarrow \infty} A_t$  almost surely, see Corollary 4.3.

As already explained, the asymptotic mean-square error for the adaptive version is the same as for the non-adaptive version, where the bias is chosen as  $A = A_\infty$ . Note that the dependence of  $\mathbf{V}_\infty(\varphi, A)$  with respect to  $A$  depends a lot on the choice of the function  $\varphi$ ; therefore no optimality result is stated.

The proof of Proposition 4.5 is postponed to Section 6; explicit expressions for  $\mathbf{V}_\infty(\varphi, A)$ , in terms of the solutions of Poisson equations, are given there.

**4.4. Interpretation as an Adaptive Biasing Force method.** The aim of this section is to describe how the ABP system (10) can be written as Adaptive Biasing Force strategy: instead of updating a function  $A : \mathbb{M}_m \rightarrow \mathbb{R}$ , one updates its gradient. We discuss how to write and interpret the consistency results, in particular Corollary 4.3, in light of this connexion.

To keep notation simple, consider the framework of Section 1.1: the diffusion process is the Brownian dynamics on  $\mathbb{T}^d$ , and  $\xi(x_1, \dots, x_d) = x_1 \in \mathbb{T}$ , *i.e.*  $m = 1$ . Assume in addition that the kernel  $K$  is symmetric,  $K(z, \zeta) = K(\zeta, z)$ .

The Free Energy function  $A_\star$  satisfies the identity

$$A'_\star(x_1) = \int_{\mathbb{T}^{d-1}} (\partial_{x_1} V(x)) e^{-\beta(V(x) - A_\star(x_1))} dx_2 \dots dx_d = \mathbb{E}_{X \sim \mu_\beta} [\partial_{x_1} V(X) | X_1 = x_1],$$

where in the conditional expectation the random variable  $X$  is distributed according to  $\mu_\beta$ . This identity is the starting point for constructions of Adaptive Biasing Force (ABF) methods mentioned in Section 1. The quantity  $A'_\star(x_1)$  is often called the equilibrium mean force in direction  $x_1$ . To write  $\partial_{x_1} A_t(x_1)$  as a conditional expectation, it is necessary to consider an additional variable (similarly to the construction of the extended dynamics in Section 2.2.3).

On the one hand, observe that the equilibrium mean force can be written as the conditional expectation

$$\partial_z A_\star(z) = \mathbb{E}_{(X, Z) \sim \eta_\star} [\partial_{x_1} V(X) | Z = z],$$

where  $(X, Z) \sim \eta_\star(dx, dz) = \mathbf{1}_{z=x_1} \mu_\beta(dx) dz$ . On the other hand, for any  $t \geq 0$ ,

$$A'_t(z) = -\beta^{-1} \frac{\int_{\mathbb{T}^{d-1}} \partial_z K(z, x_1) \bar{\mu}_t(dx)}{\int_{\mathbb{T}^{d-1}} K(z, x_1) \bar{\mu}_t(dx)} = \mathbb{E}_{(X, Z) \sim \eta_t} \left[ -\beta^{-1} \frac{\partial_z K(Z, \xi(X))}{K(Z, \xi(X))} \Big| Z = z \right],$$

where  $(X, Z) \sim \eta_t(dx, dz) = K(z, \xi(x)) \bar{\mu}_t(dx) dz$  is a probability distribution on  $\mathbb{T}^d \times \mathbb{T}$ , which depends on the kernel function  $K$ . Letting  $t \rightarrow \infty$ , thanks to Corollary 4.2, it is straightforward to check that  $\eta_t$  converges almost surely to  $\eta_\infty(dx, dz) = K(z, x_1) \mu_\beta(dx) dz$ . We thus obtain different expressions of  $A'_\infty(z)$ :

$$\begin{aligned} A'_\infty(z) &= \mathbb{E}_{(X, Z) \sim \eta_t} \left[ -\beta^{-1} \frac{\partial_z K(Z, \xi(X))}{K(Z, \xi(X))} \Big| Z = z \right] \\ &= -\beta^{-1} \frac{\int_{\mathbb{T}^{d-1}} \partial_z K(z, x_1) \mu_\beta(dx)}{\int_{\mathbb{T}^{d-1}} K(z, x_1) \mu_\beta(dx)} \\ &= \frac{\int_{\mathbb{T}^{d-1}} \partial_{x_1} V(x) K(z, x_1) \mu_\beta(dx)}{\int_{\mathbb{T}^{d-1}} K(z, x_1) \mu_\beta(dx)} \\ &= \mathbb{E}_{(X, Z) \sim \eta_\infty} [\partial_{x_1} V(X) | Z = z], \end{aligned}$$



thanks to the use of an integration by parts formula. A comparison of the expressions for  $A'_*(z)$  and  $A'_\infty(z)$  reveals that both quantities are written as conditional expectations of the same random variable  $\partial_{x_1}V(X)$ , conditional on  $Z = z$ , but for different distributions for the variables  $(X, Z)$ , respectively  $\eta_*$  and  $\eta_\infty$ . Note that the kernel function  $K$  plays a key role, and that adding the variable  $Z$  is necessary: indeed, both  $\eta_*$  and  $\eta_\infty$  have the marginal with respect to  $x$  equal to  $\mu_\beta$ .

The observation above may be the starting point for other types of Adaptive Biasing methods, based on a single realization of the stochastic process and a self-interaction mechanism using an empirical distribution.

As already explained, if  $K = K_\delta$  is constructed as a mollifier, when  $\delta \rightarrow 0$ ,  $\eta_\infty$  converges to  $\eta_*$ , and we recover the usual property that good approximation of the Free Energy function is obtained in the algorithm.

## 5. PROOF OF THEOREM 4.1

The aim of this section is to provide a detailed proof of Theorem 4.1.

First, in Sections 5.1.1 and 5.1.2, we present the strategy, and in particular we establish a connexion with the analysis of self-interacting diffusions from [4], and more generally of stochastic algorithms, see [2], [8], [21], [29]. More precisely, Section 5.1.1 presents a (random) change of time variable,  $s = \theta(t)$ , which transforms the weighted empirical distributions  $\bar{\mu}_t$  associated with the process  $X_t$ , into non-weighted empirical distributions  $\bar{\nu}_s$  associated with a process  $Y_s$ , with modified dynamics. In Section 5.1.2, we explain how the so-called ODE method can be exploited: the asymptotic behavior of  $\bar{\nu}_s$ , when  $s \rightarrow \infty$ , is related to the behavior of a differential equation of the type  $\dot{\nu} = -\nu + \Pi(\nu)$ . A crucial result, Proposition 5.2, states that  $\Pi(\nu) = \mu_\beta$  is a constant mapping, and the dynamics of the differential equation above is extremely simple.

The analysis is thus expected to be much simpler than in [4]. Indeed, in Section 5.2, we directly prove the almost sure convergence of  $\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)$  to 0 when  $t \rightarrow \infty$ . A substantial, technical, part, is devoted to the detailed analysis of the associated Poisson equation.

Even if it is not explicitly used in the technical part of the proof of Theorem 4.1, the description of the change of time variable strategy is included for pedagogical purpose. Moreover, in our opinion, it is an elegant way to justify the consistency of the approach. Moreover, it may be a useful strategy in other similar situations. Readers only interested in the proof of Theorem 4.1 may skip Sections 5.1.1 and 5.1.2 – except for Proposition 5.2 which is used in the sequel.

### 5.1. Approach from a stochastic approximation perspective.

5.1.1. *Change of time variable.* In this section, we introduce a random change of time variable, and describe some of its nice properties. This is only a mathematical tool, and does not need to be performed in practice when implementing the method. In addition, as explained above, this change of variable has only a pedagogical role, and will not be used in the technical details of the proof.

Consider the solution of the ABP system (10). Then the mapping  $t \mapsto \bar{\mu}_t \in \mathcal{P}(\mathcal{S})$  is the unique solution of the following Ordinary Differential Equation (ODE)

$$(16) \quad \frac{d\bar{\mu}_t}{dt} = \frac{\theta'(t)}{1 + \theta(t)} (\delta_{X_t} - \bar{\mu}_t), \quad \theta(t) = \int_0^t F_\tau(\xi_{\mathcal{S}}(X_\tau)) d\tau.$$

The ODE (16) is interpreted in the following weak sense: for every bounded continuous test function  $\varphi : \mathcal{S} \rightarrow \mathbb{R}$ , the real-valued function  $t \mapsto \bar{\mu}_t(\varphi) = \int_{\mathcal{S}} \varphi d\bar{\mu}_t \in \mathbb{R}$  is the unique solution of the differential equation

$$\frac{d\bar{\mu}_t(\varphi)}{dt} = \frac{\theta'(t)}{1 + \theta(t)} (\varphi(X_t) - \bar{\mu}_t),$$

with the initial condition  $\bar{\mu}_0(\varphi)$ .

Define the measure  $\mu_t = \bar{\mu}_0 + \int_0^t F_\tau(\xi_S(X_\tau))\delta_{X_\tau}d\tau$ . Then observe that  $\theta(t) = \mu_t(1)$  can be interpreted as a normalizing constant.

The presence of the random variable  $\theta(t)$  in the ODE (16) makes the analysis not trivial at first sight. However, we can get rid of this quantity by an appropriate change of time variable. As we will see, this procedure can be interpreted as a way to get rid of the weights in the definition of the measure  $\bar{\mu}_t$ . The dynamics of the stochastic process  $X_t$  will be modified accordingly.

Thanks to Theorem 3.3, there exist two deterministic real numbers  $0 < m \leq M$  such that almost surely  $\theta'(t) = F_t(\xi_S(X_t)) \in [m, M]$  for all  $t \geq 0$ . Moreover,  $\theta(0) = 0$ , and  $\theta(t) \geq mt \xrightarrow[t \rightarrow \infty]{} \infty$ . As a consequence, almost surely,  $\theta : [0, \infty) \rightarrow [0, \infty)$  is a  $\mathcal{C}^1$ -diffeomorphism, with inverse denoted by  $\theta^{-1}$ .

Define, for every  $s \geq 0$ ,  $\tilde{W}(s) = \int_0^{\theta^{-1}(s)} \sqrt{\theta(t)}dW(t)$ . Note that for every  $s \geq 0$ ,  $\theta^{-1}(s) = \inf \{t \geq 0 ; \theta(t) \geq s\}$  is a bounded stopping time, associated with the filtration generated by the Wiener process  $W$ . Then, it is straightforward to check that  $(\tilde{W}(s))_{s \geq 0}$  is a standard Wiener process on  $\mathcal{S}$ .

We introduce the following system:

$$(17) \quad \begin{cases} dY_s = \mathcal{D}(V, B_s)(Y_s) \frac{1}{G_s(\xi(Y_s))} ds + \sqrt{\frac{2\beta^{-1}}{G_s(\xi_S(Y_s))}} \Sigma d\tilde{W}_s, \\ \bar{\nu}_s = \frac{1}{1+s} (\bar{\mu}_0 + \int_0^s \delta_{Y_\sigma} d\sigma), \\ G_s = \mathcal{N}(\mathcal{K}(\bar{\nu}_s)), \\ B_s = -\frac{1}{\beta} \log(\bar{G}_s). \end{cases}$$

Then we have the following result, which can be proved with elementary computations: almost surely,

$$(18) \quad \begin{cases} X_t = Y_{\theta(t)} & , & \bar{\mu}_t = \bar{\nu}_{\theta(t)} & , & F_t = G_{\theta(t)} & , & A_t = B_{\theta(t)} & , & \forall t \geq 0 \\ Y_s = X_{\theta^{-1}(s)} & , & \bar{\nu}_s = \bar{\mu}_{\theta^{-1}(s)} & , & G_s = F_{\theta^{-1}(s)} & , & B_s = A_{\theta^{-1}(s)} & , & \forall s \geq 0. \end{cases}$$

In other words, one may consider the system (17) as the time-changed version of the original ABP system (10), with the new time variable  $s = \theta(t)$ , and the new unknowns  $Y_s$ ,  $\bar{\nu}_s$ ,  $G_s$  and  $B_s$ , replacing  $X_t$ ,  $\bar{\mu}_t$ ,  $F_t$  and  $A_t$ .

Observe that the weights  $F_t(\xi_S(X_t)) = G_s(\xi_S(Y_s))$  do not appear anymore in the definition of the measure  $\bar{\nu}_s$ . Instead, the weights appear in the dynamics of the diffusion process  $(Y_s)_{s \geq 0}$ . In terms of new variables, the ODE (16) has a simpler formulation:

$$(19) \quad \frac{d\bar{\nu}_s}{ds} = \frac{1}{1+s} (\delta_{Y_s} - \bar{\nu}_s).$$

We are interested in the convergence of  $\bar{\mu}_t$  (or  $\bar{\mu}_t(\varphi)$ ) when  $t \rightarrow \infty$ . Since  $\bar{\mu}_t = \bar{\nu}_{\theta(t)}$ , and  $\theta(t) \xrightarrow[t \rightarrow \infty]{} \infty$  almost surely, we now study the asymptotic behavior ( $s \rightarrow \infty$ ) of  $\bar{\nu}_s$ . In the remainder of this section, we work only with the system (17), and consider  $s$  as the natural time variable. From a practical point of view,  $s$  is only a fictive time variable, whereas  $t$  is the physical one.

Then observe that proving Theorem 4.1 is equivalent to proving that

$$\bar{\nu}_s(\varphi) \xrightarrow[s \rightarrow \infty]{} \mu_\beta(\varphi) , \text{ almost surely.}$$

Indeed,  $\bar{\mu}_t = \bar{\nu}_{\theta(t)}$ , and almost surely  $\theta(t) \xrightarrow[t \rightarrow \infty]{} \infty$ . In Section 5.1.2 below, we describe how the asymptotic behavior of  $\bar{\nu}_s$ , when  $s \rightarrow \infty$ , can be analyzed using the ODE method.

5.1.2. *Consistency via the ODE method.* Thanks to the change of time variable  $s = \theta(t)$  introduced above, the structure of the system (17) is closer to the formulation of self-interacting diffusions (see [4] for instance), depending on the normalized occupation measure, than for the initial system (17). However, in the specific situation considered in the present article, arguments need to be modified, in particular the coupling of the evolutions of the diffusion process and of the empirical distributions does not have the same structure (here it depends on the kernel  $K$ ).

Thanks to the ODE (19), observe that there is an asymptotic time scale separation (in the limit  $s \rightarrow \infty$ ) between slow variables  $\bar{v}_s, G_s$  and  $B_s$ , and fast variables  $Y_s$ . It is reasonable to focus on the asymptotic behavior of the diffusion process when the other variables are frozen; when its unique invariant distribution (in general depending on the frozen variables) is introduced in place of the Dirac mass in (19), a limit ODE is obtained: the rationale behind the ODE method is that its asymptotic behavior provides information on the asymptotic behavior of the solution of (19), in some situations.

The ODE method allows us to make rigorous the discussion above, and to identify the appropriate limit ODE. In this article, one of the main specific properties is that the invariant distribution of the fast equation with frozen variables is equal to  $\mu_\beta$ , the target probability distribution, and thus does not depend on the frozen variables.

**Remark 5.1.** *The asymptotic time scale separation (when  $t \rightarrow \infty$ ) between slow variables  $\bar{\mu}_t, F_t$  and  $A_t$ , and the fast variable  $X_t$ , already appears in the original system (2). The change of time variable  $s = \theta(t)$  allows us to get rid of the random quantity  $\theta(t)$ , and to easily identify the correct limit ODE.*

Precisely, for every  $G \in \mathcal{C}^\infty(\mathbb{M}_m, \mathbb{R}) \cap \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ , let  $(Y_s^G)_{s \geq 0}$  denote the diffusion process which is the unique solution of

$$(20) \quad dY_s^G = \frac{\mathcal{D}(V, B)(Y_s^G)}{G(\xi(Y_s^G))} ds + \sqrt{\frac{2\beta^{-1}}{G(\xi(Y_s^G))}} \Sigma d\tilde{W}_s,$$

where  $B = -\frac{1}{\beta} \log(\bar{G})$ . Even if there are three slow variables,  $\bar{v}, G$  and  $B$ , it is sufficient to indicate only one of them, namely  $G$ .

**Proposition 5.2.** *For every  $G \in \mathcal{C}^\infty(\mathbb{M}_m, \mathbb{R}) \cap \mathcal{C}^0(\mathbb{M}_m, (0, \infty))$ , the equation (20) admits a unique invariant probability distribution, equal to  $\mu_\beta$ .*

*Proof.* First note that  $G = \frac{\bar{G}}{n(G)} = \frac{\exp(-\beta B)}{n(G)}$  is equal to  $\exp(-\beta B)$  up to a multiplicative constant, and thus a probability distribution  $\mu$  is invariant for (20) if and only if it is invariant for

$$(21) \quad d\mathcal{Y}_s^B = \frac{\mathcal{D}(V, B)(\mathcal{Y}_s^B)}{e^{-\beta B(\xi(\mathcal{Y}_s^B))}} ds + \sqrt{\frac{2\beta^{-1}}{e^{-\beta B(\xi(\mathcal{Y}_s^B))}}} \Sigma d\tilde{W}_s.$$

Let  $\mathcal{L}_{\mathcal{Y}}^B$  denote the associated infinitesimal generator: then for every function  $\varphi \in \mathcal{C}^\infty(\mathcal{S}, \mathbb{R})$ ,

$$(22) \quad \mathcal{L}_{\mathcal{Y}}^B \varphi(y) = \frac{1}{e^{-\beta B(\xi_S(y))}} \mathcal{L}_X^B \varphi(y),$$

where  $\mathcal{L}_X^B$  is the infinitesimal generator of the biased diffusion process  $X^B$  defined by (4), with  $A = B$ .

Since the unique invariant probability distribution of (4) with  $A = B$  is  $\mu_\beta^B$ , the unique invariant probability distribution of (20) is proportional to

$$\begin{aligned} e^{-\beta B(\xi_S(y))} \mu_\beta^B(dy) &= e^{-\beta B(\xi_S(y))} \frac{\exp(-\beta \mathcal{E}(V, B)(y))}{Z^B(\beta)} \lambda(dy) \\ &= \frac{\exp(-\beta \mathcal{E}(V)(y))}{Z^B(\beta)} \lambda(dy) = \frac{Z^0(\beta)}{Z^B(\beta)} \mu_\beta(dy), \end{aligned}$$

using (5) (expression of expression of  $\mu_\beta^B$ ) and (6) (expression of  $\mathcal{E}(V, B)$ ).

Identifying the normalizing constants then concludes the proof of Proposition 5.2.  $\square$

The ODE method suggests then to introduce the following ODE:

$$(23) \quad \frac{d\gamma_s}{ds} = \frac{1}{1+s} (\Pi(\gamma_s) - \gamma_s) = \frac{1}{1+s} (\mu_\beta - \gamma_s).$$

Indeed, thanks to Proposition 5.2,  $\Pi(\gamma) = \mu_\beta$  is the unique invariant distribution of (20), where  $G = \mathcal{K}(\gamma)$ . This property justifies the consistency of the approach, *i.e.* the almost sure convergence of  $\bar{\nu}_s$  to  $\mu_\beta$ . Indeed, it is straightforward to check that, for any initial condition  $\gamma_0 \in \mathcal{P}(\mathcal{S})$ , one has

$$\gamma_s = \frac{1}{1+s} (\bar{\gamma}_0 + s\mu_\beta) \xrightarrow{s \rightarrow \infty} \mu_\beta.$$

Moreover, a rigorous connexion between the asymptotic behaviors of  $\bar{\nu}_s$  and of  $\gamma_s$  may be stated for instance using the notion of asymptotic pseudo-trajectories (see [2], [4]); or by proving direct estimates on the  $L^p$  norm of the random variable  $\bar{\nu}_s(\varphi) - \mu_\beta(\varphi)$ .

In Section 5.2 below, instead, we prove directly estimates on the  $L^p$  norm of the random variable  $\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)$ ; indeed, thanks to Proposition 5.2, the situation is rather simple and the error is analyzed using straightforward computations, combined with a powerful auxiliary tool: the use of the solutions of associated Poisson equation.

## 5.2. Analysis of the error and convergence.

5.2.1. *The error in terms of the solution of Poisson equations.* In order to prove that

$$\bar{\mu}_t(\varphi) - \mu_\beta(\varphi) = \frac{\int_0^t F_\tau(\xi_S(X_\tau)) [\varphi(X_\tau) - \mu_\beta(\varphi)] d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau}$$

converges to 0 when  $t \rightarrow \infty$ , it is standard to introduce a family of Poisson equations depending on the integrand on the numerator. Let  $\Phi : (s, y) \in [1, \infty) \times \mathcal{S} \mapsto \Phi(s, y) \in \mathbb{R}$  be a  $\mathcal{C}^{1,2}$  function, *i.e.* of class  $\mathcal{C}^1$  with respect to the variable  $s$  and of class  $\mathcal{C}^2$  with respect to the variable  $y$ , with bounded associated derivatives. The application of Itô's formula yields the equality

$$\begin{aligned} \Phi(t, X_t) - \Phi(0, X_0) &= \int_0^t \mathcal{L}_X^A(\tau, X_\tau) d\tau + \int_0^t \frac{\partial \Phi}{\partial \tau}(\tau, X_\tau) d\tau \\ &\quad + \int_0^t \sqrt{2\beta^{-1}} \langle \nabla \Phi(\tau, X_\tau), \Sigma dW(\tau) \rangle, \end{aligned}$$

where  $\mathcal{L}_X^A$  is the infinitesimal generator of the biased diffusion process  $X^A$ , see (4).

Assume that the function  $\Phi$  satisfies, for all  $t \geq 0$ ,  $x \in \mathcal{S}$

$$(24) \quad \mathcal{L}_X^A \Phi(t, X_t) = F_t(\xi_S(x)) [\varphi(x) - \mu_\beta(\varphi)];$$

then one obtains

$$\begin{aligned} \bar{\mu}_t(\varphi) - \mu_\beta(\varphi) &= \frac{\Phi(t, X_t) - \Phi(0, X_0)}{1 + \int_0^t F_\tau(\xi_S(X_\tau))d\tau} \\ &\quad - \frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla \Phi(\tau, X_\tau), \Sigma dW(\tau) \rangle}{1 + \int_0^t F_\tau(\xi_S(X_\tau))d\tau} \\ &\quad - \frac{\int_0^t \frac{\partial \Phi}{\partial \tau}(\tau, X_\tau)d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau))d\tau}. \end{aligned}$$

Recall that,  $1 + \int_0^t F_\tau(\xi_S(X_\tau))d\tau \geq mt$ , for all  $t \geq 0$ , almost surely, thanks to Theorem 3.3, with  $m > 0$ . Convergence of  $\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)$  to 0, in a  $L^p$  sense, then follows from appropriate estimates on the function  $\Phi$  and its derivatives.

We now focus on the question of existence of the function  $\Phi$ , and on the analysis of its properties. The condition (24) can be written as

$$(25) \quad \Phi(t, x) = \frac{1}{n(\bar{F}_t)} \Psi(A_t, x)$$

where, for any  $A \in \mathcal{A}$  (see (12)),  $\Psi(A, \cdot)$  is solution of the Poisson equation

$$(26) \quad \begin{cases} \mathcal{L}_x^A \Psi(A, \cdot) = e^{-\beta A(\xi_S(x))} [\varphi - \mu_\beta(\varphi)], \\ \int \Psi(A, \cdot) d\mu_\beta^A = 0. \end{cases}$$

The aim of the following section is to study well-posedness and some important estimates concerning the solution  $\Psi(A, \cdot)$  of the Poisson equation (26).

5.2.2. *Analysis of the Poisson equations* (26). Introduce the set

$$(27) \quad \mathcal{C} = \mathcal{C}_{\text{pol}}^\infty(\mathcal{S}, \mathbb{R}) = \left\{ \varphi \in \mathcal{C}^\infty(\mathcal{S}, \mathbb{R}) ; \forall k \in \mathbb{N}, \exists p_k \in \mathbb{N}, \sup_{x \in \mathcal{S}} \frac{|D^k \varphi(x)|}{1 + |x|^{p_k}} < \infty \right\}$$

of functions  $\varphi : \mathcal{S} \rightarrow \mathbb{R}$ , of class  $\mathcal{C}^\infty$ , with at most polynomial growth, and all derivatives with at most polynomial growth. Note that the average  $\mu_\beta(\varphi)$  is well-defined, since the probability distribution  $\mu_\beta$  admits finite moments of any order.

We first state and prove the following well-posedness result.

**Proposition 5.3.** *For every  $A \in \mathcal{A}$  and every  $\varphi \in \mathcal{C}$ , there exists a unique solution  $\Psi(A, \cdot) \in \mathcal{C}$  of the Poisson equation (26).*

*Proof of Proposition 5.3.* Define the auxiliary function  $\bar{\varphi}^A = e^{-A \circ \xi_S}(\varphi - \mu_\beta(\varphi))$ , and note that  $\bar{\varphi}^A \in \mathcal{C}$ .

Observe that, thanks to (8),

$$\int \bar{\varphi}^A d\mu_\beta^A = \int [\varphi - \mu_\beta(\varphi)] d\mu_\beta^0 = 0.$$

Thanks to this centering condition, we are in position to apply standard results concerning Poisson equations, either for elliptic – Brownian and extended dynamics – or hypoelliptic – Langevin dynamic – operators. We obtain that

$$(28) \quad \Psi(A, x) = - \int_0^\infty \mathbb{E}_x[\bar{\varphi}^A(X_t^A)] dt,$$

where  $X^A$  is the biased process given by (4).

This concludes the proof of Proposition 5.3. □

**Remark 5.4.** Note that, thanks to the identity (22), the Poisson equation (26) may be rewritten as

$$\mathcal{L}_Y^A \Psi(A, \cdot) = \varphi - \mu_\beta(\varphi).$$

Thanks to Proposition 5.2,  $\mu_\beta$  is the unique invariant distribution for the process (21) with infinitesimal generator  $\mathcal{L}_Y^A$ , for any choice of  $A \in \mathcal{A}$ . This interpretation gives another way of representing the solution of the Poisson equation:

$$\Psi(A, x) = - \int_0^\infty \mathbb{E}_x[\varphi(Y_s^A) - \mu_\beta(\varphi)] ds.$$

If one follows the strategy outlined in Sections 5.1.1 and 5.1.2, this way of writing the Poisson equation is more natural. By adapting the elements of proof below, using this formulation of the Poisson equation, it is clear that Proposition 5.2 and the ODE (23) justify the consistency.

We now provide several properties of the solutions  $\Psi(A, \cdot)$  which are required below. We emphasize on the property that bounds are uniform with respect to the variable  $A \in \mathcal{A}$ , the set defined by (12).

**Proposition 5.5.** Let  $\varphi \in \mathcal{C}$ , of class  $\mathcal{C}^\infty$ , with at most polynomial growth.

There exist  $C \in (0, \infty)$  and  $p \in \mathbb{N}^*$ , such that the following results hold true.

(i) For every  $A \in \mathcal{A}$  and every  $x \in \mathcal{S}$

$$(29) \quad |\Psi(A, x)| \leq C(1 + |x|^p).$$

(ii) For every  $A \in \mathcal{A}$  and every  $x \in \mathcal{S}$

$$(30) \quad |\nabla_x \Psi(A, x)| \leq C(1 + |x|^p).$$

(iii) The function  $(t, x) \in [0, \infty) \times \mathcal{S} \mapsto \Psi(A_t, x)$  is of class  $\mathcal{C}^{1,2}$ , and for every  $x \in \mathcal{S}$  and every  $t \geq 0$ , almost surely

$$(31) \quad \left| \frac{\partial \Psi(A_t, x)}{\partial t} \right| \leq \frac{C(1 + |x|^p)}{1 + t},$$

where  $(A_t)_{t \geq 0}$  is the  $\mathcal{A}$ -valued process defined in (10).

In the proof below, we only consider the case of functions  $\varphi$  which are bounded and have bounded derivatives of any order. The case of functions with polynomial growth (when the state space is not compact) requires the introduction of appropriate weight functions, under Assumptions 2.2 and 2.4. The proof is technical but standard, and we do not provide details. Note that the weight functions may be chosen uniformly with respect to  $A \in \mathcal{A}$ , thanks to Property 2.5, hence the estimates in items (i) and (ii) of Proposition 5.5 are uniform with respect to  $A \in \mathcal{A}$ .

*Proof of Proposition 5.5.*

(i) Thanks to Property 2.5, and standard techniques, there exist  $\gamma \in (0, \infty)$ ,  $C(\varphi) \in (0, \infty)$  and  $p \in \mathbb{N}^*$ , such that for every  $A \in \mathcal{A}$ , then for all  $x \in \mathcal{S}$  and  $t \geq 0$ , one has

$$(32) \quad |\mathbb{E}_x[\overline{\varphi}^A(X_t^A)]| \leq C(\varphi)e^{-\gamma t}(1 + |x|^p).$$

Integrating from  $t = 0$  to  $t = \infty$ , and using (28), gives (29).

(ii) Let  $(P_t^A)_{t \geq 0}$  denote the transition semi-group associated with the diffusion process  $(X_t^A)_{t \geq 0}$  on  $\mathcal{S}$ .

First, it is straightforward to check that, for  $t \in [0, 1]$ , one has

$$|\nabla_x P_t^A \overline{\varphi}^A(x)| \leq C(1 + |x|^p).$$

Second, for every  $t \geq 1$ ,

$$\nabla_x(P_t^A \bar{\varphi}^A) = \nabla_x\left(P_1^A(P_{t-1}^A \bar{\varphi}^A)\right);$$

then standard arguments, and the exponential convergence estimate (32), yield

$$|\nabla_x(P_t^A \bar{\varphi}^A)(x)| \leq C \sup_{\tilde{x} \in \mathcal{S}} \frac{|(P_{t-1}^A \bar{\varphi}^A)(\tilde{x})|}{1 + |\tilde{x}|^p} (1 + |x|^p) \leq C(\varphi) e^{-(t-1)} (1 + |x|^p).$$

Integrating separately from  $t = 0$  to  $t = 1$  and from  $t = 1$  to  $t = \infty$  gives (30).

Again the constants do not depend on  $A \in \mathcal{A}$ , thanks to (32) and Property 2.5.

(iii) Since  $F_t \in \mathcal{F}$  for all  $t \geq 0$ , almost surely, thanks to Theorem 3.3, then  $\min_{z \in \mathbb{M}_m} \bar{F}_t(z) \geq m$  for all  $t \geq 0$ , almost surely, for some  $m \in (0, \infty)$ .

Moreover,  $\bar{F}_t(z) = \bar{\mu}_t(K(z, \cdot))$ ; thanks to Assumption 3.1 and to the ODE (16), for every  $k \in \{0, 1, \dots\}$ , there exists  $C^{(k)} \in (0, \infty)$  such that

$$(33) \quad \sup_{z \in \mathbb{M}_m} \left| \frac{d(\partial^k A_t(z))}{dt} \right| \leq \frac{C^{(k)}}{1+t}.$$

For every  $t > 0$ , every  $\epsilon \in (-t, 1)$ , note that

$$\mathcal{L}_y^{A_t+\epsilon} \Psi(A_{t+\epsilon}, \cdot) - \mathcal{L}_y^{A_t} \Psi(A_t, \cdot) = 0,$$

thanks to (22), see Remark 5.4. Passing to the limit  $\epsilon \rightarrow 0$  yields

$$\begin{aligned} \mathcal{L}_y^{A_t} \frac{\partial \Psi(A_t, \cdot)}{\partial t} &= - \left( \frac{\partial \mathcal{L}_y^{A_t}}{\partial t} \right) \Psi(A_t, \cdot) \\ &= \left( \beta \frac{dA_t \circ \xi_S}{dt} \mathcal{L}_Y^{A_t} + \frac{d}{dt} \left( e^{\beta A_t \circ \xi_S} \langle \mathcal{D}(V, A_t), \nabla \cdot \rangle \right) \right) (\Psi(A_t, \cdot)). \end{aligned}$$

Considering each example for the definition of the drift function  $\mathcal{D}(V, B_s)$ , it is straightforward to check that  $-\left(\frac{d\mathcal{L}_y^{A_t}}{ds}\right)\Psi(A_t, \cdot) \in \mathcal{C}$  is a function of class  $\mathcal{C}^\infty$  with polynomial growth; and, moreover, that for every  $k \in \{0, 1, \dots\}$ , there exist  $p_k \geq 0$  and  $C^{(k)} \in (0, \infty)$  such that

$$\sup_{y \in \mathcal{S}} \left| D^k \left( \frac{\partial \mathcal{L}_y^{A_t}}{\partial t} \right) \Psi(A_t, \cdot)(x) \right| \leq \frac{C^{(k)}(1 + |x|^{p_k})}{1+t},$$

thanks to the inequality (33), and the estimate (30) on the gradient  $\nabla_x \Psi(A_t, x)$ .

Thanks to Proposition 5.3, one then concludes the proof of (31). □

**5.2.3. Proof of convergence.** We have introduced Poisson equations (26) in order to construct an auxiliary function  $(t, x) \mapsto \Phi(t, x)$  to solve condition (24), using (25).

However, due to the possible low regularity properties of the normalization operator  $\mathfrak{n}$  (see Assumption 3.2), we cannot in general define  $\Phi$  with (25). Instead, we rely on an approximation procedure, thanks to the sequence  $(\mathfrak{n}^{(k)})_{k \in \mathbb{N}}$  introduced in Assumption 3.2. We thus define, for every  $k \in \mathbb{N}$

$$\Phi^{(k)}(t, x) = \frac{1}{\mathfrak{n}^{(k)}(\bar{F}_t)} \Psi(A_t, x).$$

Then observe that that

$$\begin{aligned}
\bar{\mu}_t(\varphi) - \mu_\beta(\varphi) &= \frac{\int_0^t F_\tau(\xi_S(X_\tau)) [\varphi(X_\tau) - \mu_\beta(\varphi)] d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&= \frac{\int_0^t \frac{1}{n(\bar{F}_\tau)} \mathcal{L}_X^{A_\tau} \Psi(A_\tau, \cdot)(X_\tau) d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&= \lim_{k \rightarrow \infty} \frac{\int_0^t \frac{1}{n^{(k)}(\bar{F}_\tau)} \mathcal{L}_X^{A_\tau} \Psi^{(k)}(A_\tau, X_\tau) d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&= \lim_{k \rightarrow \infty} \frac{\int_0^t \mathcal{L}_X^{A_\tau} \Phi^{(k)}(\tau, X_\tau) d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} =: \lim_{k \rightarrow \infty} \epsilon_t^{(k)}(\varphi),
\end{aligned}$$

where the limit  $k \rightarrow \infty$  is understood in an almost sure sense, thanks to Assumption 3.2, and the fact that  $\bar{F}_t \in \mathcal{F}$  for all  $t \geq 0$ , almost surely, thanks to Theorem 3.3.

Itô's formula can be used, since  $(t, x) \mapsto \Phi^{(k)}(t, x)$  is of class  $\mathcal{C}^{1,2}$  thanks to Proposition 5.5. Then

$$\begin{aligned}
\epsilon_t^{(k)}(\varphi) &= \frac{\Phi^{(k)}(t, X_t) - \Phi^{(k)}(0, X_0)}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&\quad - \frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla \Phi^{(k)}(\tau, X_\tau), \Sigma dW(\tau) \rangle}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&\quad - \frac{\int_0^t \frac{\partial \Phi^{(k)}}{\partial \tau}(\tau, X_\tau) d\tau}{1 + \int_0^t F_\tau(\xi_S(X_\tau)) d\tau} \\
&=: \epsilon_t^{(k),1}(\varphi) + \epsilon_t^{(k),2}(\varphi) + \epsilon_t^{(k),3}(\varphi).
\end{aligned}$$

We now prove the following result.

**Lemma 5.6.** *Let  $\varphi \in \mathcal{C}$ . There exists  $C(\varphi) \in (0, \infty)$  such that for every  $t \geq 0$  and  $k \in \mathbb{N}$*

$$\mathbb{E} |\epsilon_t^{(k)}(\varphi)|^2 \leq \frac{C(\varphi)}{t}.$$

Observe that Lemma 5.6 is valid for test functions  $\varphi$  in the set  $\mathcal{C}$  defined by (27). To prove Theorem 4.1, *i.e.* an almost sure convergence result, we will only use it with test functions which are bounded and have bounded derivatives. However, to prove Proposition 4.5, we will need this Lemma for test functions with polynomial growth.

*Proof of Lemma 5.6.* The proof of that result consists in using the estimates of Proposition 5.5.

- Thanks to item (i) from Proposition 5.5, for every  $t \geq 0$ ,  $\Phi^{(k)}(t, \cdot)$  has at most polynomial growth, and moments of the process  $X$  are bounded, see Theorem 3.3. More precisely, the parameters  $C$  and  $p$  in the right-hand side of the inequality (29) do not depend on  $A = A_\tau$ . Moreover, thanks to Assumption 3.2, for every  $k \in \mathbb{N}$  and  $\tau \geq 0$ , one has  $n^{(k)}(\bar{F}_\tau) \geq m > 0$  almost surely.

It is then straightforward to conclude that

$$\mathbb{E} |\epsilon_t^{(k),1}(\varphi)|^2 \leq \frac{C(\varphi)}{t^2}.$$



- To have an estimate of the stochastic integral, we use Itô's formula, and we obtain

$$\begin{aligned}\mathbb{E}|\epsilon_t^{(k),2}(\varphi)|^2 &\leq C \frac{1 + \int_0^t |\Sigma^* \nabla_x \Phi^{(k)}(\tau, X_\tau)|^2 d\tau}{1 + t^2} \\ &\leq \frac{C(\varphi)}{t},\end{aligned}$$

thanks to (30), and arguments similar to the term above.

- Finally, using (31), and similar arguments, one obtains

$$\mathbb{E}|\epsilon_t^{(k),3}(\varphi)|^2 \leq \frac{C(\varphi) \left( \int_0^t \frac{1}{1+\tau} d\tau \right)^2}{t^2} \leq \frac{C(\varphi)(1 + \log(t))^2}{t^2}.$$

Gathering estimates then concludes the proof of Lemma 5.6.  $\square$

We are now in position to deduce Theorem 4.1 from Lemma 5.6. First, note that it is straightforward to obtain

$$\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 \leq \frac{C(\varphi)}{t}.$$

Indeed, the right-hand side in the estimate of Lemma 5.6 does not depend on  $k$ , and taking the limit  $k \rightarrow \infty$  in the right-hand side gives the result, thanks to Assumption 3.2 which ensures the required uniform convergence properties for the application of the bounded convergence theorem.

This estimate ensures the convergence in mean-square sense, and in probability, of  $\bar{\mu}_t(\varphi)$  to  $\mu_\beta(\varphi)$ . To go further, and obtain the almost sure convergence, we use the following arguments. First, note that it is sufficient to prove that  $\bar{\mu}_{\exp(t)}$  converges almost surely to  $\mu_\beta(\varphi)$  when  $t \rightarrow \infty$ . Using the estimate

$$\mathbb{E}|\bar{\mu}_{\exp(t)}(\varphi) - \mu_\beta(\varphi)|^2 \leq C(\varphi)e^{-t},$$

and Borel-Cantelli's Lemma, then almost surely, for every  $\delta \in \mathbb{Q} \cap (0, \infty)$ ,

$$\bar{\mu}_{\exp(n\delta)}(\varphi) \xrightarrow{n \rightarrow \infty} \mu_\beta(\varphi).$$

Finally, thanks to the differential equation (16) and boundedness of the function  $\varphi$ , the mapping  $t \mapsto \bar{\mu}_{\exp(t)}(\varphi)$  is Lipschitz continuous, with constant smaller than  $C(\varphi)$ , almost surely, for some  $C(\varphi) \in (0, \infty)$  depending only on  $\varphi$ , and on the parameters appearing in the definition of the set  $\mathcal{F}$ , see (12).

It is then straightforward to obtain the almost sure convergence

$$\bar{\mu}_{\exp(t)}(\varphi) \xrightarrow{t \rightarrow \infty} \mu_\beta(\varphi).$$

This concludes the proof of Theorem 4.1.

## 6. ANALYSIS OF THE MEAN-SQUARE ERROR. PROOF OF PROPOSITION 4.5

In this section, we give a proof of Proposition 4.5, concerning the asymptotic behavior of the mean-square error, which is decomposed as the sum of the bias squared and of the variance,

$$(34) \quad \mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 = (\mathbb{E}\bar{\mu}_t(\varphi) - \mu_\beta(\varphi))^2 + \text{Var}(\bar{\mu}_t(\varphi)),$$

when  $t \rightarrow \infty$ , for functions  $\varphi \in \mathcal{C}$ , of class  $\mathcal{C}^\infty$ , with at most polynomial growth.

In Section 6.1, we prove that the bias satisfies

$$(35) \quad \mathbb{E}\bar{\mu}_t(\varphi) - \mu_\beta(\varphi) = O\left(\frac{1 + \log(t)}{t}\right).$$

In Section 6.2, we then prove that

$$(36) \quad t\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 \xrightarrow{t \rightarrow \infty} \mathbf{V}_\infty(\varphi) \in [0, \infty).$$

In particular, thanks to (35), we may interpret the limit as the asymptotic variance, since

$$\mathbf{V}_\infty(\varphi) = \lim_{t \rightarrow \infty} t \text{Var}(\bar{\mu}_t(\varphi)).$$

The asymptotic variance is expressed in terms of the solution of a Poisson equation (26), with  $A = A_\infty = \lim_{t \rightarrow \infty} A_t$  (defined in Corollary 4.3).

In Section 6.3, we check that  $\mathbf{V}_\infty(\varphi) = \mathbf{V}_\infty(\varphi, A_\infty)$ , where  $\mathbf{V}_\infty(\varphi, A) \in [0, \infty)$  is the asymptotic variance associated with the non-adaptively biasing method, using (4) and (9), with  $A = A_\infty$ .

**6.1. Asymptotic behavior of the bias.** Let us prove (35). Using the same arguments as in Section 5.2.3, note that

$$\begin{aligned} \mathbb{E}\bar{\mu}_t(\varphi) - \mu_\beta(\varphi) &= \mathbb{E}\left[\lim_{k \rightarrow \infty} \epsilon_t^{(k)}(\varphi)\right] \\ &= \lim_{k \rightarrow \infty} \mathbb{E}\left[\epsilon_t^{(k)}(\varphi)\right] \\ &= \lim_{k \rightarrow \infty} \mathbb{E}\left[\epsilon_t^{(k),1}(\varphi) + \epsilon_t^{(k),3}(\varphi)\right]. \end{aligned}$$

Indeed, using Assumption 3.2 and the property that  $\bar{F}_t \in \mathcal{F}$ , for all  $t \geq 0$ , almost surely, allows us to use the bounded convergence theorem. Moreover, note that  $\mathbb{E}[\epsilon_t^{(k),2}(\varphi)] = 0$ , for all  $k \in \mathbb{N}$  and  $t \geq 0$ .

It then remains to use Lemma 5.6 to conclude the proof of (35). Note that we have considered test functions  $\varphi \in \mathcal{C}$ , which are not assumed to be bounded.

**6.2. Asymptotic behavior of the mean-square error.** Let us now prove (36). Like in Sections 5.2.3 and 6.1 above, we use the decomposition of  $\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)$  in terms of the auxiliary function  $\Phi^{(k)}$ ; we prove error bounds which are uniform with respect to  $k \in \mathbb{N}$ , and pass to the limit  $k \rightarrow \infty$ , thanks to Assumption 3.2 and Theorem 3.3.

It is straightforward to check that, uniformly in  $k \in \mathbb{N}$ ,

$$t\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 - t\mathbb{E}\left|\frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Phi^{(k)}(\tau, X_\tau), \Sigma dW(\tau) \rangle}{1 + \int_0^t F_\tau(\xi_S(X_\tau^A)) d\tau}\right|^2 = \mathcal{O}\left(\frac{1}{t}\right),$$

thanks to Lemma 5.6, *i.e.* only the stochastic integral contributes to the asymptotic variance. We can directly pass to the limit  $k \rightarrow \infty$  at this stage.

Let  $R(t) = \frac{1 + \int_0^t F_\tau(\xi_S(X_\tau^A)) d\tau}{t}$ . Then almost surely, thanks to Corollary 4.3 (uniform convergence of  $F_\tau$  to  $F_\infty$  when  $\tau \rightarrow \infty$ ), Cesaro's Lemma, and Corollary 4.4,

$$R(t) \xrightarrow{t \rightarrow \infty} \int_{\mathcal{S}} F_\infty \circ \xi_S d\mu_\beta^{A_\infty} = \frac{\mu_\beta^{A_\infty}(e^{-\beta A_\infty \circ \xi_S})}{\mathfrak{n}(\bar{F}_\infty)} = \frac{1}{\mu_\beta^0(e^{\beta A_\infty \circ \xi_S}) \mathfrak{n}(\bar{F}_\infty)},$$

using (8). We then obtain, using standard arguments,

$$\begin{aligned}
t\mathbb{E}\left|\frac{\int_0^t \frac{\sqrt{2\beta^{-1}}}{n(\bar{F}_\tau)} \langle \nabla_x \Psi(A_\tau, X_\tau), \Sigma dW(\tau) \rangle}{1 + \int_0^t F_\tau(\xi_S(X_\tau^A)) d\tau}\right|^2 &= \mathbb{E} \frac{\left| \int_0^t \frac{\sqrt{2\beta^{-1}}}{n(\bar{F}_\tau)} \langle \nabla_x \Psi(A_\tau, X_\tau), \Sigma dW(\tau) \rangle \right|^2}{t|R(t)|^2} \\
&= \mathbb{E} \frac{\left| \int_0^t \frac{\sqrt{2\beta^{-1}}}{n(\bar{F}_\tau)} \langle \nabla_x \Psi(A_\tau, X_\tau), \Sigma dW(\tau) \rangle \right|^2}{t} + o(1) \\
&= \frac{2\beta^{-1} \int_0^t \mathbb{E} \frac{|\Sigma^* \nabla_x \Psi(A_\tau, X_\tau)|^2}{n(\bar{F}_\tau)^2} d\tau}{t} + o(1) \\
&= \frac{\frac{2\beta^{-1}}{n(\bar{F}_\infty)^2} \int_0^t \mathbb{E} |\Sigma^* \nabla_x \Psi(A_\infty, X_\tau)|^2 d\tau}{t} + o(1) \\
&= 2\beta^{-1} \mathbb{E} \bar{\mu}_t(|\Sigma^* \nabla_x \Psi(A_\infty, \cdot)|^2) + o(1) \\
&\xrightarrow{t \rightarrow \infty} 2\beta^{-1} \mu_\beta(|\Sigma^* \nabla_x \Psi(A_\infty, \cdot)|^2),
\end{aligned}$$

thanks to Lemma 5.6, applied to the function  $|\Sigma^* \nabla_x \Psi(A_\infty, \cdot)|^2 \in \mathcal{C}$ , thanks to Proposition 5.3. Observe that the limit does not depend on the normalization operator  $n$ .

We thus obtain (36), more precisely,

$$t\mathbb{E}|\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 \xrightarrow{t \rightarrow \infty} \mathbf{V}_\infty(\varphi) = 2\beta^{-1} \mu_\beta(|\Sigma^* \nabla_x \Psi(A_\infty, \cdot)|^2).$$

**6.3. Comparison with the non-adaptive biasing method.** We now prove that the expression obtained above for the asymptotic variance in the adaptive method, coincides with the expression of the asymptotic variance in the non-adaptive method, when choosing  $A = A_\infty$ .

Let  $A \in \mathcal{A}$  (see (12)), and  $\varphi \in \mathcal{C}$ . Using (9), and the solution of the Poisson equation (26),

$$\begin{aligned}
\bar{\mu}_t^A(\varphi) - \mu_\beta(\varphi) &= \frac{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau^A)) [\varphi(X_\tau^A) - \mu_\beta(\varphi)] d\tau}{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau)) d\tau} \\
&= \frac{1 + \int_0^t \mathcal{L}_X^A \Psi(A, X_\tau^A) d\tau}{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau)) d\tau} \\
&= \frac{\Psi(A, X_t^A) - \Psi(A, X_0^A) - \int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Psi(A, X_\tau^A), \Sigma dW(\tau) \rangle}{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau)) d\tau}
\end{aligned}$$

Since  $A \in \mathcal{A}$ ,  $\exp(-\beta A \circ \xi_S) \geq m > 0$ , for some  $m > 0$ . Then

$$t\mathbb{E}|\bar{\mu}_t^A(\varphi) - \mu_\beta(\varphi)|^2 - t\mathbb{E}\left|\frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Psi(A, X_\tau^A), \Sigma dW(\tau) \rangle}{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau^A)) d\tau}\right|^2 = O\left(\frac{1}{t}\right).$$

Let  $R^A(t) = \frac{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau^A)) d\tau}{t}$ . Then almost surely  $R^A(t) \xrightarrow{t \rightarrow \infty} \mu_\beta^A(e^{-\beta A \circ \xi_S}) = \frac{1}{\mu_\beta(e^{\beta A \circ \xi_S})}$ .

With the same arguments as in Section 6.2 above,

$$\begin{aligned}
t\mathbb{E}\left|\frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Psi(A, X_\tau^A), \Sigma dW(\tau) \rangle}{1 + \int_0^t \exp(-\beta A \circ \xi_S(X_\tau^A)) d\tau}\right|^2 &= \mathbb{E}\left|\frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Psi(A, X_\tau^A), \Sigma dW(\tau) \rangle}{t|R^A(t)|^2}\right|^2 \\
&= \mathbb{E}\left|\frac{\int_0^t \sqrt{2\beta^{-1}} \langle \nabla_x \Psi(A, X_\tau^A), \Sigma dW(\tau) \rangle}{t}\right|^2 + o(1) \\
&= \frac{2\beta^{-1} \int_0^t \mathbb{E}|\Sigma^* \nabla_x \Psi(A, X_\tau^A)|^2 d\tau}{t} + o(1) \\
&= 2\beta^{-1} \mathbb{E} \bar{\mu}_t^A (|\Sigma^* \nabla_x \Psi(A, \cdot)|^2) + o(1) \\
&\xrightarrow{t \rightarrow \infty} 2\beta^{-1} \mu_\beta (|\Sigma^* \nabla_x \Psi(A, \cdot)|^2).
\end{aligned}$$

We thus conclude that

$$t\mathbb{E}|\bar{\mu}_t^A(\varphi) - \mu_\beta(\varphi)|^2 \xrightarrow{t \rightarrow \infty} \mathbf{V}_\infty(\varphi, A) = 2\beta^{-1} \mu_\beta (|\Sigma^* \nabla_x \Psi(A, \cdot)|^2).$$

The asymptotic variance in the ABP method is thus equal to the asymptotic variance in the non-adaptive method with  $A = A_\infty = \lim_{t \rightarrow \infty} A_t$ , as expected:

$$\mathbf{V}_\infty(\varphi) = 2\beta^{-1} \mu_\beta (|\Sigma^* \nabla_x \Psi(A_\infty, \cdot)|^2) = \mathbf{V}_\infty(\varphi, A_\infty).$$

## 7. THE SPDE CASE

The aim of this section is to generalize the approach developed in other sections of this article, to deal with metastable stochastic processes in infinite dimension. More precisely, we describe an ABP method designed to compute averages  $\mu_\beta(\varphi)$ , where  $\mu_\beta$  is a probability distribution defined on an infinite dimensional (Hilbert) space; the corresponding diffusion processes are given by some parabolic semilinear Stochastic Partial Differential Equations (SPDEs).

In Section 7.1, we describe the model, and we explain how it fits in the framework of Section 2. In particular, this description justifies the introduction of the abstract objects in Section 2.

Some arguments and some statements need to be substantially modified, compared with the finite dimensional situation. We make these modifications precise in Section 7.2 and 7.3. However, we do not provide the associated detailed proofs.

**7.1. The model.** In this section, we consider infinite dimensional diffusion processes, which are solutions of parabolic, semilinear, SPDEs, driven by space-time white noise, in space dimension 1, which may be written in the following form:

$$(37) \quad du^0(t, x) = \frac{\partial^2 u^0(t, x)}{\partial x^2} dt - \nabla \mathcal{V}(u^0(t, x)) dt + \sqrt{2\beta^{-1}} dW(t, x),$$

for  $x \in (0, 1)$ , with (for instance) homogeneous Dirichlet boundary conditions.

The function  $\mathcal{V} : \mathbb{R} \rightarrow \mathbb{R}$  is a smooth mapping. With the choice  $\mathcal{V}(x) = \frac{x^4}{4} - \frac{x^2}{2}$ , one obtains the Allen-Cahn equation, which is the paradigmatic example of metastable SPDE considered in the literature: see for instance [9], [10], [11], [22].

In this article,  $\mathcal{V}$  is assumed to have bounded derivatives, in order to simplify the presentation and the functional setting. Metastable states are solutions of the stationary PDE

$$\frac{\partial^2 u(x)}{\partial x^2} - \nabla \mathcal{V}(u(tx)) = 0.$$

Assume that the potential energy function  $\mathcal{V}$  is even; then  $x \mapsto u^0(x) = 0$  is one solution. Moreover, if there exists another solution  $x \mapsto u^+(x)$ ,  $x \mapsto u^-(x) = -u^+(x)$  is also a solution. These solutions

are critical points of the energy functional

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

and may be local minima, saddle points, etc...

It is convenient and standard to write (37) as a Stochastic Evolution Equation in the Hilbert space  $H = L^2(0, 1)$ , see for instance the monograph [17]:

$$(38) \quad du_t^0 = Lu_t^0 dt - DV(u_t^0)dt + \sqrt{2\beta^{-1}}dW(t),$$

where  $D$  denotes the Fréchet derivative, and

- $(e_n)_{n \in \mathbb{N}^*}$  is the complete orthonormal system of  $H$  given by  $e_n(x) = \sqrt{2} \sin(n\pi x)$ ;
- the unbounded linear operator  $L : H \rightarrow H$  satisfies  $Lu = -\sum_{n \in \mathbb{N}^*} \pi^2 n^2 \langle u, e_n \rangle e_n$ ;
- $V(u) = \int_0^1 \langle \nabla V(\theta u), u \rangle d\theta$  for all  $u \in H$ ;
- $(W(t))_{t \geq 0}$  is a cylindrical Wiener process on  $H$ , i.e.  $W(t) = \sum_{n \in \mathbb{N}^*} \beta_n(t) e_n$  for a family  $(\beta_n)_{n \in \mathbb{N}^*}$  of independent, one-dimensional, standard Wiener processes.

Equation (38) admits a unique mild solution (see [17]) with values in  $H$ , defined for  $t \geq 0$ , i.e.  $u^0$  is the unique process solution satisfying the equation

$$u_t^0 = e^{tL}u_0 - \int_0^t e^{(t-s)L} DV(u_s^0)ds + \sqrt{2\beta^{-1}} \int_0^t e^{(t-s)L} dW(s),$$

where  $(e^{tL})_{t \in [0, +\infty)}$  is the semi-group on  $H$  generated by  $L$ :  $e^{tL}u = \sum_{n \in \mathbb{N}^*} e^{-\pi^2 n^2 t} \langle u, e_n \rangle e_n$ .

In the context of this section, Assumption 2.2 is rephrased as follows:

$$(39) \quad \sup_{x \in \mathbb{R}} |V''(x)| < \pi^2.$$

In other words, the Lipschitz constant of the non-linear coefficient  $u \in H \mapsto DV(u) \in H$  is bounded from above by all the eigenvalues of  $-L$ . Inequality (39) is a standard sufficient (but not necessary) condition for ergodicity of the SPDE (38), and for exponential convergence to equilibrium.

We are now in position to explain how the SPDE dynamics fits into the general framework presented in this article, in Section 2.

7.1.1. *Setting.* In the SPDE example, one has the following elements, see Section 2.1.

- **State space:**  $\mathcal{S} = H$  (infinite dimensional, separable, Hilbert space).
- **Reaction coordinate:** assume that  $\mathbb{M}_m = \mathbb{T}$  (with  $m = 1$ ),  $E_1 = \mathbb{R}$  (with  $d = 1$ ). Then for instance  $\xi(u) = \xi_{\mathcal{S}}(u) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{1}{2} \int_0^1 u(x) dx\right)$ .
- **Drift coefficient:**  $\mathcal{D}(V, A) = Lu - D(V - A \circ \xi_{\mathcal{S}})$ . **Diffusion operator:**  $\Sigma$  is the identity on  $\mathcal{S}$ .

Since  $L$  is an unbounded linear operator on  $H$ , note that the drift is only defined on a domain  $D(L) \subset H$ . This is one of the technical issues which are specific to the infinite dimensional framework.

**Remark 7.1.** Note that, in general, there does not exist a function  $\mathcal{V}_A : \mathbb{R} \rightarrow \mathbb{R}$  such that the function  $V - A \circ \xi : L^2(0, 1) \rightarrow \mathbb{R}$  satisfies  $D(V - A \circ \xi)(u)(x) = \nabla \mathcal{V}_A(u(x))$ : the bias is a nonlocal function of  $u$ , since it depends on the spatial average  $\int_0^1 u(y) dy$ , instead of  $u(x)$  only.

The biased version (4) of the SPDE (37) is written as

$$(40) \quad du^A(t) = Lu^A(t)dt - D(V - A \circ \xi_S)(u^A(t))dt + \sqrt{2\beta^{-1}}dW(t),$$

with mild formulation

$$u^A(t) = e^{tL}u_0 - \int_0^t e^{(t-s)L}DV(u^A(s))ds + \int_0^t e^{(t-s)L}D(A \circ \xi_S)(u^A(s))ds + \sqrt{2\beta^{-1}} \int_0^t e^{(t-s)L}dW(s).$$

**7.1.2. Invariant probability distribution.** We now construct the Total Energy function, and the reference measure  $\lambda$  on  $H$ .

First, the definition of the mapping  $V \mapsto \mathcal{E}(V)$  is straightforward:  $\mathcal{E}(V) = V$ . The reference measure  $\lambda$  on  $\mathcal{S}$  is defined as follows: it is the centered Gaussian probability distribution on  $H$  with covariance operator  $\beta^{-1}L^{-1}$ . This measure can be constructed as follows: let  $(g_n)_{n \in \mathbb{N}^*}$  be a sequence of independent standard real-valued Gaussian random variables (centered and with variance 1); then  $\lambda$  is the probability distribution of the  $H$ -valued random variable  $\sum_{n \in \mathbb{N}^*} \frac{\sqrt{\beta^{-1}}}{n\pi} g_n e_n$ .

**Remark 7.2.** *One may check that  $\lambda$  defined as above is the distribution of the Brownian Bridge on  $(0, 1)$ . This interpretation is specific to the choice of  $L$  and plays no role in this article. On the contrary, the construction above, based on eigenvalues and eigenvectors of  $L$ , is general.*

It is straightforward to check that  $\lambda$  is the unique invariant distribution of (38) when  $V = 0$ . More generally, for any function  $A : \mathbb{T} \rightarrow \mathbb{R}$  of class  $C^\infty$ , the probability distribution  $\mu_\beta^A$  on  $H$ , defined by

$$\mu_\beta^A(du) = \frac{\exp(-\beta(V(u) - A(\xi(u))))}{Z^A(\beta)} \lambda(du)$$

where  $Z^A(\beta) \in (0, \infty)$  thanks to (39), is the unique invariant distribution of the biased SPDE (40), see for instance [16].

**7.1.3. Free Energy function.** It remains to discuss how the Free Energy function  $A_\star$  is defined. The only thing to be careful about is the choice of reference measure  $\pi$  defined on  $\mathbb{M}_m = \mathbb{T}$ , such that Assumption 2.6 is satisfied. Indeed, the free energy function  $A_\star$  is then obtained, as in finite dimensional cases, by Definition 2.8.

It is legitimate here to choose  $\pi$  to be the Lebesgue measure on  $\mathbb{T}$ . Indeed, the image measure of the Gaussian distribution  $\lambda$  by the linear mapping  $u \mapsto \int_0^1 u(y)dy$  is a non-degenerate Gaussian distribution on  $\mathbb{R}$ ; thus the image of  $\lambda$  by  $\xi_S$  is equivalent to the Lebesgue measure on  $\mathbb{T}$ . Then  $\pi_\beta^0$  the image of  $\mu_\beta^0$  by  $\xi$  is equivalent to  $\pi$ .

**7.2. ABP dynamics and convergence results.** Let us first describe the dynamics of the ABP method, which generalizes (10) in the case where the diffusion process is governed by a SPDE:

$$(41) \quad \begin{cases} du(t) = Lu(t)dt - D(V - A_t \circ \xi_S)(u(t))dt + \sqrt{2\beta^{-1}}dW(t) \\ \bar{\mu}_t = \frac{\bar{\mu}_0 + \int_0^t \exp(-\beta A_r \circ \xi(X_r)) \delta_{X_r} dr}{1 + \int_0^t \exp(-\beta A_r \circ \xi(X_r)) dr} \\ \exp(-\beta A_t(z)) = \int_{\mathbb{T}^d} K(z, \xi(x)) \bar{\mu}_t(dx), \quad \forall z \in \mathbb{T}^m, \end{cases}$$

For simplicity, we have chosen the normalization operator  $\mathcal{N}$ , with  $n(F) = \int_{\mathbb{T}} F(z)dz$ . The kernel function  $K : \mathbb{T} \times \mathbb{T} \rightarrow (0, \infty)$  satisfies Assumption 3.1.

As explained in Section 7.1 above, it is convenient to consider the mild formulation for the SPDE dynamics: the first equation in (41) is understood as

$$u(t) = e^{tL}u_0 - \int_0^t e^{(t-s)L}D(V - A_s \circ \xi)(u(s))ds + \sqrt{2\beta^{-1}} \int_0^t e^{(t-s)L}dW(s).$$

Using Lemma 3.5 and standard techniques, the following generalization of Theorem 3.3 is obtained.

**Theorem 7.3.** *Consider the framework of Section 7.1 (in particular assume that (39) is satisfied), and assume that the kernel function  $K$  satisfies Assumption 3.1.*

- *There exists a unique process  $t \mapsto (u(t), \bar{\mu}_t, A_t)$ , taking values in  $\mathcal{C}(\mathbb{R}^+, H \times \mathcal{P}(H) \times \mathcal{C}^0(\mathbb{M}_m, (0, \infty)))$ , which is solution of the ABP system (41).*
- *For all  $n \geq 2$ ,  $\sup_{t \geq 0} \mathbb{E} \|u(t)\|_H^n < +\infty$ .*
- *There exist  $m \in (0, \infty)$  and  $(M^{(r)})_{r \in \{0, 1, \dots\}} \in (0, \infty)$  such that, almost surely,  $A_t \in \mathcal{A}$ , for all  $t \in \mathbb{R}^+$ , where*

$$(42) \quad \begin{cases} \mathcal{F} = \{F \in \mathcal{C}^\infty(\mathbb{M}_m); \min F \geq m, \max |\partial^k F| \leq M^{(k)}, k \geq 0\}, \\ \mathcal{A} = \{A = -\frac{1}{\beta} \log(\bar{F}); F \in \mathcal{F}\}. \end{cases}$$

We are able to prove generalizations of Theorem 4.1 and of Corollary 4.3

**Theorem 7.4.** • *Let  $\varphi \in \mathcal{C}^\infty(H, \mathbb{R})$  be a bounded function, with bounded derivatives of any order. Then, almost surely,*

$$\bar{\mu}_t(\varphi) \xrightarrow[t \rightarrow \infty]{} \mu_\beta(\varphi).$$

- *Define, for all  $z \in \mathbb{M}_m$ ,*

$$A_\infty(z) = -\frac{1}{\beta} \log(\mu_\beta(K(z, \cdot))).$$

*Then, almost surely, for every  $\ell \in \{0, 1, \dots\}$ , uniformly on  $\mathbb{M}_m$ ,*

$$\partial^\ell A_t \xrightarrow[t \rightarrow \infty]{} \partial^\ell A_\infty.$$

The efficiency results from Section 4.3 may also be generalized: more precisely, the convergence result (14), and Proposition 4.5 remain valid.

**7.3. Some modifications for SPDEs.** The first type of modifications, compared with the finite dimensional situation, is concerned with the statement of the convergence results. Indeed, note that we have only described the (almost sure) convergence of averages  $\bar{\mu}_t(\varphi)$ , not the convergence of the probability distribution  $\bar{\mu}_t$ . The arguments used in the proof of Corollary 4.2 do not easily generalize to the infinite dimensional setting.

The second type of modifications, is concerned with the proof of Theorem 7.4, more precisely of the convergence of  $\bar{\mu}_t(\varphi)$ . The strategy of the proof follows the steps developed in Section 5.2; however care is required when dealing with the Poisson equations (26).

To simplify the discussion, assume first that  $A = 0$ . Then the Poisson equation can be written as

$$\langle Lu - DV(u), D\Psi(u) \rangle + \frac{1}{2} \text{Trace}(D^2\Psi(u)) = \varphi(u) - \mu_\beta(\varphi), \quad \forall u \in H,$$

where the unknown is the function  $\Psi : H \rightarrow \mathbb{R}$ , and we use standard identifications, in order to consider the first order derivative  $D\Psi(x) \in H$  as an element of  $H$ , and the second order derivative  $D^2\Psi(x) \in \mathcal{L}(H)$  as a bounded linear operator on  $H$ . For a function  $\Psi$  of class  $\mathcal{C}^2$  on  $H$ , it is not true in general that the left-hand side is well-defined, for all  $u \in H$ , or even when  $u = u(t)$  is the diffusion process evaluated at a time  $t \geq 0$ . Indeed,  $L$  is an unbounded operator, so  $Lu$  is not an element of  $H$  in general. Moreover, the trace term may be infinite.

In fact, solutions of Poisson equations in infinite dimensions have appropriate regularity properties, such that all the terms are well-defined. See for instance [13, Chapters 4,5], for smoothing properties of the transition semi-group. For details concerning the Poisson equation, see, for instance, [12, Proposition 6.1], and [28, Chapter 4, Section 8]. Rigorous properties are often stated for

spatial Galerkin approximations, with bounds not depending on the dimension. We do not provide such details here, and directly write the results in the Hilbert space  $H$ .

Using arguments from the references mentioned above, and taking care of the dependence with respect to the function  $A$  to obtain uniform bounds on the set  $\mathcal{A}$ , generalizations of Propositions 5.3 and 5.5 are obtained.

For any  $\alpha \in (0, 1)$ , introduce the unbounded linear operator  $L^\alpha$ , such that

$$L^\alpha u = \sum_n (-n^2 \pi^2)^\alpha \langle u, e_n \rangle e_n.$$

**Proposition 7.5.** *Let  $A \in \mathcal{A}$ , and  $\varphi : H \rightarrow \mathbb{R}$ , of class  $C^\infty$ , bounded and with bounded derivatives of any order.*

*There exists a unique solution  $\Psi(A, \cdot)$  of the Poisson equation (26),*

$$\langle Lu - D(V - A \circ \xi)(u), D\Psi(u) \rangle + \frac{1}{2} \text{Trace}(D^2\Psi(u)) = e^{-\beta A(\xi(u))} [\varphi(u) - \mu_\beta(\varphi)], \quad \forall u \in H,$$

*with the condition  $\int_H \Psi d\mu_\beta = 0$ .*

*This solution is given by*

$$\Psi(u) = - \int_0^\infty \mathbb{E}_u [\bar{\varphi}^A(u(t))] dt,$$

*for all  $u \in H$ , where  $\bar{\varphi}^A(u) = e^{-\beta A(\xi(u))} [\varphi(u) - \mu_\beta(\varphi)]$ .*

*Moreover, the following properties are satisfied.*

- *There exists  $C(\varphi) \in (0, \infty)$  such that, for all  $A \in \mathcal{A}$  and  $u \in H$ ,*

$$|\Psi(A, u)| \leq C(\varphi)(1 + \|u\|_H^2).$$

- *For every  $\alpha \in (0, \frac{1}{2})$ , there exists  $C(\alpha, \varphi) \in (0, \infty)$ , such that, for all  $A \in \mathcal{A}$  and  $u \in H$*

$$(43) \quad \begin{cases} \|L^{2\alpha} D_u \Psi(A, u)\|_H \leq C(\alpha, \varphi)(1 + \|u\|_H^2) \\ \|L^\alpha D_u^2 \Psi(A, u) L^\alpha\|_{\mathcal{L}(H)} \leq C(\alpha, \varphi)(1 + \|u\|_H^2). \end{cases}$$

- *For every  $\alpha \in (0, \frac{1}{4})$  and every  $n \in \mathbb{N}$ , there exists  $C(\alpha, n, \varphi) \in (0, \infty)$ , such that  $\mathbb{E} \|L^\alpha u(t)\|^n \leq C(\alpha, n, \varphi)(1 + \frac{\|u(0)\|_H}{t^\alpha})^n$ .*
- *The function  $(t, u) \in [0, \infty) \times H \mapsto \Psi(A_t, u)$  is of class  $\mathcal{C}^{1,2}$ , and for every  $u \in H$  and every  $t \geq 0$ , almost surely*

$$\left| \frac{\partial \Psi(A_t, u)}{\partial t} \right| \leq \frac{C(1 + \|u\|^2)}{1 + t},$$

*where  $(A_t)_{t \geq 0}$  is the  $\mathcal{A}$ -valued process defined in (41).*

We do not provide the details for the proof of Proposition 7.5. We emphasize on the estimates (43). They are the only results which are specific to the infinite dimensional case. These estimates are fundamental at two stages: first, to justify the well-posedness of the Poisson equation, second to control stochastic integral which appears in the analysis of the error  $\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)$ .

Adapting the strategy of proof of Theorem 4.1, developed in Section 5.2, and using Proposition 7.5 to control the terms, it is then straightforward to prove that

$$\mathbb{E} |\bar{\mu}_t(\varphi) - \mu_\beta(\varphi)|^2 \leq \frac{C(\varphi)}{t} \xrightarrow{t \rightarrow \infty} 0.$$

The proof of the almost sure convergence result is obtained using the boundedness of  $\varphi$ , and the same argument as in the finite dimensional case. This concludes the proof of the first part of Theorem 7.4.

The second part of Theorem 7.4, concerning the almost sure convergence of  $A_t$ , is proved exactly as Corollary 4.3.



Thanks to the general framework developed in Section 2, the application and the analysis of the ABP method also applies in the infinite dimensional setting, for metastable Stochastic PDEs.

#### ACKNOWLEDGMENTS

The authors would like to thank Tony Lelièvre and Gabriel Stoltz for stimulating discussions on this article. The work is partially supported by SNF grants 200020149871 and 200021163072.

#### REFERENCES

- [1] A. Barducci, G. Bussi, and M. Parrinello. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical review letters*, 100(2):020603, 2008.
- [2] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999.
- [3] M. Benaïm and C.-E. Bréhier. Convergence of adaptive biasing potential methods for diffusions. *C. R. Math. Acad. Sci. Paris*, 354(8):842–846, 2016.
- [4] M. Benaïm, M. Ledoux, and O. Raimond. Self-interacting diffusions. *Probab. Theory Related Fields*, 122(1):1–41, 2002.
- [5] M. Benaïm and O. Raimond. Self-interacting diffusions. II. Convergence in law. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(6):1043–1055, 2003.
- [6] M. Benaïm and O. Raimond. Self-interacting diffusions. III. Symmetric interactions. *Ann. Probab.*, 33(5):1717–1759, 2005.
- [7] M. Benaïm and O. Raimond. Self-interacting diffusions IV: Rate of convergence. *Electron. J. Probab.*, 16:no. 66, 1815–1843, 2011.
- [8] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. Translated from the French by Stephen S. Wilson.
- [9] N. Berglund and B. Gentz. Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond. *Electron. J. Probab.*, 18:no. 24, 58, 2013.
- [10] A. Bovier and F. den Hollander. *Metastability*, volume 351 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2015. A potential-theoretic approach.
- [11] C.-E. Bréhier, M. Gazeau, L. Goudenège, and M. Rousset. Analysis and simulation of rare events for SPDEs. In *CEMRACS 2013—modelling and simulation of complex systems: stochastic and deterministic approaches*, volume 48 of *ESAIM Proc. Surveys*, pages 364–384. EDP Sci., Les Ulis, 2015.
- [12] C.-E. Bréhier and M. Kopec. Approximation of the invariant law of spdes: error analysis using a poisson equation for a full-discretization scheme. *IMA Journal of Numerical Analysis*, 2016.
- [13] S. Cerrai. *Second order PDE’s in finite and infinite dimension*, volume 1762 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001. A probabilistic approach.
- [14] C. Chipot and T. Lelièvre. Enhanced sampling of multidimensional free-energy landscapes using adaptive biasing forces. *SIAM J. Appl. Math.*, 71(5):1673–1695, 2011.
- [15] 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*, 119(3):1129–1151, 2014.
- [16] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.
- [17] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
- [18] E. Darve and A. Pohorille. Calculating free energies using average force. *The Journal of Chemical Physics*, 115(20):9169–9183, 2001.
- [19] 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*, 114:5823–5830, 2010.
- [20] B. M. Dickson. Survey of adaptive biasing potentials: comparisons and outlook. *Current Opinion in Structural Biology*, 43:63–67, 2017.
- [21] M. Dufo. *Random iterative models*, volume 34 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. Translated from the 1990 French original by Stephen S. Wilson and revised by the author.
- [22] W. G. Faris and G. Jona-Lasinio. Large fluctuations for a nonlinear heat equation with noise. *Journal of Physics A: Mathematical and General*, 15(10):3025, 1982.

- [23] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz. Efficiency of the Wang-Landau algorithm: a simple test case. *Appl. Math. Res. Express. AMRX*, (2):275–311, 2014.
- [24] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre, and G. Stoltz. Convergence of the Wang-Landau algorithm. *Math. Comp.*, 84(295):2297–2327, 2015.
- [25] G. Fort, B. Jourdain, T. Lelièvre, and G. Stoltz. Self-healing umbrella sampling: convergence and efficiency. *Stat. Comput.*, 27(1):147–168, 2017.
- [26] J. Hénin and C. Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *The Journal of chemical physics*, 121(7):2904–2914, 2004.
- [27] B. Jourdain, T. Lelièvre, and R. Roux. Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process. *M2AN Math. Model. Numer. Anal.*, 44(5):831–865, 2010.
- [28] M. Kopec. Quelques contributions à l’analyse numérique d’équations stochastiques. PhD Thesis, *ENS Rennes.*, 2014.
- [29] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.
- [30] A. Laio and M. Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002.
- [31] T. Lelièvre and K. Minoukadeh. Long-time convergence of an adaptive biasing force method: the bi-channel case. *Arch. Ration. Mech. Anal.*, 202(1):1–34, 2011.
- [32] T. Lelièvre, M. Rousset, and G. Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21(6):1155–1181, 2008.
- [33] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations: A mathematical perspective*. Imperial College Press, London, 2010.
- [34] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numer.*, 25:681–880, 2016.
- [35] A. Lesage, T. Lelièvre, G. Stoltz, and J. Hénin. Smoothed biasing forces yield unbiased free energies with the extended-system adaptive biasing force method. *The Journal of Physical Chemistry B*, 2016.
- [36] 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*, 110(29):14011–14013, 2006.
- [37] 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*, 64(5):056101, 2001.
- [38] F. Wang and D. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Physical review letters*, 86(10):2050, 2001.

UNIVERSITÉ DE NEUCHÂTEL, INSTITUT DE MATHÉMATIQUES, RUE EMILE ARGAND 11, CH-2000 NEUCHÂTEL, SWITZERLAND

*E-mail address:* michel.benaim@unine.ch

UNIV LYON, UNIVERSITÉ CLAUDE BERNARD LYON 1, CNRS UMR 5208, INSTITUT CAMILLE JORDAN, 43 BLVD. DU 11 NOVEMBRE 1918, F-69622 VILLEURBANNE CEDEX, FRANCE

*E-mail address:* brehier@math.univ-lyon1.fr