

# Long-time convergence of an Adaptive Biasing Force method: the bi-channel case

T. Lelièvre and K. Minoukadeh

CERMICS, École des Ponts ParisTech, 6-8 avenue Blaise-Pascal,  
77455 Champs-sur-Marne, Marne-la-Valle cedex 2, France  
MICMAC Project-Team, INRIA Rocquencourt, 78153 Le Chesnay, France

## Abstract

We present convergence results for an adaptive algorithm to compute free energies, namely the adaptive biasing force (ABF) method [5, 8]. The free energy is the effective potential associated to a so-called reaction coordinate  $\xi(q)$ , where  $q = (q_1, \dots, q_{3N})$  is the position vector of a  $N$ -particle system. Computing free energy differences remains an important challenge in molecular dynamics due to the presence of meta-stable regions in the potential energy surface. The ABF method uses an on-the-fly estimate of the free energy to bias dynamics and overcome metastability. Using entropy arguments and logarithmic Sobolev inequalities, previous results have shown that the rate of convergence of the ABF method is limited by the metastable features of the canonical measures conditioned to being at fixed values of  $\xi$  [14]. In this paper, we present an improvement on the existing results, in the presence of such metastabilities, which is a generic case encountered in practice. More precisely, we study the so-called bi-channel case, where two channels along the reaction coordinate direction exist between an initial and final state, the channels being separated from each other by a region of very low probability. With hypotheses made on ‘channel-dependent’ conditional measures, we show on a bi-channel model that we introduce, that the convergence of the ABF method is in fact not limited by metastabilities in directions orthogonal to  $\xi$  under two crucial assumptions: (i) exchange between the two channels is possible for *some* values of  $\xi$  and (ii) the free energy is a good bias in each channel. This theoretical result supports recent numerical experiments [15], where the efficiency of the ABF approach is demonstrated for such a multiple-channel situation.

## 1 Introduction

We consider a system of  $N$  particles with positions  $q \in \mathcal{D} \subset \mathbb{R}^{3N}$ . In statistical physics, one is interested in calculating averages with respect to the Boltzmann-Gibbs measure

$$d\nu(q) = Z^{-1} \exp(-\beta V(q)) dq, \quad (1)$$

with  $V : \mathcal{D} \rightarrow \mathbb{R}$  the potential energy function,  $Z = \int_{\mathcal{D}} \exp(-\beta V(q)) dq$  the partition function,  $\mathcal{D} = \{q, V(q) < \infty\}$  the configuration space, and  $\beta = 1/(k_B T)$ , where  $k_B$  is the Boltzmann constant and  $T$  is the temperature. The function  $V$  is the energy associated with the positions of the particles and is assumed to be such that  $Z < \infty$ . The probability measure  $\nu$  represents the equilibrium measure sampled by the particles in the canonical ensemble. A typical dynamics that can be used to sample this measure through trajectorial averages is the overdamped Langevin dynamics

$$dQ_t = -\nabla V(Q_t) dt + \sqrt{2\beta^{-1}} dB_t, \quad (2)$$

where  $(B_t)_{t \geq 0}$  is a  $3N$ -dimensional standard Brownian motion. Indeed, under loose assumptions on  $V$ , one has the ergodic property: for any smooth test function  $\varphi$ ,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \varphi(Q_t) dt = \int_{\mathcal{D}} \varphi d\nu.$$

The efficiency of this sampling procedure, which can be shown to be related to the convergence rate to equilibrium of the above dynamics is often hindered by *metastabilities in the potential function*  $V$ , namely regions of low energy are separated by high energy barriers. Equivalently, in terms of the probability measure,  $\nu$  is typically a *multimodal measure*, with regions of high probability separated by regions of low probability. To circumvent this issue, a one-dimensional collective variable (or reaction coordinate)  $\xi : \mathcal{D} \rightarrow \mathcal{M}$  is introduced, which will be used to define a biasing potential for (2). In the following, we will assume that  $|\nabla \xi| > 0$  on  $\mathcal{D}$ , and that  $\mathcal{M} = \mathbb{T}$ , where  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  is the one-dimensional torus (which typically corresponds to the case where the reaction coordinate represents a dihedral angle, for example to characterize the conformation of a molecule). Before defining more precisely the biased dynamics in the next section, we need to introduce a few notation.

The image of the measure  $\nu$  in  $\xi$  is given by

$$d\nu^\xi(z) = Z^{-1} \exp(-\beta A(z)) dz,$$

where  $A$  is the so-called *free energy*, defined by

$$A(z) = -\beta^{-1} \ln(Z_z) \quad (3)$$

where

$$Z_z = \int_{\Sigma_z} \exp(-\beta V(q)) \delta_{\xi(q)-z}(dq) \quad (4)$$

is the partition function on the submanifold  $\Sigma_z = \{q \in \mathcal{D} \mid \xi(q) = z\}$ . The measure  $\delta_{\xi(q)-z}(dq)$  is defined through the conditioning formula: for any smooth test function  $\varphi : \mathcal{D} \rightarrow \mathbb{R}$ ,

$$\int_{\mathcal{D}} \varphi(q) dq = \int_{\mathcal{M}} \int_{\Sigma_z} \varphi(q) \delta_{\xi(q)-z}(dq) dz.$$

Using the co-area formula (see [14, Appendix A]), one can also identify this measure as  $\delta_{\xi(q)-z}(dq) = |\nabla \xi|^{-1} d\sigma_{\Sigma_z}$ , where  $\sigma_{\Sigma_z}$  is the Lebesgue measure on  $\Sigma_z$ . We assume in the following that  $\xi$  and  $V$  are such that  $Z_z < \infty$ , for all  $z \in \mathcal{M}$ .

Practitioners are typically interested in free energy differences  $A(z) - A(z_0)$ , which can be obtained by computing (and integrating) the derivative  $A'(z)$ , the so-called *mean force*

$$A'(z) = Z_z^{-1} \int_{\Sigma_z} f(q) \exp(-\beta V(q)) \delta_{\xi(q)-z}(dq) \quad (5)$$

where  $f$  is the *local mean force* defined by

$$f = \frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \operatorname{div} \left( \frac{\nabla \xi}{|\nabla \xi|^2} \right). \quad (6)$$

The function  $f$  can be understood as the negative force projected onto the reaction coordinate, plus some correction term related to the curvature of the submanifolds  $\Sigma_z$ . Notice that the mean force (5) is in fact a conditional expectation

$$A'(z) = \mathbb{E}_\nu[f(Q) \mid \xi(Q) = z] = \int_{\Sigma_z} f d\nu|_z, \quad (7)$$

where

$$d\nu|_z = \frac{\exp(-\beta V(q))\delta_{\xi(q)-z}(dq)}{\exp(-\beta(A(z)))}$$

denotes the probability measure  $\nu$  conditioned to a fixed value  $z$  of  $\xi(q)$ . This measure is supported on the submanifold  $\Sigma_z$ . For the derivation of (5)–(6)–(7) which is again based on the co-area formula, the reader is referred to [4, 17, 6].

## 1.1 The adaptive biasing force method

The adaptive biasing force (ABF) method [5, 8] uses an estimate of the mean force  $A'$  to bias the standard overdamped Langevin dynamics (2) in order to overcome metastabilities in  $\xi$ . The bottom line of the approach is that it should be easier to sample the probability measure with density proportional to  $\exp(-\beta(V - A \circ \xi))$  than the original Boltzmann-Gibbs measure  $\nu$ , since the marginal probability of the former in  $\xi$  is a uniform probability measure on  $\mathbb{T}$ , while the marginal of the latter (namely  $\exp(-\beta A(z)) dz$ ) is typically multimodal.

The ABF dynamics is given by

$$\begin{cases} dX_t = -\nabla(V - A_t \circ \xi)(X_t)dt + \sqrt{2\beta^{-1}}dB_t, \\ A'_t(z) = \mathbb{E}[f(X_t)|\xi(X_t) = z], \end{cases} \quad (8)$$

where  $A'_t$  is an on-the-fly estimate of the mean force, which, in view of the definition (7), is expected to be a good estimate of  $A'$ . The law of  $X_t$  has density  $\psi(t, \cdot)$ , which satisfies the non-linear Fokker-Planck equation:

$$\begin{cases} \partial_t \psi = \operatorname{div}(\nabla(V - A_t \circ \xi)\psi) + \beta^{-1}\Delta\psi, \\ A'_t(z) = \frac{\int_{\Sigma_z} f \psi \delta_{\xi(q)-z}(dq)}{\int_{\Sigma_z} \psi \delta_{\xi(q)-z}(dq)}. \end{cases} \quad (9)$$

Roughly speaking, the biasing force  $\nabla(A_t \circ \xi)$  “flattens the free-energy barriers in  $\xi$ ”. To support this claim, let us simply indicate that if  $|\nabla\xi|$  is constant, the marginal density in  $\xi$  satisfies a simple heat equation, with zero bias, see [13, 14] and also Equation (16) below. Existence and uniqueness of solutions to (8) are studied in [11] and a study of the longtime convergence of (9) can be found in [14], the results of which are briefly discussed below.

## 1.2 Existing convergence results, and the multiple channel scenario

It has been shown in [14] that, under appropriate assumptions, the biasing force  $A'_t$  in (8) (actually for a slightly different dynamics which reduces to (8) if  $|\nabla\xi|$  is constant for example) converges to the mean force  $A'$  exponentially fast in the longtime limit. The rate of convergence was estimated as the minimum of (i) the rate at which the law of  $\xi(X_t)$  converges to equilibrium, and (ii) the smallest logarithmic Sobolev inequality constant (LSI constant, discussed in Section 2.3) of the conditional probability measures  $\nu|_z$ , for  $z \in \mathcal{M}$ . Thanks to the bias in the direction of the reaction coordinate, it can be shown that the rate of convergence of the marginal in  $\xi$  is actually not the limiting rate in practice since it satisfies a simple diffusion equation. The real limitation is thus the metastable features (*i.e.* the multimodality) of the family of laws  $\nu|_z$ , which is quantified through the LSI constants associated to these measures: roughly speaking, the smaller the constant the more multimodal the probability measure. These constants may be in some cases extremely small, at least for some values of  $z \in \mathcal{M}$ . The question we address in this paper is the *optimality* of this theoretical rate of convergence for the ABF method.

The generic situation is indeed that the LSI constants for the measures  $\nu|_z$  are not small uniformly in  $z$ . Typically, there exists some values of  $z$  for which these measures are multimodal.

This situation is often encountered in practice due to two reasons. First, finding a suitable reaction coordinate, namely in our context one that ensures that there are no metastabilities associated to the equilibrium measures  $\nu|_z$ , is not trivial for a large-dimensional system. Secondly, a low-dimensional reaction coordinate may simply not be sufficient to describe all metastabilities of the system. In such cases, the results of [14] predict a very small rate of convergence for the ABF dynamics (8), and thus, the inefficiency of this biasing procedure.

As a typical case for which such difficulties appear, we will consider in the following the so-called ‘multiple channel situation’ (see Figure 1 on the left, for a bi-channel case). In such a situation, starting from a metastable state and as the system evolves in the direction of increasing values of the reaction coordinate  $\xi$ , it can follow different channels, which are separated (in the ‘orthogonal direction to  $\xi$ ’) by arbitrary high energy barriers. In other words, the energy landscape features parallel valleys which are orthogonal to the isocontours of the reaction coordinate. In such a prototypical situation, the conditional probability measures  $\nu|_z$  are indeed multimodal, for the values of  $z$  corresponding to the system being in one of these channels.

However, recent numerical experiments [15] (based on a discretization of the ABF dynamics (8) by multiple walkers simulated in parallel) suggest that in fact, high energy barriers in  $\Sigma_z$  do not always hinder the convergence of the ABF method. The multiple walkers are made to follow similar dynamics (8), but driven by independent Brownian motions. The chemical system considered in [15] is the compact states of the deca-alanine peptide (the reaction coordinate  $\xi$  being thus the end-to-end distance of the peptide). Due to some ‘buckling effects’, this is a typical multiple channel situation, since the molecule can shrink to various compact states (see [9]). In [15], numerical results show that the ABF approach indeed yields reliable results in such a situation. We interpret this as follows. When encountered with a fork in the channel, each walker is likely to travel down a different channel. Thus, it is indeed almost impossible for a given walker to switch from one channel to another, once it has entered one of them, but this appears not to be necessary to obtain reliable results. It suggests that the theoretical rate of convergence obtained in [14] is actually not optimal.

Inspired by these numerical results, we present herein an improved theoretical rate of convergence of the ABF method. The rate will be shown to depend on the LSI constants of the family of equilibrium measures conditioned to being in  $\Sigma_z$  and a channel. By doing so, we show that high energy barriers separating the channels do not in fact affect the rate of convergence of the method. The crucial assumptions needed to show our result are: (i) exchange between the two channels is possible for *some* values of  $\xi$  (see [H1] below) and (ii) the free energy is a ‘good bias’ in each channel (see [H4] below). This is formalized in the main result of this paper, namely Theorem 1 below.

For some technical reasons, we were actually unable to prove this result on the original ABF dynamics (8). We will thus consider a slightly different system (that we call *the bi-channel model*) which retains the most important features of the dynamics (8) when applied to a potential exhibiting two parallel channels in the direction of  $\xi$ , separated by a high energy barrier.

The paper is organized as follows. In Section 2 we give details of the bi-channel model, define some probability measures and recall some entropy definitions before presenting the main result. Finally, the proof of the main result is given in Section 3.

## 2 The bi-channel model and statement of the main results

In this section, we present a model to describe the bi-channel scenario. In the following, we treat the case  $d = 2$  (so that the position is  $q = (x, y)$ ),  $\mathcal{D} = \mathbb{T} \times \mathbb{R}$ ,  $\mathcal{M} = \mathbb{T}$  and  $\xi : \mathcal{D} \rightarrow \mathbb{T}$ , where  $\xi(x, y) = x$ . We further assume without loss of generality that  $\beta = 1$ , which can be done by a change of variables  $\tilde{t} = \beta^{-1}t$ ,  $\tilde{\psi}(\tilde{t}, x) = \psi(t, x)$ ,  $\tilde{V}(x, y) = \beta V(x, y)$ . With these assumptions, some notation may be simplified:  $|\nabla\xi| = 1$ ,  $\Sigma_z = \mathbb{R}$  and  $\delta_{\xi(q)-z}(dq) = dy$ . Furthermore, the definition

of the local mean force in (6) simplifies to

$$f = \partial_x V$$

and the free energy and its derivative are given by

$$A(x) = -\ln \int_{\mathbb{R}} \exp(-V(x, y)) dy \quad \text{and} \quad A'(x) = \frac{\int_{\mathbb{R}} \partial_x V(x, y) \exp(-V(x, y)) dy}{\int_{\mathbb{R}} \exp(-V(x, y)) dy}. \quad (10)$$

We would like to emphasize that the choice of the domain  $\mathcal{M} = \mathbb{T}$  and reaction coordinate  $\xi(x, y) = x$  is merely to reduce technicalities, see [14] for appropriate tools to treat general  $\xi$ . In particular, the results can be straightforwardly generalized to the case  $\mathcal{D} = \mathbb{T} \times \mathbb{R}^{d-1}$  and  $\xi(q_1, \dots, q_d) = q_1$ . Likewise, the generalization to a situation with multiple channels (more than two) is straightforward (see Remark 1 below for another generalization).

## 2.1 The bi-channel model

The bi-channel situation is characterized by the existence of two channels joining an initial and final state on a potential energy surface  $V$ , separated from each other by a region of high energy, as depicted in the left of Figure 1. As explained above, we were not able to analyze the original ABF dynamics (8) in this situation, because of some technical difficulties in expressing the probability density flux from one channel to the other.

We therefore analyze the convergence of the ABF method for a slightly different model, which is schematically represented in right of Figure 1. Each channel is described by a potential energy function  $V_i : \mathcal{D} \rightarrow \mathbb{R}$ , where  $i \in \{0, 1\}$  denotes the channel index. The stochastic process we consider now is actually a couple  $(X_t, I_t)$ , where the position vector  $X_t$  lives at time  $t$  on the potential  $V_{I_t}$ ,  $I_t \in \{0, 1\}$  being the index of the visited channel at time  $t$ . The channel index  $I_t$  is allowed to switch to  $1 - I_t$  only if  $\xi(X_t)$  lies in some designated regions (typically at the two ends of the two channels).

The dynamics for the ABF dynamics in the bi-channel model is then

$$\left\{ \begin{array}{l} dX_t = -\nabla(V_{I_t} - A_t \circ \xi)(X_t)dt + \sqrt{2} dB_t, \\ A'_t(x) = \mathbb{E}[\partial_x V_{I_t}(X_t) | \xi(X_t) = x], \\ I_t \in \{0, 1\} \text{ is a jump process with generator} \\ L\varphi(x, y, i) = -\lambda(x)(\varphi(x, y, i) - \varphi(x, y, 1 - i)). \end{array} \right. \quad (11)$$

In terms of the stochastic process  $I_t$ , switching between the two potentials (namely change of  $I_t$  to  $1 - I_t$ ) occurs at times

$$\tau_{n+1} = \inf \left\{ t > \tau_n \mid \int_{\tau_n}^t \lambda(\xi(X_s)) ds > T_n \right\},$$

where  $\tau_0 = 0$  and  $T_n$  are i.i.d. random variables drawn from the exponential distribution with parameter 1. In this way  $\lambda(x)$  denotes the rate at which the trajectories jump from potential  $V_i$  to potential  $V_{1-i}$ . Note that this rate depends only on the position,  $x$ , in the reaction coordinate. The bi-channel feature of the model is related to the fact that we assume that the rate  $\lambda$  is zero (there is no switching) in some region of the reaction coordinate. Outside of this region, the rate is supposed to be constant, and the potential functions are assumed to be identical (the particles live in the same potential). Let us state this as a formal assumption,

$$[\mathbf{H1}] \quad \exists \mathcal{E} \subset \mathbb{T}, \quad \lambda(x) = \lambda \mathbf{1}_{\mathbb{T} \setminus \mathcal{E}}(x) \text{ and } \forall x \in \mathbb{T} \setminus \mathcal{E}, V_0(x, \cdot) = V_1(x, \cdot).$$

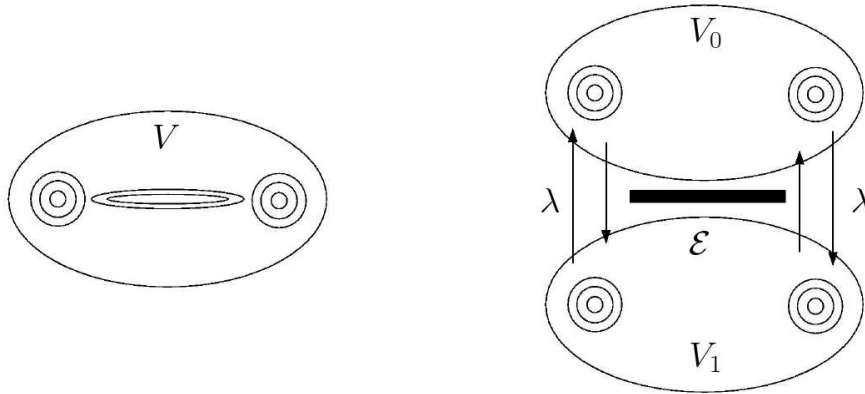


Figure 1: Left: Contour plot of a 2-dimensional potential energy surface demonstrating the bi-channel scenario. Right: In the bi-channel model the two channels are described by two potential functions  $V_i : \mathcal{D} \rightarrow \mathbb{R}$ ,  $i \in \{0, 1\}$ . Exchange between the two channels is allowed only in regions  $\mathbb{T} \setminus \mathcal{E}$  at a rate  $\lambda > 0$ .

The region  $\mathcal{E} \subset \mathbb{T}$  in the above hypothesis represents the region where the two channels are separated by high energy barriers, see Figure 1. It is assumed to have a Lebesgue measure different from 0 and 1.

The main qualitative difference between the bi-channel model we study, and the original ABF dynamics (8) is that the switching only depends on the  $x$ -position and not on the  $y$ -position. However, the ABF dynamics (11) conserves the main difficulty of the original one, namely the metastability of the dynamics in terms of visited channels for some values of  $\xi$ . At times  $t$  such that  $\xi(X_t) \in \mathcal{E}$ ,  $I_t$  cannot switch to  $1 - I_t$ . In particular, it can be checked that the proof of [14] applied to (11) in the case  $\mathcal{E} = \emptyset$  leads to an estimated rate of convergence limited by  $\lambda$  and is thus eventually zero if  $\lambda$  goes to zero (see Remark 3 below). The aim of this work is to study the case  $\mathcal{E} \neq \emptyset$  and to obtain an exponential rate of convergence even if  $\lambda = 0$  in some region (see Theorem 1 below).

**Remark 1** *On Figure 1, we represent the bi-channel case with two metastable states linked by two different channels. The region where the stochastic process can jump from one channel to another has two connected components. We would like to emphasize that our result also applies to the case where this region is simply connected, namely a situation where two channels start from a metastable state (along the reaction coordinate value) but do not end in another metastable state. This situation is of course less favorable in terms of speed of convergence to equilibrium, which would be reflected in our theoretical result through the parameter  $\theta$  (see [H4] below). This is actually typically the situation of the numerical experiments in [15] mentioned above since the various compact states obtained do not belong to the same metastable basin.*

## 2.2 A partial differential equation formulation

Let us introduce the time marginal of the process  $(X_t, I_t)$  in (11):

$$d\mu_t = \sum_{i=0}^1 \psi(t, x, y, i) \delta_i dx dy,$$

where  $\delta_i$  is the Dirac measure on the singleton  $\{i\}$ . When necessary, we shall denote the  $i$ -dependency of the density by a subscript  $\psi_i$ , so that  $\psi_{1-i}$  denotes  $\psi(t, x, y, 1 - i)$  for example. The

evolution of the densities are described by a system of non-linear partial differential equations:  
 $\forall i \in \{0, 1\}$ ,

$$\partial_t \psi = \operatorname{div}(\nabla(V_i - A_t \circ \xi)\psi) + \Delta\psi - \lambda \circ \xi [\psi - \psi_{1-i}] \text{ on } \mathcal{D}, \quad (12)$$

where, we recall, the last term is zero for  $\xi(x, y) \in \mathcal{E}$ . The non-linearity is due to the definition of the mean force estimate, given by

$$A'_t(x) = \frac{\sum_{i=0}^1 \int_{\mathbb{R}} \partial_x V_i(x, y) \psi(t, x, y, i) dy}{\sum_{i=0}^1 \int_{\mathbb{R}} \psi(t, x, y, i) dy}. \quad (13)$$

Using hypothesis [H1], it can be checked that if  $\psi_\infty$  is defined as a probability density proportional to  $e^{-(V_i - A_\infty \circ \xi)}$  where  $A_\infty$  is a given long-time limit for  $A_t$ , then  $\psi_\infty$  is a stationary solution to (12). Then, by replacing  $\psi$  by  $\psi_\infty$  in (13) and comparing with the definition of the mean force (10), it is clear that by choosing  $A'_\infty = A'$ , one obtains a stationary solution of the system (12)–(13) written as:

$$\psi_\infty(x, y, i) = \frac{e^{-(V_i(x, y) - A(x))}}{\sum_{i=0}^1 \int_{\mathcal{D}} e^{-(V_i - A \circ \xi)} dx dy}. \quad (14)$$

The associated equilibrium measure for the process  $(X_t, I_t)$  writes:

$$d\mu_\infty = \sum_{i=0}^1 \psi_\infty(x, y, i) \delta_i dx dy.$$

We will need further notation for marginal and conditional laws associated to  $\mu_t$  and  $\mu_\infty$ . The next two sections give precise definitions for these measures, and Table 1 summarizes the notation.

### 2.2.1 Marginal laws

We are now in the position to define marginal probability densities. The image of the probability measure  $\mu_t$  in  $\xi$  and  $I$  is denoted by

$$d\mu_t^{\xi, I} = \sum_{i=0}^1 \psi^{\xi, I}(t, x, i) \delta_i dx \quad \text{where} \quad \psi^{\xi, I}(t, x, i) = \int_{\mathbb{R}} \psi(t, x, y, i) dy.$$

The evolution of the density  $\psi^{\xi, I}$  is described by the (non-closed) partial differential equation

$$\partial_t \psi^{\xi, I} = \int_{\mathbb{R}} \partial_x (\partial_x (V_i - A_t \circ \xi) \psi) dy + \partial_{xx} \psi^{\xi, I} - \lambda(x) (\psi^{\xi, I} - \psi_{1-i}^{\xi, I}), \quad (15)$$

obtained by integrating (12) in  $y$ . The associated equilibrium measure is

$$d\mu_\infty^{\xi, I} = \sum_{i=0}^1 \psi_\infty^{\xi, I}(x, i) \delta_i dx \quad \text{where} \quad \psi_\infty^{\xi, I}(x, i) = \int_{\mathbb{R}} \psi_\infty(x, y, i) dy.$$

The marginal measure in  $\xi$  only is denoted by

$$d\mu_t^\xi = \psi^\xi(t, x) dx \quad \text{where} \quad \psi^\xi(t, x) = \sum_{i=0}^1 \psi^{\xi, I}(t, x, i) = \sum_{i=0}^1 \int_{\mathbb{R}} \psi(t, x, y, i) dy.$$

Measure	Definition	Description of measure
$d\mu_t$	$\sum_{i=0}^1 \psi(t, x, y, i) \delta_i dx dy$	Probability measure in domain $\mathcal{D} \times \{0, 1\}$
$d\mu_t^{\xi, I}$	$\sum_{i=0}^1 \psi^{\xi, I}(t, x, i) \delta_i dx$	Marginal measure in reaction coordinate and channel
$d\mu_t^{\xi}$	$\psi^{\xi}(t, x) dx$	Marginal measure in reaction coordinate
$d\mu_{t x, i}$	$\frac{\psi(t, x, y, i) dy}{\psi^{\xi, I}(t, x, i)}$	Measure conditioned to being at $x$ and $i$
$d\mu_{t x}$	$\frac{\sum_{i=0}^1 \psi(t, x, y, i) \delta_i dy}{\psi^{\xi}(t, x)}$	Measure conditioned to being at $x$
$d\mu_{t x}^I$	$\frac{\sum_{i=0}^1 \psi^{\xi, I}(t, x, i) \delta_i}{\psi^{\xi}(t, x)}$	Marginal in $I$ , conditioned to being at $x$

Table 1: A list of probability measures.

By summing (15) over  $i$  and using the definition (13) of  $A'_t$ , it is easy to check that  $\psi^{\xi}$  satisfies a closed, very simple partial differential equation (this is similar to the original ABF dynamics, see [13, 14]):

$$\partial_t \psi^{\xi} = \partial_{xx} \psi^{\xi} \quad \text{on } \mathbb{T}. \quad (16)$$

Thanks to the adaptive bias, along  $\xi$ , the barriers have been flattened. The long-time limit of  $\psi^{\xi}$  is given by

$$\psi_{\infty}^{\xi}(x) = \sum_{i=0}^1 \int_{\mathbb{R}} \psi_{\infty}(x, y, i) dy = 1,$$

which corresponds to the uniform probability measure on the torus  $\mathbb{T}$ .

### 2.2.2 Conditional laws

Let us introduce the measure  $\mu_{t|x}$  of  $(X_t, I_t)$  conditioned to being at a specified point  $x$  in the reaction coordinate:

$$d\mu_{t|x} = \frac{\sum_{i=0}^1 \psi(t, x, y, i) \delta_i dy}{\psi^{\xi}(t, x)}. \quad (17)$$

Its long-time limit is

$$d\mu_{\infty|x} = \frac{\sum_{i=0}^1 \psi_{\infty}(x, y, i) \delta_i dy}{\psi_{\infty}^{\xi}(x)}. \quad (18)$$

To study the bi-channel model, we will also need to introduce the measures conditioned to being at fixed  $\xi$  and a particular channel  $i \in \{0, 1\}$ . This measure and its long-time limit are defined

respectively by

$$d\mu_{t|x,i} = \frac{\psi(t, x, y, i) dy}{\psi^{\xi, I}(t, x, i)} \quad \text{and} \quad d\mu_{\infty|x,i} = \frac{\psi_{\infty}(x, y, i) dy}{\psi_{\infty}^{\xi, I}(x, i)}. \quad (19)$$

Finally, we will also need the marginals in  $I$  of the probability measures  $\mu_{t|x}$  and  $\mu_{\infty|x}$ . These Bernoulli probability measures  $\mu_{t|x}^I$  and  $\mu_{\infty|x}^I$  represent the proportion of the marginal distribution  $\psi^{\xi}(t, \cdot)$  and  $\psi^{\xi}$  in each channel. They are formally defined as

$$d\mu_{t|x}^I = \frac{\sum_{i=0}^1 \psi^{\xi, I}(t, x, i) \delta_i}{\psi^{\xi}(t, x)} \quad \text{and} \quad d\mu_{\infty|x}^I = \frac{\sum_{i=0}^1 \psi_{\infty}^{\xi, I}(x, i) \delta_i}{\psi_{\infty}^{\xi}(x)}. \quad (20)$$

### 2.3 Entropy and Fisher information

In this section we recall some well-known results for defining relative entropy between two probability measures, which can be seen as a measure of the ‘distance’ between those. A general introduction to this topic can be found in [1, 18] and applications to study the longtime behavior of Fokker-Planck type equations are presented in [2].

**Definition 1 (Entropy).** For any two probability measures  $\mu$  and  $\nu$  such that  $\mu$  is absolutely continuous with respect to  $\nu$  (denoted as  $\mu \ll \nu$ ), the relative entropy is defined as

$$H(\mu|\nu) = \int \ln \left( \frac{d\mu}{d\nu} \right) d\mu.$$

The positivity of the relative entropy can be shown using the inequality  $\ln(1/x) \geq 1 - x$ . Furthermore,  $H(\mu|\nu) = 0$  if and only if  $\mu = \nu$ .

**Definition 2 (Csiszar-Kullback inequality).** For measures  $\mu$  and  $\nu$  which have densities with respect to the Lebesgue measure, the following holds

$$\|\mu - \nu\|_{L^1} \leq \sqrt{2H(\mu|\nu)}.$$

This allows us to control the  $L^1$ -norm of the difference of two probability measures by their relative entropy.

**Definition 3 (Fisher information).** For any probability measure  $\mu$  absolutely continuous with respect to  $\nu$ , the Fisher information is given by

$$F(\mu|\nu) = \int \left| \nabla \ln \left( \frac{d\mu}{d\nu} \right) \right|^2 d\mu.$$

**Definition 4 (Logarithmic Sobolev inequality).** The probability measure  $\nu$  is said to satisfy a logarithmic Sobolev inequality with constant  $\rho > 0$  (in short:  $LSI(\rho)$ ) if for all probability measures  $\mu$  such that  $\mu \ll \nu$ ,

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

Such an inequality holds for Gaussian measures [7] for example, and more generally [3] for any measure  $\nu$  with density proportional to  $e^{-V}$ , where  $V$  is  $\alpha$ -convex (in which case the LSI constant  $\rho$  is equal to  $\alpha$ ). Besides, there exists a perturbation result [10]: if  $\tilde{\nu}$  is a probability measure such that  $d\tilde{\nu}/d\nu = e^U$ , where  $\nu$  satisfies a  $LSI(\rho)$  and  $U$  is a bounded function, then  $\tilde{\nu}$  satisfies a LSI with constant  $\tilde{\rho} = \rho e^{-\text{osc}(U)}$ , where  $\text{osc}(U) = \sup(U) - \inf(U)$ . Thus, a very large class of probability measures satisfy a LSI. An important feature of the LSI constant is that it degenerates to zero in the case where the probability measure is multimodal. For example, if  $d\nu = Z^{-1} \exp(-\beta W(x)) dx$

and  $W(x) = x^4/4 - x^2/2$  is the double-well potential in dimension 1, then the LSI constant scales as  $\exp(-\beta\Delta W)$  where  $\Delta W = W(0) - W(1) > 0$  is the height of the barrier. Such inequalities thus hold under rather loose assumptions, but the constant  $\rho$  is very small for a multimodal measure. For example, in typical situations encountered in molecular dynamics, the LSI constant for the measure  $\nu$  defined in (1) is extremely small, which is related to the fact that it is difficult to sample directly this measure.

To analyze the convergence of the ABF dynamics for the bi-channel model, we make use of LSIs for the conditional measures  $\mu_{\infty|x,i}$  (see assumption [H3] below), which are the equilibrium canonical measures, conditioned to being at a fixed value of  $\xi$  and in a channel. In [14], log-Sobolev inequalities for the equilibrium canonical measures conditioned only to being at a fixed value of  $\xi$  were considered, but in the bi-channel case, those are typically very small due to the presence of high energy barriers ‘orthogonal’ to the isocontours of the reaction coordinate (see again Figure 1).

Let us now define the Wasserstein distance between two probability measures.

**Definition 5 (Wasserstein distance).** *The Wasserstein distance with linear cost between probability measures  $\mu$  and  $\nu$  is defined as*

$$W(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{D} \times \mathcal{D}} |y - y'| \pi(dy, dy'),$$

where  $\Pi(\mu, \nu)$  denotes the set of coupling probability measures on  $\mathcal{D} \times \mathcal{D}$ , with marginals  $\mu$  and  $\nu$ .

**Definition 6 (Talagrand inequality).** *The probability measure  $\nu$  is said to satisfy a Talagrand inequality with constant  $\rho > 0$  (or  $T(\rho)$ ) if for all probability measures  $\mu$  such that  $\mu \ll \nu$ ,*

$$W(\mu, \nu) \leq \sqrt{\frac{2}{\rho} H(\mu|\nu)}. \quad (21)$$

Logarithmic Sobolev inequalities and Talagrand inequalities are related (see [16]):

**Lemma 1** *If  $\nu$  satisfies  $LSI(\rho)$ , then  $\nu$  satisfies  $T(\rho)$ .*

Below, we present entropies that prove useful in obtaining convergence results of the bi-channel ABF model. In this paper, we are primarily interested in the convergence to a stationary state of the Fokker-Planck equation (12)–(13), and thus of the associated partial differential equations (15) and (16). Relative entropies will therefore be defined for some probability measures with respect to their long-time limits.

The *total entropy* will be denoted by

$$E(t) = H(\mu_t|\mu_\infty) = \sum_{i=0}^1 \int_{\mathcal{D}} \ln \left( \frac{\psi}{\psi_\infty} \right) \psi \, dx \, dy.$$

The so-called *macroscopic entropy* is defined as

$$E_M(t) = H(\mu_t^\xi|\mu_\infty^\xi) = \int_{\mathbb{T}} \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \psi^\xi \, dx,$$

the *local entropy* at a fixed value  $x$  in the reaction coordinate by

$$e_m(t, x) = H(\mu_{t|x}|\mu_{\infty|x}) = \sum_{i=0}^1 \int_{\mathbb{R}} \ln \left( \frac{\psi}{\psi^\xi} / \frac{\psi_\infty}{\psi_\infty^\xi} \right) \frac{\psi}{\psi^\xi} \, dy,$$

and the *microscopic entropy* by

$$E_m(t) = \int_{\mathbb{T}} e_m(t, x) \psi^\xi(t, x) \, dx.$$

With the above definitions, it is easy to show that

$$E(t) = E_M(t) + E_m(t). \quad (22)$$

In order to treat the bi-channel case, we define a *channel-local entropy*, defined by

$$e_{cl}(t, x, i) = H(\mu_{t|x,i} | \mu_{\infty|x,i}) = \int_{\mathbb{R}} \ln \left( \frac{\psi}{\psi^{\xi,I}} \bigg/ \frac{\psi_{\infty}}{\psi_{\infty}^{\xi,I}} \right) \frac{\psi}{\psi^{\xi,I}} dy.$$

Two hypothesis that will be essential to the results presented below are: a so-called ‘bounded coupling’ assumption on the cross derivative  $\partial_{x,y} V_i$  (see [14, 12]), and an assumption on the logarithmic Sobolev inequality constants for the probability measures  $\mu_{\infty|x,i}$ . The hypotheses read

$$[\mathbf{H2}] \quad \left\{ \begin{array}{l} \forall i \in \{0, 1\}, V_i \text{ and } \xi \text{ are sufficiently differentiable functions such that } \exists C, M > 0 \\ \|\partial_{x,y} V_i\|_{L^\infty(\mathbb{T} \times \mathbb{R})} \leq M < \infty \text{ and } \left\| \frac{\int_{\mathbb{R}} \partial_x V_i e^{-V_i} dy}{\int_{\mathbb{R}} e^{-V_i} dy} \right\|_{L^\infty(\mathbb{T})} \leq C < \infty, \end{array} \right.$$

$$[\mathbf{H3}] \quad \left\{ \begin{array}{l} \forall i \in \{0, 1\}, V_i \text{ and } \xi \text{ are such that } \exists \rho > 0, \forall x \in \mathbb{T}, \forall i \in \{0, 1\}, \\ \text{the conditional measure } \mu_{\infty|x,i} \text{ satisfies LSI}(\rho). \end{array} \right.$$

The hypothesis [H3] gives us that  $\forall x \in \mathbb{T}, \forall i \in \{0, 1\}$

$$H(\mu_{t|x,i} | \mu_{\infty|x,i}) \leq \frac{1}{2\rho} F(\mu_{t|x,i} | \mu_{\infty|x,i}),$$

or equivalently

$$\int_{\mathbb{R}} \ln \left( \frac{\psi}{\psi^{\xi,I}} \bigg/ \frac{\psi_{\infty}}{\psi_{\infty}^{\xi,I}} \right) \frac{\psi}{\psi^{\xi,I}} dy \leq \frac{1}{2\rho} \int_{\mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_{\infty}} \right) \right|^2 \frac{\psi}{\psi^{\xi,I}} dy. \quad (23)$$

Finally, an entropy that appears in later calculations, is that of the Bernoulli measure  $\mu_{t|x}^I$  with respect to its long-time limit

$$e_c(t, x) = H(\mu_{t|x}^I | \mu_{\infty|x}^I) = \sum_{i=0}^1 \ln \left( \frac{\psi^{\xi,I}}{\psi^{\xi}} \bigg/ \frac{\psi_{\infty}^{\xi,I}}{\psi_{\infty}^{\xi}} \right) \frac{\psi^{\xi,I}}{\psi^{\xi}}.$$

The so-called *channel entropy* is then given by

$$\begin{aligned} E_c(t) &= \int_{\mathbb{T}} e_c(t, x) \psi^{\xi} dx \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \ln \left( \frac{\psi^{\xi,I}}{\psi^{\xi}} \bigg/ \frac{\psi_{\infty}^{\xi,I}}{\psi_{\infty}^{\xi}} \right) \psi^{\xi,I}. \end{aligned} \quad (24)$$

## 2.4 The free energy as a bias in each channel

As well as the hypothesis [H3] on the conditional measures  $\mu_{\infty|x,i}$ , an assumption will also be necessary to ensure that the free energy is a ‘good bias’ in each channel. More precisely, once the bias has converged, the marginals along  $\xi$  in each channel must converge sufficiently quickly to their long-time limit. Roughly speaking, the channels should not be too ‘asymmetrical’. The aim of this section is to state this more formally, see assumption [H4] below.

Consider that the system is already nearly at equilibrium, in the sense that

$$i) \quad A'_t = A'_\infty = A',$$

$$\text{ii) } \forall i \in \{0, 1\}, \int_{\mathbb{R}} \partial_x V_i d\mu_{t|x,i} = \int_{\mathbb{R}} \partial_x V_i d\mu_{\infty|x,i}.$$

Then the marginal density  $\psi^{\xi,I}$  can be shown to satisfy (see (15) and (40) below)

$$\partial_t \psi^{\xi,I} = -\mathcal{L}_i \psi^{\xi,I} = \partial_x (\psi_{\infty}^{\xi,I} \partial_x (\psi^{\xi,I} / \psi_{\infty}^{\xi,I})) - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}). \quad (25)$$

It can be shown that the operator  $\mathcal{L} = (\mathcal{L}_0, \mathcal{L}_1)$  is symmetric and positive definite with respect to the inner product  $\langle f, g \rangle = \sum_{i=0}^1 \int_{\mathbb{T}} f_i(x) g_i(x) \frac{1}{\psi_{\infty}^{\xi,I}(x,i)} dx$  and has a spectral gap  $\theta > 0$  (see Lemma 12 below). In other words,  $\mathcal{L}$  is such that for all functions  $f : \mathbb{T} \times \{0, 1\} \rightarrow \mathbb{R}$ , with  $f_i \in H^1\left(\frac{1}{\psi_{\infty}^{\xi,I}(x,i)} dx\right)$ ,  $f \neq 0$  and  $\sum_{i=0}^1 \int_{\mathbb{T}} f_i(x) dx = 0$ , we have

$$\langle f, f \rangle \leq \frac{1}{\theta} \langle f, \mathcal{L}f \rangle. \quad (26)$$

As will be apparent in the proof, it will be necessary for the spectral gap  $\theta$  to be sufficiently large.

$$[\mathbf{H4}] \quad \left\{ \begin{array}{l} \theta > \theta_{\min} \text{ with } \theta_{\min} = \frac{8(C + M\rho^{-1/2})^2 \tilde{M}}{c} \\ \text{where } 0 < c, \tilde{M} < \infty \text{ are such that } \inf_{x,i} \psi_{\infty}^{\xi,I} = c \text{ and } \sup_x \psi^{\xi}(0, x) = \tilde{M}. \end{array} \right.$$

We recall that the constants  $C$ ,  $M$  and  $\rho$  have been introduced in assumptions [H2] and [H3] above. The fact that  $c > 0$  is not restrictive since  $\forall i \in \{0, 1\}$ ,  $V_i$  is a continuous function and  $\psi_{\infty}^{\xi,I} \propto \int_{\mathbb{R}} e^{-V_i} dy > 0$  is continuous and defined on the compact space  $\mathbb{T}$ . Similarly, since  $\psi^{\xi}$  satisfies the heat equation (16), the assumption  $\tilde{M} < \infty$  is not restrictive. If, for example, the initial condition has a Dirac mass marginal in  $\xi$ , for any positive time  $t_0 > 0$ ,  $\psi^{\xi}(t_0, \cdot)$  is a bounded function, and one has simply to consider the dynamics on  $[t_0, \infty)$ .

**Remark 2** *In the hypothesis [H2], the assumption on the cross derivative  $\partial_{x,y} V_i$  could in fact be replaced by  $\|\partial_x V_i\|_{L^\infty(\mathbb{T} \times \mathbb{R})} \leq M < \infty$ , in which case, in [H4], the minimum value for  $\theta$  would be changed to  $\theta_{\min} = \frac{8M^2 \tilde{M}}{c}$  and the Talagrand inequalities in Lemmas 6 and 11 would be replaced by Csiszar-Kullback inequalities.*

## 2.5 Main result

We are now in position to present the main result of the paper.

**Theorem 1** *Assume hypotheses [H1]-[H4]. There exists a smooth function  $\Lambda : (\theta_{\min}, \infty) \rightarrow (0, \rho)$  which is increasing and such that:*

$$\Lambda(\rho + 2\theta_{\min}) = \frac{\rho}{2} \quad \text{and} \quad \Lambda(\theta) \rightarrow \begin{cases} 0 & \text{as } \theta \rightarrow \theta_{\min} \\ \rho & \text{as } \theta \rightarrow \infty \end{cases}$$

for which we can prove the following convergence result: for any  $\varepsilon \in (0, \Lambda(\theta))$ , there exists a constant  $K > 0$  such that,  $\forall t > 0$ ,

$$E_m(t) \leq K \exp(-2 \min\{(\Lambda(\theta) - \varepsilon), 4\pi^2\} t). \quad (27)$$

This implies that the total entropy  $E$  and thus  $\|\psi(t, \cdot) - \psi_{\infty}\|_{L^1}^2$  converge exponentially fast to zero with the same rate. Furthermore, the biasing force  $A'_t$  converges to the mean force  $A'$  in the following sense:  $\forall t \geq 0$ ,

$$\int_{\mathbb{T}} |A'_t(x) - A'(x)|^2 \psi^{\xi}(t, x) dx \leq 2(C + M\rho^{-1/2})^2 E_m(t).$$

As a consequence, for any positive time  $t_0 > 0$  and  $\varepsilon \in (0, \Lambda(\theta))$ , there exists a constant  $\bar{K}$  such that  $\forall t \geq t_0$ ,

$$\int_{\mathbb{T}} |A'_t(x) - A'(x)|^2 dx \leq \bar{K} \exp(-2 \min\{\Lambda(\theta) - \varepsilon, 4\pi^2\} t).$$

The term  $4\pi^2$  corresponds to the exponential rate of convergence of  $\psi^\xi$  to  $\psi_\infty^\xi$  (see (16)), and is clearly not the bottleneck. There are actually various ways to make this rate as small as needed (see Remark 11 in [14]).

Thus, this result essentially shows that the ABF method converges at a rate which is limited by the multimodality of the equilibrium canonical measures conditioned to being at fixed value of  $\xi$  and in a channel (quantified by the constant  $\rho$ ), if the free energy is a bias which enables a fast exploration of each channel (this is quantified by the constant  $\theta$ , which should be sufficiently large for  $\Lambda(\theta)$  to be indeed close to  $\rho$ ). Thus, the convergence may be fast even if switching between the two channels is impossible for some values of  $\xi$ . If the spectral gap  $\theta$  is sufficiently large, we thus recover a similar expression for the rate of convergence of the ABF method as the one derived in [14], with  $\rho$  being now the LSI constant of the canonical measures  $\mu_{|x,i}$ .

We would like to emphasize that our arguments hold under the following assumption of existence of regular solutions: We assume that we are given a process  $(X_t, I_t)$  and a function  $A'_t$  which satisfy (11), and such that  $X_t$ , conditionally on  $I_t = i$  has a smooth density  $\psi(t, x, y, i)$ . We suppose that this density is sufficiently regular so that the entropy estimates below are valid. We refer for example to [2] for an appropriate functional framework in which such entropy estimates hold.

### 3 Proof of main result

In order to prove the exponential decay of the microscopic entropy in Theorem 1, we use the fact that the time evolution of the microscopic entropy can be expressed as a combination of the evolution of the total and the marginal entropies, from (22)

$$\frac{dE_m}{dt} = \frac{dE}{dt} - \frac{dE_M}{dt}.$$

In order to obtain results for the microscopic entropy, we begin by treating the time evolutions of the total entropy and the macroscopic entropy separately.

#### 3.1 Preliminary computations on the total entropy $E$

In order to study the evolution of the total entropy, some auxiliary results will be needed and are given in the lemmas below. First, it will be useful to write the Fokker-Planck equation associated to  $\psi$  in a different form.

**Lemma 2** *The Fokker-Planck equation (12) for  $\psi$  can be rewritten as*

$$\partial_t \psi = \operatorname{div}(\psi_\infty \nabla(\psi/\psi_\infty)) + \partial_x((A' - A'_t)\psi) - \lambda(x)(\psi - \psi_{1-i}).$$

*Proof :* By developing the right hand side, we obtain

$$\begin{aligned} \partial_t \psi &= \operatorname{div}\left(\nabla\psi - \frac{\psi}{\psi_\infty}\nabla\psi_\infty\right) + \partial_x((A' - A'_t)\psi) - \lambda(x)(\psi - \psi_{1-i}) \\ &= \operatorname{div}(\nabla\psi + \nabla(V_i - A \circ \xi)\psi) + \partial_x(A'\psi) - \partial_x(A'_t\psi) - \lambda(x)(\psi - \psi_{1-i}) \\ &= \operatorname{div}(\nabla(V_i - A_t \circ \xi)\psi) + \Delta\psi - \lambda(x)(\psi - \psi_{1-i}), \end{aligned}$$

which is indeed the Fokker-Planck equation (12). ◇

The above may now be used to estimate the time evolution of the total entropy.

**Lemma 3** *The total entropy  $E$  satisfies*

$$\frac{dE}{dt} \leq - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \nabla \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} (A' - A'_t) \partial_x \left[ \ln \left( \frac{\psi}{\psi_\infty} \right) \right] \psi. \quad (28)$$

*Proof :* First, by definition of the total entropy, we have

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt} \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \ln \left( \frac{\psi}{\psi_\infty} \right) \psi \\ &= \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \ln \left( \frac{\psi}{\psi_\infty} \right) \partial_t \psi, \end{aligned} \quad (29)$$

using the fact that  $\psi$  is a probability density. Next, we use Lemma 2 to obtain

$$\begin{aligned} \frac{dE}{dt} &= \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \operatorname{div} \left( \psi_\infty \nabla \left( \frac{\psi}{\psi_\infty} \right) \right) \ln \left( \frac{\psi}{\psi_\infty} \right) + \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \partial_x ((A' - A'_t) \psi) \ln \left( \frac{\psi}{\psi_\infty} \right) \\ &\quad - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \lambda(x) \ln \left( \frac{\psi}{\psi_\infty} \right) (\psi - \psi_{1-i}) \\ &\leq - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \nabla \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} (A' - A'_t) \partial_x \left[ \ln \left( \frac{\psi}{\psi_\infty} \right) \right] \psi, \end{aligned} \quad (30)$$

where the last line is a result of integration by parts and the fact that (30) is non-positive, which is proved below in Lemma 4.  $\diamond$

**Lemma 4** *For  $\psi$  satisfying (12) and  $\psi_\infty$  its long-time limit, the following holds*

$$- \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \lambda(x) \ln \left( \frac{\psi}{\psi_\infty} \right) (\psi - \psi_{1-i}) \leq 0.$$

*Proof :* First, recall that  $\lambda(x) = 0, \forall x \in \mathcal{E}$ . We consider therefore the left hand side of the inequality at fixed  $(x, y) \in (\mathbb{T} \setminus \mathcal{E}) \times \mathbb{R}$ . At fixed  $(x, y)$ , we consider renormalized (Bernoulli) probabilities denoted by  $\tilde{\psi}_i = \psi_i / (\psi_0 + \psi_1)$ , so that  $\tilde{\psi}_0 + \tilde{\psi}_1 = 1$ .

$$\begin{aligned} & - \sum_{i=0}^1 \ln \left( \frac{\psi}{\psi_\infty} \right) (\psi - \psi_{1-i}) \\ &= (-\psi_0 + \psi_1) \left[ \ln \left( \frac{\psi_0}{\psi_{\infty,0}} \right) - \ln \left( \frac{\psi_1}{\psi_{\infty,1}} \right) \right] \\ &= (\psi_0 + \psi_1) (-\tilde{\psi}_0 + \tilde{\psi}_1) \left[ \ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) - \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \right] \\ &= (\psi_0 + \psi_1) \left[ -\ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) \tilde{\psi}_0 - \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \tilde{\psi}_1 + \ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) \tilde{\psi}_1 + \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \tilde{\psi}_0 \right] \\ &= (\psi_0 + \psi_1) \left[ -2 \ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) \tilde{\psi}_0 - 2 \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \tilde{\psi}_1 + (\tilde{\psi}_0 + \tilde{\psi}_1) \ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) + (\tilde{\psi}_0 + \tilde{\psi}_1) \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \right] \\ &= (\psi_0 + \psi_1) \left[ -2 \sum_{i=0}^1 \ln \left( \frac{\tilde{\psi}_i}{\tilde{\psi}_{\infty,i}} \right) \tilde{\psi}_i + \ln \left( \frac{\tilde{\psi}_0}{\tilde{\psi}_{\infty,0}} \right) + \ln \left( \frac{\tilde{\psi}_1}{\tilde{\psi}_{\infty,1}} \right) \right]. \end{aligned}$$

From hypothesis [H1],  $\forall(x, y) \in (\mathbb{T} \setminus \mathcal{E}) \times \mathbb{R}$ ,  $\tilde{\psi}_{\infty,1} = \tilde{\psi}_{\infty,0} = \frac{1}{2}$ , which allows the above to be written as

$$\begin{aligned} -\sum_{i=0}^1 \ln\left(\frac{\psi}{\psi_\infty}\right) (\psi - \psi_{1-i}) &= (\psi_0 + \psi_1) \left[ -2 \sum_{i=0}^1 \ln\left(\frac{\tilde{\psi}_i}{\tilde{\psi}_\infty}\right) \tilde{\psi}_i + 2 \sum_{i=0}^1 \ln\left(\frac{\tilde{\psi}_i}{\tilde{\psi}_\infty}\right) \tilde{\psi}_\infty \right] \\ &= -2(\psi_0 + \psi_1) \left[ \sum_{i=0}^1 \ln\left(\frac{\tilde{\psi}_i}{\tilde{\psi}_\infty}\right) \tilde{\psi}_i + \sum_{i=0}^1 \ln\left(\frac{\tilde{\psi}_\infty}{\tilde{\psi}_i}\right) \tilde{\psi}_\infty \right] \\ &\leq 0. \end{aligned} \quad (31)$$

The last line is due to the fact that the two terms between brackets are non-negative, since they are relative entropies.  $\diamond$

**Remark 3** Let us consider the case where  $\mathcal{E} = \emptyset$  and thus  $\forall x \in \mathbb{T}$ ,  $\lambda(x) = \lambda > 0$ , which implies (see [H1])  $V_0 = V_1$  and therefore  $\psi_{\infty,0} = \psi_{\infty,1}$  everywhere on  $\mathbb{T} \times \mathbb{R}$ . In this case, it follows from (31) that

$$\begin{aligned} -\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \lambda(x) \ln\left(\frac{\psi}{\psi_\infty}\right) (\psi - \psi_{1-i}) \\ \leq -2\lambda \int_{\mathbb{T} \times \mathbb{R}} (\psi_0 + \psi_1) \sum_{i=0}^1 \ln\left(\frac{\tilde{\psi}_i}{\tilde{\psi}_\infty}\right) \tilde{\psi}_i \\ \leq -2\lambda \int_{\mathbb{T} \times \mathbb{R}} \sum_{i=0}^1 \ln\left(\frac{\psi}{\psi_\infty}\right) \psi + 2\lambda \int_{\mathbb{T} \times \mathbb{R}} \ln\left(\frac{\psi_0 + \psi_1}{\psi_{\infty,0} + \psi_{\infty,1}}\right) (\psi_0 + \psi_1). \end{aligned}$$

Furthermore, since the marginal  $\psi^{x,y} := \psi_0 + \psi_1$  satisfies in this specific case

$$\partial_t \psi^{x,y} = \operatorname{div}(\psi_\infty^{x,y} \nabla(\psi^{x,y}/\psi_\infty^{x,y})) + \partial_x((A' - A'_t)\psi^{x,y}), \quad (32)$$

one can show, using the results of [14], that  $\exists C > 0$ ,  $\forall t \geq 0$ ,

$$-\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \lambda(x) \ln\left(\frac{\psi}{\psi_\infty}\right) (\psi - \psi_{1-i}) \leq -2\lambda E + 2\lambda C e^{-2 \min\{\rho, 4\pi^2\} t}.$$

Using this, and the fact that  $-E \leq -E_m$ , one can show that  $E_m$  converges to zero exponentially fast with rate

$$2 \min\{\rho, 4\pi^2, \lambda\}.$$

The convergence rate thus depends on  $\lambda > 0$ , the rate at which switching occurs between the two channels. This is comparable to the original result obtained for the ABF algorithm, see [14]. The aim of what follows is to obtain a result in the case where  $\mathcal{E} \neq \emptyset$ , namely when  $\lambda = 0$  in some region.

In order to estimate the last term on the right hand side of (28), it will be helpful to express the difference of the biasing force and the mean force in terms of densities

**Lemma 5** The difference between the biasing force  $A'_t$  and the mean force  $A'$  can be expressed in the following way

$$A'_t - A' = \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln\left(\frac{\psi}{\psi_\infty}\right) \frac{\psi}{\psi^\xi} dy - \partial_x \ln\left(\frac{\psi^\xi}{\psi_\infty^\xi}\right).$$

*Proof* : We develop the expression on the right hand side and use the fact that  $\psi_\infty^\xi \equiv 1$

$$\begin{aligned}
& \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \frac{\psi}{\psi^\xi} dy - \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \\
&= \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln(\psi) \frac{\psi}{\psi^\xi} dy - \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln(\psi_\infty) \frac{\psi}{\psi^\xi} dy - \partial_x \ln(\psi^\xi) \\
&= \sum_{i=0}^1 \int_{\mathbb{R}} \frac{\partial_x \psi}{\psi^\xi} dy + \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x (V_i - A \circ \xi) \frac{\psi}{\psi^\xi} dy - \frac{\partial_x \psi^\xi}{\psi^\xi} \\
&= \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x V_i \frac{\psi}{\psi^\xi} dy - \sum_{i=0}^1 \int_{\mathbb{R}} A' \frac{\psi}{\psi^\xi} dy \\
&= A'_t - A',
\end{aligned}$$

The last line is a result of the definition of  $A'_t$  in (13) and the fact that  $A$  is a function of  $x$  only.  $\diamond$

Another useful estimate for the difference between  $A'_t$  and  $A'$  is given in the following lemma.

**Lemma 6** *The difference of the biasing force and the mean force can be bounded by the microscopic entropy as:*

$$\int_{\mathbb{T}} |A'_t - A'|^2 \psi^\xi dx \leq 2R^2 E_m(t),$$

where

$$R = \left( C + M\rho^{-1/2} \right).$$

*Proof :* We begin by showing that  $|A'_t(x) - A'(x)| \leq M\sqrt{2e_m(t, x)/\rho}$ . By definition, we have

$$\begin{aligned}
A'_t(x) - A'(x) &= \frac{\sum_{i=0}^1 \int_{\mathbb{R}} \partial_x V_i \psi dy}{\sum_{i=0}^1 \int_{\mathbb{R}} \psi dy} - \frac{\sum_{i=0}^1 \int_{\mathbb{R}} \partial_x V_i \psi_\infty dy}{\sum_{i=0}^1 \int_{\mathbb{R}} \psi_\infty dy} \\
&= \sum_{i=0}^1 \int_{\mathbb{R}} \left( \left( \partial_x V_i \frac{\psi^{\xi, I}}{\psi^{\xi, I}} \right) \frac{\psi}{\psi^{\xi, I}} - \left( \partial_x V_i \frac{\psi_\infty^{\xi, I}}{\psi_\infty^{\xi, I}} \right) \frac{\psi_\infty}{\psi_\infty^{\xi, I}} \right) dy \\
&= \sum_{i=0}^1 \int_{\mathbb{R} \times \mathbb{R}} \left( \partial_x V_i(x, y) \frac{\psi^{\xi, I}}{\psi^\xi} - \partial_x V_i(x, y') \frac{\psi_\infty^{\xi, I}}{\psi_\infty^\xi} \right) \pi(dy, dy')
\end{aligned}$$

where  $\pi(dy, dy')$  is any coupling measure on  $\mathbb{R} \times \mathbb{R}$  with marginals  $\mu_{t|x, i}$  and  $\mu_{\infty|x, i}$ . Next, using Taylor's expansion on  $\partial_x V_i(x, y)$ , we have

$$\begin{aligned}
A'_t(x) - A'(x) &= \sum_{i=0}^1 \int_{\mathbb{R} \times \mathbb{R}} \left( (\partial_x V_i(x, y') + \partial_{x, y} V_i(x, \eta(y, y'))(y - y')) \frac{\psi^{\xi, I}}{\psi^\xi} - \partial_x V_i(x, y') \frac{\psi_\infty^{\xi, I}}{\psi_\infty^\xi} \right) \pi(dy, dy') \\
&= \sum_{i=0}^1 \int_{\mathbb{R} \times \mathbb{R}} \left( \partial_x V_i(x, y') \left( \frac{\psi^{\xi, I}}{\psi^\xi} - \frac{\psi_\infty^{\xi, I}}{\psi_\infty^\xi} \right) + \partial_{x, y} V_i(x, \eta(y, y'))(y - y') \frac{\psi^{\xi, I}}{\psi^\xi} \right) \pi(dy, dy') \\
&= \sum_{i=0}^1 \left( \frac{\psi^{\xi, I}}{\psi^\xi} - \frac{\psi_\infty^{\xi, I}}{\psi_\infty^\xi} \right) \int_{\mathbb{R}} \partial_x V_i(x, y') \frac{\psi_\infty}{\psi_\infty^{\xi, I}} dy' + \sum_{i=0}^1 \frac{\psi^{\xi, I}}{\psi^\xi} \int_{\mathbb{R} \times \mathbb{R}} \partial_{x, y} V_i(x, \eta(y, y'))(y - y') \pi(dy, dy')
\end{aligned}$$

where  $\eta(y, y') \in [y, y']$ . Recall from [H2] that  $\exists C, M > 0$  such that  $\left\| \int_{\mathbb{R}} \partial_x V_i d\mu_{\infty|x, i} \right\|_{L^\infty(\mathbb{T})} \leq C$  and  $\|\partial_{x,y} V_i\|_{L^\infty} \leq M$ . Furthermore, with the use of the Csiszar-Kullback inequality, we have

$$\begin{aligned} |A'_t(x) - A'(x)| &\leq C \sum_{i=0}^1 \left| \frac{\psi^{\xi, I}}{\psi^\xi} - \frac{\psi_{\infty}^{\xi, I}}{\psi_{\infty}^\xi} \right| + M \sum_{i=0}^1 \frac{\psi^{\xi, I}}{\psi^\xi} \int_{\mathbb{R} \times \mathbb{R}} |y - y'| \pi(dy, dy') \\ &\leq C \sqrt{2H(\mu_{t|x}^I, \mu_{\infty|x}^I)} + M \sum_{i=0}^1 \frac{\psi^{\xi, I}}{\psi^\xi} \int_{\mathbb{R} \times \mathbb{R}} |y - y'| \pi(dy, dy'). \end{aligned}$$

Next, by the Talagrand inequality (21) and the concavity of the function  $x \mapsto \sqrt{x}$ ,

$$\begin{aligned} |A'_t(x) - A'(x)| &\leq C \sqrt{2H(\mu_{t|x}^I, \mu_{\infty|x}^I)} + M \sum_{i=0}^1 \frac{\psi^{\xi, I}}{\psi^\xi} \sqrt{\frac{2}{\rho} H(\mu_{t|x, i}, \mu_{\infty|x, i})} \\ &\leq C \sqrt{2H(\mu_{t|x}, \mu_{\infty|x})} + M \sqrt{\frac{2}{\rho} \sum_{i=0}^1 H(\mu_{t|x, i}, \mu_{\infty|x, i}) \frac{\psi^{\xi, I}}{\psi^\xi}} \\ &\leq C \sqrt{2H(\mu_{t|x}, \mu_{\infty|x})} + M \sqrt{\frac{2}{\rho} H(\mu_{t|x}, \mu_{\infty|x})} \end{aligned}$$

Therefore

$$|A'_t(x) - A'(x)|^2 \leq 2 \left( C + M\rho^{-1/2} \right)^2 e_m(t, x).$$

The result follows immediately, since  $E_m(t) = \int_{\mathbb{T}} e_m(t, x) \psi^\xi dx$ .  $\diamond$

We are now equipped with the right tools to control  $dE/dt$ , and are left to handle the evolution of the macroscopic entropy.

### 3.2 Controlling $E_M$

Due to the free diffusion equation (16) satisfied by  $\psi^\xi$ , the macroscopic entropy  $E_M$  is easily controlled.

**Lemma 7** *The macroscopic entropy satisfies*

$$\frac{dE_M}{dt} = - \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_{\infty}^\xi} \right) \right|^2 \psi^\xi = -F(\psi^\xi | \psi_{\infty}^\xi). \quad (33)$$

*Proof :* Using (16) and integration by parts

$$\begin{aligned} \frac{dE_M}{dt} &= \frac{d}{dt} \int_{\mathbb{T}} \ln \left( \frac{\psi^\xi}{\psi_{\infty}^\xi} \right) \psi^\xi \\ &= \int_{\mathbb{T}} \ln \left( \frac{\psi^\xi}{\psi_{\infty}^\xi} \right) \partial_{xx} \psi^\xi \\ &= - \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_{\infty}^\xi} \right) \right|^2 \psi^\xi. \end{aligned}$$

**Lemma 8** *The Fisher information for the marginal density  $\psi^\xi$  decays exponentially fast with rate  $r = 8\pi^2$*

$$F(\psi^\xi(t, \cdot) | \psi_{\infty}^\xi) \leq F_0 \exp(-8\pi^2 t),$$

where  $F_0 = F(\psi^\xi(0, \cdot) | \psi_{\infty}^\xi)$ .

*Proof* : See Lemma 12 of reference [14].

◇

In light of the lemmas presented above, hypotheses [H1]-[H4] may now be used to control the evolution of the microscopic entropy.

### 3.3 Controlling $E_m$

The aim of this section is to obtain an estimation on the evolution of th microscopic entropy  $E_m$ , see (36) below. We first begin by using the fact that the channel-local conditional measures  $\mu_{\infty|x,i}$  satisfy LSI( $\rho$ ) to bound the microscopic entropy.

**Lemma 9** *Under the hypothesis [H3], the microscopic entropy  $E_m$  satisfies*

$$E_m \leq \frac{1}{2\rho} \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi \, dx \, dy + E_c,$$

where  $E_c$  is defined in (24).

*Proof* : We fist consider the local entropy  $e_m(t, x)$ , which can be decomposed into the sum of the channel-local entropy and the entropy of the Bernoulli measures.

$$\begin{aligned} e_m(t, x) &= H(\mu_{t|x} | \mu_{\infty|x}) \\ &= \sum_{i=0}^1 \int_{\mathbb{R}} \ln \left( \frac{\psi}{\psi^\xi} / \frac{\psi_\infty}{\psi_\infty^\xi} \right) \frac{\psi}{\psi^\xi} \, dy \\ &= \sum_{i=0}^1 \left[ \int_{\mathbb{R}} \ln \left( \frac{\psi}{\psi^{\xi,I}} / \frac{\psi_\infty}{\psi_\infty^{\xi,I}} \right) \frac{\psi}{\psi^\xi} \, dy + \int_{\mathbb{R}} \ln \left( \frac{\psi^{\xi,I}}{\psi^\xi} / \frac{\psi_\infty^{\xi,I}}{\psi_\infty^\xi} \right) \frac{\psi}{\psi^\xi} \, dy \right] \\ &\leq \sum_{i=0}^1 \left[ \frac{1}{2\rho} \int_{\mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \frac{\psi}{\psi^\xi} \, dy \right] + \sum_{i=0}^1 \left[ \ln \left( \frac{\psi^{\xi,I}}{\psi^\xi} / \frac{\psi_\infty^{\xi,I}}{\psi_\infty^\xi} \right) \frac{\psi^{\xi,I}}{\psi^\xi} \right] \\ &= \frac{1}{2\rho} \sum_{i=0}^1 \int_{\mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \frac{\psi}{\psi^\xi} \, dy + e_c(t, x), \end{aligned}$$

where the inequality is a direct result of [H3]. The microscopic entropy is then

$$\begin{aligned} E_m(t) &= \int_{\mathbb{T}} e_m(t, x) \psi^\xi \, dx \\ &\leq \frac{1}{2\rho} \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi \, dx \, dy + E_c, \end{aligned}$$

as required.

◇

The time evolution of  $E_m$  may now be expressed using results of Lemmas 3 and 7.

$$\begin{aligned} \frac{dE_m}{dt} &= \frac{dE}{dt} - \frac{dE_M}{dt} \\ &\leq - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \nabla \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} (A' - A'_i) \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \psi + \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right|^2 \psi^\xi. \end{aligned}$$

We may now apply Lemma 5 and use integration by parts

$$\begin{aligned}
\frac{dE_m}{dt} &= -\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi \\
&\quad + \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left( \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \frac{\psi}{\psi^\xi} dy - \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right) \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \psi \\
&\quad + \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right|^2 \psi^\xi \\
&= -\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi \\
&\quad - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi + \int_{\mathbb{T}} \left( \sum_{i=0}^1 \int_{\mathbb{R}} \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \psi dy \right)^2 \frac{1}{\psi^\xi} \\
&\quad - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \psi + \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right|^2 \psi^\xi.
\end{aligned} \tag{34}$$

Notice that (34) is non-positive by the Cauchy-Schwartz inequality. We therefore have

$$\begin{aligned}
\frac{dE_m}{dt} &\leq -\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi \\
&\quad - \sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \partial_x \ln \left( \frac{\psi}{\psi_\infty} \right) \psi + \int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right|^2 \psi^\xi \\
&= -\sum_{i=0}^1 \int_{\mathbb{T} \times \mathbb{R}} \left| \partial_y \ln \left( \frac{\psi}{\psi_\infty} \right) \right|^2 \psi - \int_{\mathbb{T}} \left[ \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \psi^\xi \right] (A'_t - A') \\
&\leq -2\rho E_m + 2\rho E_c + \sqrt{\int_{\mathbb{T}} |A'_t - A'|^2 \psi^\xi} \sqrt{\int_{\mathbb{T}} \left| \partial_x \ln \left( \frac{\psi^\xi}{\psi_\infty^\xi} \right) \right|^2 \psi^\xi}.
\end{aligned} \tag{35}$$

Line (35) is a result of Lemma 5 and the last inequality is due to Lemma 9 and a further application of the Cauchy-Schwartz inequality. Now, using Lemmas 6 and 8, we obtain

$$\frac{dE_m}{dt} \leq -2\rho E_m + 2\rho E_c + R\sqrt{2E_m} \sqrt{F_0 e^{-8\pi^2 t}},$$

where we recall  $F_0 = F(\psi^\xi(0, \cdot) | \psi_\infty^\xi)$ . Finally, using Young's inequality:  $\forall \varepsilon > 0, \forall a, b \in \mathbb{R}, ab < \varepsilon a^2 + \frac{1}{4\varepsilon} b^2$ , we obtain

$$\frac{dE_m}{dt} \leq -2(\rho - R^2\varepsilon) E_m + 2\rho E_c + \frac{1}{4\varepsilon} F_0 e^{-8\pi^2 t}, \tag{36}$$

where  $\varepsilon > 0$  will be chosen optimally later in the proof. We are left to control the channel entropy term  $E_c$  in order to conclude.

### 3.4 Controlling $E_c$

The aim of this section is to obtain a control on the evolution of  $E_c$  defined by (38), and more precisely an upper bound of  $E_c$ , denoted  $P$ , which is the weighted  $\chi^2$ -distance between  $\psi^{\xi, I}$  and  $\psi_\infty^{\xi, I}$ , see (38) and (49) below.

Recall that  $E_c = \int_{\mathbb{T}} H(\mu_{t|x}^I | \mu_{\infty|x}^I) \psi^\xi dx$ , where the integrand is the relative entropy of a Bernoulli measure. Poincaré and logarithmic Sobolev inequalities have been studied for Bernoulli measures [1]. In order to obtain an exponentially decaying relative entropy, however, a suitable semi-group and its infinitesimal generator is needed for the measure. In the case of the bi-channel model (in particular [H1]), we face problems due to the region  $\mathcal{E} \subset \mathbb{T}$ , where no exchange is permitted between the two channels: in this region, the speed at which the measure  $\mu_{t|x}^I$  reaches equilibrium cannot directly be controlled. To circumvent this issue, we consider the spectral gap of an adequate operator and resort to the Poincaré inequality.

By the definition of  $E_c$  and using the inequality  $\forall x > 0, x \ln(x) \leq x(x-1)$  and the fact that  $\sum_{i=0}^1 \int_{\mathbb{T}} \psi^{\xi,I}(t, x, i) dx = 1$ , we obtain

$$\begin{aligned} E_c &= \sum_{i=0}^1 \int_{\mathbb{T}} \ln \left( \frac{\psi^{\xi,I}}{\psi^\xi} / \frac{\psi_{\infty}^{\xi,I}}{\psi_{\infty}^\xi} \right) \psi^{\xi,I} \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \ln (\psi^{\xi,I} / \psi_{\infty}^{\xi,I}) \psi^{\xi,I} - E_M \\ &\leq \sum_{i=0}^1 \int_{\mathbb{T}} \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} - 1 \right)^2 \psi_{\infty}^{\xi,I} dx - E_M. \end{aligned} \quad (37)$$

We therefore have

$$E_c \leq P, \quad (38)$$

where

$$P = \sum_{i=0}^1 \int_{\mathbb{T}} \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} - 1 \right)^2 \psi_{\infty}^{\xi,I} dx. \quad (39)$$

In order to proceed and consider the time derivative of  $P$ , we will need some further results to express the evolution of the marginal density  $\psi^{\xi,I}$ . The idea is to compare the evolution of  $\psi^{\xi,I}$  with the dynamics of this density if  $A'_t$  and  $\int_{\mathbb{T}} \partial_x V d\mu_{t|x,i}$  were already at equilibrium (see Section 2.4).

**Lemma 10** *The Fokker-Planck equation (15) for  $\psi^{\xi,I}$  can be rewritten as*

$$\begin{aligned} \partial_t \psi^{\xi,I} &= \partial_x \left( \psi_{\infty}^{\xi,I} \partial_x \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} \right) \right) + \partial_x \left( \left( \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi,I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi,I}} \right) \psi^{\xi,I} \right) \\ &\quad + \partial_x ((A' - A'_t) \psi^{\xi,I}) - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}) \end{aligned} \quad (40)$$

*Proof* : First we will show that

$$\partial_t \psi^{\xi,I} = \partial_x \left( \int_{\mathbb{R}} \psi_{\infty} \partial_x \left( \frac{\psi}{\psi_{\infty}} \right) dy \right) + \partial_x ((A' - A'_t) \psi^{\xi,I}) - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}). \quad (41)$$

By developing the right hand side, we have

$$\begin{aligned} \partial_t \psi^{\xi,I} &= \partial_x \left( \int_{\mathbb{R}} \partial_x \psi - \frac{\psi}{\psi_{\infty}} \partial_x \psi_{\infty} \right) + \partial_x \int_{\mathbb{R}} (A' - A'_t) \psi - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}) \\ &= \partial_x \left( \int_{\mathbb{R}} \partial_x (V_i - A \circ \xi) \psi \right) + \partial_{xx} \psi^{\xi,I} + \partial_x \int_{\mathbb{R}} (A' - A'_t) \psi - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}) \\ &= \int_{\mathbb{R}} \partial_x (\partial_x (V_i - A_t \circ \xi) \psi) + \partial_{xx} \psi^{\xi,I} - \lambda(x) (\psi^{\xi,I} - \psi_{1-i}^{\xi,I}) \end{aligned}$$

which is indeed the Fokker-Planck equation (15) associated to  $\psi^{\xi,I}$ .

Next, we show that  $\forall x \in \mathbb{T}, \forall i \in \{0, 1\}$

$$\int_{\mathbb{R}} \psi_{\infty} \partial_x \left( \frac{\psi}{\psi_{\infty}} \right) dy = \psi_{\infty}^{\xi, I} \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) + \left( \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right) \psi^{\xi, I}. \quad (42)$$

To prove the above, notice that

$$\begin{aligned} \int_{\mathbb{R}} \psi_{\infty} \partial_x \left( \frac{\psi}{\psi_{\infty}} \right) dy &= \psi_{\infty}^{\xi, I} \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \\ &= \int_{\mathbb{R}} \partial_x \psi dy - \int_{\mathbb{R}} \frac{\psi}{\psi_{\infty}} \partial_x \psi_{\infty} - \partial_x \psi^{\xi, I} + \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \partial_x \psi_{\infty}^{\xi, I} \\ &= \int_{\mathbb{R}} (\partial_x (V_i - A \circ \xi)) \psi dy - \frac{\int_{\mathbb{R}} \partial_x (V_i - A \circ \xi) e^{-(V_i - A \circ \xi)} dy}{\int_{\mathbb{R}} e^{-(V_i - A \circ \xi)} dy} \psi^{\xi, I} \\ &= \int_{\mathbb{R}} (\partial_x V_i) \psi dy - A'(x) \psi^{\xi, I} - \frac{\int_{\mathbb{R}} \partial_x V_i e^{-(V_i - A \circ \xi)} dy}{\int_{\mathbb{R}} e^{-(V_i - A \circ \xi)} dy} \psi^{\xi, I} + A'(x) \psi^{\xi, I} \end{aligned}$$

Finally, by using the fact that the free energy  $A$  is independent of  $y$ , we obtain

$$\int_{\mathbb{R}} \psi_{\infty} \partial_x (\psi / \psi_{\infty}) dy - \psi_{\infty}^{\xi, I} \partial_x (\psi^{\xi, I} / \psi_{\infty}^{\xi, I}) = \left( \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right) \psi^{\xi, I}, \quad (43)$$

as required. The final result (40) is obtained by substituting (42) into (41).  $\diamond$

Notice that (40) is comparable to (25), only with additional terms due to the fact that  $A_t$  and  $\int_{\mathbb{T}} \partial_x V d\mu_{t|x, i}$  have not yet converged. The difference of the biasing force and mean force,  $A'_t - A'$ , was already estimated in Lemmas 5 and 6. We are therefore left to control the remaining term.

**Lemma 11**  $\forall x \in \mathbb{T}, \forall i \in \{0, 1\}$ ,

$$\left| \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right| \leq M \sqrt{\frac{2}{\rho}} H(\mu_{t|x, i} | \mu_{\infty|x, i}).$$

As a consequence,

$$\sum_{i=0}^1 \int_{\mathbb{T}} \left| \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right|^2 \psi^{\xi, I} dx \leq \frac{2M^2}{\rho} E_m. \quad (44)$$

*Proof* : Let  $\Pi(\mu_{t|x, i}, \mu_{\infty|x, i})$  be the set of coupling measures on  $\mathbb{R} \times \mathbb{R}$  with marginals  $\mu_{t|x, i}$  and  $\mu_{\infty|x, i}$  respectively and let  $\pi \in \Pi$ . Then

$$\begin{aligned} \left| \int_{\mathbb{R}} \left( \partial_x V_i \frac{\psi}{\psi^{\xi, I}} - \partial_x V_i \frac{\psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right) \right| &= \left| \int_{\mathbb{R} \times \mathbb{R}} (\partial_x V_i(x, y) - \partial_x V_i(x, y')) \pi(dy, dy') \right| \\ &\leq \|\partial_{x, y} V_i\|_{L^\infty} \int_{\mathbb{R} \times \mathbb{R}} |y - y'| \pi(dy, dy') \\ &\leq M \sqrt{\frac{2}{\rho}} H(\mu_{t|x, i} | \mu_{\infty|x, i}) \end{aligned}$$

where we have used Lemma 1 since  $\mu_{\infty|x, i}$  satisfies LSI( $\rho$ ). Equation (44) follows immediately from the fact that  $\sum_{i=0}^1 \int_{\mathbb{T}} H(\mu_{t|x, i} | \mu_{\infty|x, i}) \psi^{\xi, I} dx \leq E_m$ .  $\diamond$

One result that is now needed to derive estimates on the evolution of  $P$  is the existence of a spectral gap of the operator describing the dynamics once  $A'_t$  and  $\int_{\mathbb{R}} \partial_x V_i d\mu_{t|x,i}$  have converged (see (25)). We now justify the existence of such a spectral gap.

In order to do so, let us define the vector spaces

$$\mathcal{V}_l = \left\{ v : \mathbb{T} \times \{0, 1\} \rightarrow \mathbb{R} \mid \forall i \in \{0, 1\}, \frac{v_i}{\psi_{\infty}^{\xi, I}(x, i)} \in L^2(\mathbb{T}, \psi_{\infty}^{\xi, I}(x, i) dx), \sum_{i=0}^1 \int_{\mathbb{T}} v_i(x) dx = l \right\}$$

and

$$\mathcal{W}_l = \left\{ w \in \mathcal{V}_l \mid \forall i \in \{0, 1\}, \frac{w_i}{\psi_{\infty}^{\xi, I}(x, i)} \in H^1(\mathbb{T}, \psi_{\infty}^{\xi, I}(x, i) dx), \sum_{i=0}^1 \int_{\mathbb{T}} w_i(x) dx = l \right\}.$$

A function  $\phi$  in  $\mathcal{V}_l$  (or in  $\mathcal{W}_l$ ) will also be considered as a vector valued function as  $\phi : \begin{cases} \mathbb{T} \rightarrow \mathbb{R}^2 \\ x \mapsto (\phi_0(x), \phi_1(x)) \end{cases}$ .

Notice that  $\phi \in \mathcal{W}_1$  if and only if  $f := \phi - \psi_{\infty}^{\xi, I} \in \mathcal{W}_0$ .

**Lemma 12** Recall the operator  $\mathcal{L} = (\mathcal{L}_0, \mathcal{L}_1)$ , with  $\mathcal{L}_i$  defined as in (25),

$$\mathcal{L}_i \phi = - \left[ \partial_x \left( \psi_{\infty, i}^{\xi, I} \partial_x \left( \phi_i / \psi_{\infty, i}^{\xi, I} \right) \right) - \lambda(x) (\phi_i - \phi_{1-i}) \right].$$

Then

i) The operator  $\mathcal{L}$  is symmetric and positive definite with respect to the inner product

$$\langle f, g \rangle = \sum_{i=0}^1 \int_{\mathbb{T}} f_i(x) g_i(x) \frac{1}{\psi_{\infty}^{\xi, I}(x, i)} dx.$$

ii)  $\mathcal{L}$  has a spectral gap  $\theta > 0$  in the sense that

$$\inf_{f \in \mathcal{W}_0, f \neq 0} \frac{\langle f, \mathcal{L}f \rangle}{\langle f, f \rangle} = \theta > 0. \quad (45)$$

*Proof :* i) To show symmetry of the operator  $\mathcal{L}$ , consider functions  $\varphi, \phi \in \mathcal{W}_0$ . Now, using the fact that  $\forall x \in \mathbb{T} \setminus \mathcal{E}, \psi_{\infty, 0}^{\xi, I}(x) = \psi_{\infty, 1}^{\xi, I}(x)$ ,

$$\begin{aligned} & \sum_{i=0}^1 \int_{\mathbb{T}} \varphi_i \mathcal{L}_i \phi \frac{1}{\psi_{\infty, i}^{\xi, I}} dx \\ &= - \sum_{i=0}^1 \int_{\mathbb{T}} \varphi_i \partial_x \left( \psi_{\infty, i}^{\xi, I} \partial_x \left( \frac{\phi_i}{\psi_{\infty, i}^{\xi, I}} \right) \right) \frac{1}{\psi_{\infty, i}^{\xi, I}} dx + \sum_{i=0}^1 \int_{\mathbb{T}} \lambda(x) \varphi_i (\phi_i - \phi_{1-i}) \frac{1}{\psi_{\infty, i}^{\xi, I}} dx \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \partial_x \left( \frac{\varphi_i}{\psi_{\infty, i}^{\xi, I}} \right) \partial_x \left( \frac{\phi_i}{\psi_{\infty, i}^{\xi, I}} \right) \psi_{\infty, i}^{\xi, I} dx + \int_{\mathbb{T}} \lambda(x) (\varphi_0 - \varphi_1) (\phi_0 - \phi_1) \frac{1}{\psi_{\infty, 0}^{\xi, I}} dx \quad (46) \\ &= - \sum_{i=0}^1 \int_{\mathbb{T}} \phi_i \partial_x \left( \psi_{\infty, i}^{\xi, I} \partial_x \left( \frac{\varphi_i}{\psi_{\infty, i}^{\xi, I}} \right) \right) \frac{1}{\psi_{\infty, i}^{\xi, I}} dx + \int_{\mathbb{T}} \lambda(x) (\phi_0 - \phi_1) (\varphi_0 - \varphi_1) \frac{1}{\psi_{\infty, 0}^{\xi, I}} dx \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \phi_i \mathcal{L}_i \varphi \frac{1}{\psi_{\infty, i}^{\xi, I}} dx. \end{aligned}$$

From (46), we conclude positive definiteness of  $\mathcal{L}$

$$\sum_{i=0}^1 \int_{\mathbb{T}} \phi_i \mathcal{L}_i \phi \frac{1}{\psi_{\infty,i}^{\xi,I}} dx = \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\phi_i}{\psi_{\infty,i}^{\xi,I}} \right) \right|^2 \psi_{\infty,i}^{\xi,I} dx + \int_{\mathbb{T}} \lambda(x) (\phi_0 - \phi_1)^2 \frac{1}{\psi_{\infty,0}^{\xi,I}} dx > 0. \quad (47)$$

Notice that the above is strictly positive for any  $\phi \in \mathcal{W}_0$ , since  $\langle \phi, \mathcal{L}\phi \rangle = 0$  if and only if  $\phi_0 = \phi_1 = 0$ .

ii) In fact, one can check that  $\exists \kappa > 0$ , such that  $\forall \phi \in \mathcal{W}_0$ ,  $\phi \neq 0$ ,

$$\sum_{i=0}^1 \int_{\mathbb{T}} \phi_i \mathcal{L}_i \phi \frac{1}{\psi_{\infty}^{\xi,I}} dx \geq \kappa \sum_{i=0}^1 \int_{\mathbb{T}} (|\phi_i|^2 + |\nabla \phi_i|^2) \frac{1}{\psi_{\infty}^{\xi,I}} dx.$$

Therefore, by the Lax-Milgram theorem,  $\mathcal{L}^{-1}$  is well defined from  $\mathcal{V}_0$  to  $\mathcal{W}_0$  and thus compact from  $\mathcal{V}_0$  to  $\mathcal{V}_0$ . From the symmetry and positive definiteness of  $\mathcal{L}$ , and the fact that its inverse is a compact operator from  $\mathcal{V}_0$  to  $\mathcal{V}_0$ , it has a strictly positive and discrete spectrum. There exists a set of eigenvectors  $(v_n)_{n \geq 1}$ , orthonormal with respect to the inner product  $\langle \cdot, \cdot \rangle$ , forming a basis of  $\mathcal{V}_0$  and  $\mathcal{W}_0$ , and associated to an increasing sequence of eigenvalues  $(\sigma_n)_{n \geq 1}$ , such that  $\lim_{n \rightarrow \infty} \sigma_n = \infty$ . In particular, there exists a spectral gap:  $\theta = \sigma_1 > 0$ .  $\diamond$

**Remark 4** In the case where a function  $\phi \in \mathcal{W}_1$  satisfies  $\partial_t \phi = \mathcal{L}\phi$ , a consequence of Lemma 12 is

$$\forall t \geq 0, \quad \|\phi(t, \cdot) - \psi_{\infty}^{\xi,I}\|^2 \leq K e^{-2\theta t}, \quad (48)$$

where  $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$  and  $K = \sum_{n \geq 1} \langle \phi(0, \cdot) - \psi_{\infty}^{\xi,I}, v_n \rangle^2$ . This is easily obtained by noticing that  $\phi - \psi_{\infty}^{\xi,I} \in \mathcal{W}_0$  and therefore can be expressed in terms of the orthonormal eigenvectors  $(v_n)_{n \geq 1}$

$$\phi(t, \cdot) - \psi_{\infty}^{\xi,I} = \sum_{n \geq 1} \langle \phi(0, \cdot) - \psi_{\infty}^{\xi,I}, v_n \rangle v_n e^{-\sigma_n t}.$$

The result (48) follows immediately since  $\forall n \geq 2$ ,  $\sigma_n \geq \sigma_1 = \theta$ .

Notice in particular that the Fokker-Planck equation (40) satisfied by  $\psi^{\xi,I}$  can be written as

$$\partial_t \psi^{\xi,I}(t, x, i) = \mathcal{L}_i \psi^{\xi,I} + \partial_x \left( \left( \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi,I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi,I}} \right) \psi^{\xi,I} \right) + \partial_x ((A' - A'_t) \psi^{\xi,I}).$$

We will show that the last two terms can be controlled by  $E_m$ , see (49).

With these tools at hand, let us consider the time evolution of the functional  $P$  defined in (39)

$$\begin{aligned} \frac{1}{2} \frac{dP}{dt} &= \sum_{i=0}^1 \int_{\mathbb{T}} \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} - 1 \right) \partial_t \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} \right) \psi_{\infty}^{\xi,I} dx \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \psi^{\xi,I} \partial_t \left( \frac{\psi^{\xi,I}}{\psi_{\infty}^{\xi,I}} \right) dx - \sum_{i=0}^1 \int_{\mathbb{T}} \partial_t \psi^{\xi,I} dx \\ &= \sum_{i=0}^1 \int_{\mathbb{T}} \psi^{\xi,I} \partial_t \psi^{\xi,I} \frac{1}{\psi_{\infty}^{\xi,I}} dx. \end{aligned}$$

Using equation (40) of Lemma 10, we get

$$\begin{aligned}
\frac{1}{2} \frac{dP}{dt} &= \sum_{i=0}^1 \int_{\mathbb{T}} \psi^{\xi, I} \partial_t \psi^{\xi, I} \frac{1}{\psi_{\infty}^{\xi, I}} dx \\
&= - \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx - \sum_{i=0}^1 \int_{\mathbb{T}} \lambda(x) \psi_i^{\xi, I} (\psi_i^{\xi, I} - \psi_{1-i}^{\xi, I}) \frac{1}{\psi_{\infty, i}^{\xi, I}} dx \\
&\quad - \sum_{i=0}^1 \int_{\mathbb{T}} \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) (A' - A'_t) \psi^{\xi, I} dx - \sum_{i=0}^1 \int_{\mathbb{T}} \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \left( \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right) \psi^{\xi, I} dx.
\end{aligned}$$

Notice that, by developing the sum and using the fact that  $\psi_{\infty, 0}^{\xi, I} = \psi_{\infty, 1}^{\xi, I}$  for  $\lambda(x) \neq 0$ , the second term may be replaced by  $\int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty, 0}^{\xi, I}} dx$ . Finally by using Young's inequality on the last two terms, we obtain for a parameter  $\alpha > 0$  to be chosen later on,

$$\begin{aligned}
\frac{1}{2} \frac{dP}{dt} &= - \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx - \int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty, 0}^{\xi, I}} dx \\
&\quad + \frac{1}{4\alpha} \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi^{\xi, I} dx + \alpha \int_{\mathbb{T}} |A' - A'_t|^2 \psi^{\xi} dx \\
&\quad + \frac{1}{4\alpha} \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi^{\xi, I} dx + \alpha \sum_{i=0}^1 \int_{\mathbb{T}} \left| \frac{\int_{\mathbb{R}} \partial_x V_i \psi}{\psi^{\xi, I}} - \frac{\int_{\mathbb{R}} \partial_x V_i \psi_{\infty}}{\psi_{\infty}^{\xi, I}} \right|^2 \psi^{\xi, I} dx.
\end{aligned}$$

Next, by Lemmas 6 and 11,

$$\begin{aligned}
\frac{1}{2} \frac{dP}{dt} &= - \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx - \int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty}^{\xi, I}} dx \\
&\quad + \frac{1}{2\alpha} \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi^{\xi, I} dx + 2\alpha \frac{M^2}{\rho} E_m + 2\alpha R^2 E_m.
\end{aligned}$$

Notice that, in the third term,  $\psi^{\xi, I} \leq \psi^{\xi} \leq \tilde{M}$  for  $\tilde{M} = \|\psi^{\xi}(0, \cdot)\|_{L^{\infty}}$  and  $1 \leq \psi_{\infty}^{\xi, I}/c$  for  $c = \min_{x, i} \psi_{\infty}^{\xi, I}$ . This gives

$$\begin{aligned}
\frac{1}{2} \frac{dP}{dt} &\leq - \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx - \int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty}^{\xi, I}} dx \\
&\quad + \frac{\tilde{M}}{2\alpha c} \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx + 4\alpha R^2 E_m.
\end{aligned}$$

Finally, by grouping terms together and using the fact that  $\alpha$  may be chosen such that  $\tilde{M}/2\alpha c < 1$

(an appropriate choice for  $\alpha$  is given later in the proof), we have

$$\begin{aligned}
\frac{1}{2} \frac{dP}{dt} &\leq - \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx - \int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty}^{\xi, I}} dx + 4\alpha R^2 E_m \\
&\leq - \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \left[ \sum_{i=0}^1 \int_{\mathbb{T}} \left| \partial_x \left( \frac{\psi^{\xi, I}}{\psi_{\infty}^{\xi, I}} \right) \right|^2 \psi_{\infty}^{\xi, I} dx + \int_{\mathbb{T}} \lambda(x) \left| \psi_0^{\xi, I} - \psi_1^{\xi, I} \right|^2 \frac{1}{\psi_{\infty}^{\xi, I}} dx \right] + 4\alpha R^2 E_m \\
&\leq - \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \sum_{i=0}^1 \int_{\mathbb{T}} (\psi^{\xi, I} \mathcal{L}_i \psi^{\xi, I}) \frac{1}{\psi_{\infty}^{\xi, I}} dx + 4\alpha R^2 E_m \\
&\leq - \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \theta P + 4\alpha R^2 E_m,
\end{aligned} \tag{49}$$

where the last line is a result of (45), with  $f := \psi^{\xi, I} - \psi_{\infty}^{\xi, I}$ . Notice that  $f \in \mathcal{W}_0$  since the normalization for  $\psi^{\xi, I}$  is  $\sum_{i=0}^1 \int_{\mathbb{T}} \psi^{\xi, I} dx = 1$ .

To complete the proof of Theorem 1, we now need to study the system of inequalities (36) and (49).

### 3.5 Completing the proof

To show that  $E_m$  decays exponentially fast, we study the system of two inequalities (36) and (49). Since, from (38),  $E_c \leq P$ , the system to be studied is

$$\begin{cases} \frac{dE_m}{dt} \leq -2(\rho - R^2\varepsilon) E_m + 2\rho P + \frac{1}{4\varepsilon} F_0 e^{-8\pi^2 t}, \\ \frac{dP}{dt} \leq 8\alpha R^2 E_m - 2 \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \theta P. \end{cases} \tag{50}$$

The parameters  $\alpha > \frac{\tilde{M}}{2c}$  and  $\varepsilon > 0$  remain to be chosen in order to obtain an exponential convergence with the best possible rate. To fix  $\alpha$ , let us first study the eigenvalues of the matrix of coefficients, neglecting the terms in  $\varepsilon$ .

**Lemma 13** *Let us assume [H4]. The matrix*

$$A = \begin{pmatrix} -\rho & \rho \\ 4\alpha R^2 & -\left(1 - \frac{\tilde{M}}{2\alpha c}\right) \theta \end{pmatrix}$$

*is negative definite and  $\alpha$  may be chosen so that the eigenvalues  $-\lambda_{\pm}$  of  $A$  are such that*

$$-\lambda_- \leq -\lambda_+ = -\Lambda(\theta) < 0$$

*where  $\Lambda : (\theta_{\min}, \infty) \rightarrow (0, \rho)$  is a positive, increasing function. We recall that  $\theta_{\min} = \frac{8\tilde{M}R^2}{c}$ , where  $R = C + M\rho^{-1/2}$ . The function  $\Lambda$  is such that  $\Lambda \rightarrow 0$  as  $\theta \rightarrow \theta_{\min}$  and  $\Lambda \rightarrow \rho$  as  $\theta \rightarrow \infty$ . Moreover,  $\Lambda(\rho + 2\theta_{\min}) = \frac{\rho}{2}$ .*

*Proof :* In order to prove the negative definiteness of the matrix  $A$ , we show that for certain values of  $\alpha > 0$ ,  $\text{tr}(A) < 0$  and  $\det(A) > 0$ . In the following, we only consider positive values of  $\alpha$  (which is imposed by the previous computations). We have

$$\text{tr}(A) = -\rho - \left(1 - \frac{\tilde{M}}{2\alpha c}\right) \theta < 0 \quad \text{iff} \quad \alpha > \frac{\tilde{M}\theta}{2c(\rho + \theta)} \tag{51}$$

and

$$\det(A) = \theta\rho \left(1 - \frac{\tilde{M}}{2\alpha c}\right) - 4\alpha R^2\rho > 0 \text{ iff } \alpha \in (\alpha_-, \alpha_+), \alpha_{\pm} = \frac{\theta c \pm \sqrt{\theta^2 c^2 - 8\tilde{M}\theta R^2 c}}{8R^2 c}. \quad (52)$$

The interval  $(\alpha_-, \alpha_+)$  is indeed well defined and included in  $[0, \infty)$  since  $\theta > \theta_{\min} = 8\tilde{M}R^2/c$  (hypothesis [H4]). We seek an optimal  $\alpha$  that minimizes eigenvalue  $-\lambda_+$  and satisfies (51) and (52). An analytical solution cannot be easily obtained. We choose

$$\alpha = \alpha^* := \frac{\tilde{M}}{c},$$

which appears to be very close to the optimal choice, from numerical computations. Notice that  $\alpha^*$  satisfies (51) and (52) since, for  $\alpha = \alpha^*$ ,  $\text{tr}(A) = -\rho - \theta/2 < 0$  and  $\det(A) = \theta\rho/2 - 4R^2\tilde{M}\rho/c > 0$ . The eigenvalues of the matrix are now given by

$$-\lambda_{\pm} = \frac{1}{2} \left( -\left(\rho + \frac{\theta}{2}\right) \pm \sqrt{\left(\rho - \frac{\theta}{2}\right)^2 + \frac{16R^2\tilde{M}\rho}{c}} \right) < 0.$$

The rate of convergence of the system is given by the largest of the two eigenvalues  $-\lambda_+$ . Let us introduce the function

$$\Lambda(\theta) = -\frac{1}{2} \left( -\left(\rho + \frac{\theta}{2}\right) + \sqrt{\left(\rho - \frac{\theta}{2}\right)^2 + \frac{16R^2\tilde{M}\rho}{c}} \right)$$

such that  $\lambda_+ = \Lambda(\theta)$ . It is easily shown that  $\Lambda$  is an increasing function of  $\theta$  with

$$\Lambda(\theta) \rightarrow \begin{cases} 0 & \text{as } \theta \rightarrow \frac{8\tilde{M}R^2}{c}, \\ \rho & \text{as } \theta \rightarrow \infty. \end{cases}$$

Moreover, it is easy to check that  $\Lambda(\rho + 2\theta_{\min}) = \frac{\rho}{2}$ , which concludes the proof.  $\diamond$

We are now in position to complete the proof of Theorem 1. Let us define  $Y(t) = (E_m(t), P(t))$ . Using (50) the fact that  $E_m \leq \|Y\|_2$  (where  $\|Y\|_2$  denotes the Euclidean norm of the two-dimensional vector  $Y$ ), we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|Y\|_2^2 &= \frac{1}{2} \frac{d}{dt} (E_m^2 + P^2) \\ &\leq 2Y^T AY + 2R^2\varepsilon E_m^2 + \frac{1}{4\varepsilon} F_0 e^{-8\pi^2 t} E_m \\ &\leq -2\Lambda(\theta) \|Y\|_2^2 + 2R^2\varepsilon \|Y\|_2^2 + \frac{1}{4\varepsilon} F_0 e^{-8\pi^2 t} \|Y\|_2, \end{aligned}$$

and as a result,

$$\frac{d\|Y\|_2}{dt} \leq -2(\Lambda(\theta) - R^2\varepsilon) \|Y\|_2 + \frac{1}{4\varepsilon} F_0 e^{-8\pi^2 t}. \quad (53)$$

For arbitrary small  $\varepsilon > 0$ , let us consider  $\lambda_\varepsilon = \Lambda(\theta) - R^2\varepsilon < \Lambda(\theta)$ . We may assume without loss of generality that  $\lambda_\varepsilon \neq 4\pi^2$ . Then, from (53), one gets:

$$E_m \leq \|Y\|_2 \leq K_\varepsilon e^{-2\min\{\lambda_\varepsilon, 4\pi^2\}t}, \quad (54)$$

where

$$K_\varepsilon = 2 \max \left\{ \sqrt{E_m^2(0) + P^2(0)}, \frac{F_0}{8\varepsilon|\lambda_\varepsilon - 4\pi^2|} \right\},$$

which concludes the proof of (27).

The exponential convergence of the total entropy  $E$  results from the relation  $E = E_M + E_m$ , (54) and Lemmas 7 and 8. The Csiszar-Kullback inequality implies the same for  $\|\psi(t, \cdot) - \psi_\infty\|_{L^1}^2$ .

Finally, the convergence results on  $A'_t$  are easily obtained from Lemma 6 and the fact that  $\psi^\zeta$  is bounded from below by a positive constant for times larger than an arbitrary small positive time, see the beginning of Section 3.3.2 in [14] for more details.

## Acknowledgements

This work is supported in part by the Agence Nationale de la Recherche, under the grant MEGAS (ANR-09-BLAN-0216-01). The authors are grateful to Benjamin Jourdain, Eric Cancès, Andreas Eberle, Felix Otto and Gabriel Stoltz for helpful discussions.

## References

- [1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. SMF, 2000.
- [2] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On logarithmic Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Differential Equations*, 26:35–43, 2001.
- [3] D. Bakry and M. Emery. Hypercontractivité de semi-groupes de diffusion. *C. R. Acad. Sci Paris Sér. I*, 299:775–778, 1984.
- [4] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. *Comm. Pure Appl. Math.*, 61:3, 2008.
- [5] E. Darve and A. Pohorille. Calculating free energies using average forces. *J. Chem. Phys.*, 115(20):9196–9183, 2001.
- [6] W.K. den Otter and W.J. Briels. The calculation of free energy differences by constrained molecular dynamics simulations. *J. Chem. Phys.*, 109:4139, 1998.
- [7] L. Gross. Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97:1061–1083, 1975.
- [8] J. Hénin and C. Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.*, 121:2904, 2004.
- [9] J. Hénin, G. Fiorin, C. Chipot, and M. L. Klein. Exploring multidimensional free energy landscapes using time-dependent biases on collective variables. *J. Chem. Theory Comput.*, 6:35–47, 2010.
- [10] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46:1159–1194, 1987.
- [11] B. Jourdain, T. Lelièvre, and R. Roux. Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process. *to appear in ESAIM: M2AN*, 2010.
- [12] T. Lelièvre. A general two-scale criteria for logarithmic Sobolev inequalities. *J. Func. Anal.*, 256(7):2211–2221, 2009.
- [13] T. Lelièvre, M. Rousset, and G. Stoltz. Computation of free energy profiles with parallel adaptive dynamics. *J. Chem. Phys.*, 126:134111, 2007.

- [14] T. Lelièvre, M. Rousset, and G. Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21(6):1155–1181, 2008.
- [15] K. Minoukadeh, C. Chipot, and T. Lelièvre. Potential of mean force calculations: A multiple-walker Adaptive Biasing Force approach. *J. Chem. Theory Comput.*, 6:1008–1017, 2010.
- [16] F. Otto and C. Villani. Generalization of an inequality by Talagrand, viewed as a consequence of the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.
- [17] M. Sprik and G. Cicotti. Free energy from constrained molecular dynamics. *J. Chem. Phys.*, 109:7737–7744, 1998.
- [18] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.