

## MEAN-FIELD LIMIT VERSUS SMALL-NOISE LIMIT FOR SOME INTERACTING PARTICLE SYSTEMS

SAMUEL HERRMANN AND JULIAN TUGAUT

**ABSTRACT.** In the nonlinear diffusion framework, stochastic processes of McKean-Vlasov type play an important role. Such diffusions can be obtained by taking the hydrodynamic limit in a huge system of linear diffusions in interaction. In both cases, for the linear and the nonlinear processes, small-noise asymptotics have been emphasized by specific large deviation phenomena. The natural question, therefore, is: is it possible to interchange the mean-field limit with the small-noise limit? The aim here is to consider this question by proving that the rate function of the first particle in a mean-field system converges to the rate function of the hydrodynamic limit as the number of particles becomes large.

### 1. Introduction

In the stochastic convergence framework, the large deviation theory plays an essential role for describing the rate at which the probability of certain rare events decays. Each convergence result therefore leads to find the large deviation rate associated with. In suitable cases, the knowledge of the so-called *large deviation principle (LDP)* even permits to obtain information about the convergence itself (see the central limit theorem [3]).

This paper is concerned with the convergence of continuous stochastic processes defined as small random perturbations of dynamical systems. In the classical diffusion case, the stochastic process converges in the small-noise limit to the deterministic solution of the dynamical system and the large deviation theory developed by Freidlin and Wentzell [11] emphasizes the behaviour of the rare event probabilities. More recently, Herrmann, Imkeller and Peithmann [14] studied the large deviation phenomenon associated with the McKean-Vlasov process, a particular nonlinear diffusion which is attracted by its own law (the so-called self-stabilizing effect). This process appears for instance in the probabilistic interpretation of the granular media equation. They presented the explicit expression of the rate function  $J_\infty$  and the Kramers' rate which is related to the time needed by the diffusion to exit a given bounded domain.

The aim of this paper is to better understand the link between the large deviation principle of the nonlinear diffusion and the classical theory developed by Freidlin and Wentzell. More precisely, the McKean-Vlasov equation describes the

---

Received 2016-1-22; Communicated by the editors.

2010 *Mathematics Subject Classification.* Primary 46N30, 60H10 ; Secondary 60F10, 60J60.

*Key words and phrases.* Inversion of limits, large deviations, interacting particle systems, hydrodynamic limit, nonlinear diffusions, small-noise limit.

behaviour of one particle in a huge system of particles in interaction, as a result of the hydrodynamic limit in a mean-field system. The natural question, therefore, is to emphasize the link between the rate function (or entropy function)  $J_\infty$  of the nonlinear diffusion and the Freidlin-Wentzell rate function  $J_N$  associated to one particle in a mean-field system of size  $N$ . Our approach is essentially based on functional analysis tools.

The material is organized as follows: first we discuss and recall different notions associated with the large deviation theory. Secondly we present the model and point out the link between nonlinear diffusion and high dimensional classical diffusions: the so-called mean-field effect. The third section will be devoted to the main result: the convergence of the rate functions  $J_N \rightarrow J_\infty$  as  $N$  becomes large. Finally, we present some immediate consequences and a generalization result.

**1.1. A large deviation principle.** In this paper, the framework concerns continuous time processes depending on a parameter  $\sigma$  and we describe the behaviour of this family in the small-parameter limit. We need to find out a rate function (entropy) which describes the probability of a trajectory to be far away from typical paths.

Let us consider a family of continuous stochastic processes  $X^\sigma := (X_t^\sigma)_{t \in [0, T]}$  with  $T < \infty$ . In the following, the family of processes  $(X^\sigma)_{\sigma > 0}$  is said to satisfy a *large deviation principle with speed  $\frac{\sigma^2}{2}$*  if there exists a lower semi-continuous mapping (called *rate function*)  $I$  from  $\mathcal{C}([0, T]; \mathbb{R}^d)$  to  $\mathbb{R}_+$  such that

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log [\mathbb{P} \{X^\sigma \in \mathbb{F}\}] \leq - \inf_{\varphi \in \mathbb{F}} I(\varphi)$$

for any closed subset  $\mathbb{F} \subset \mathcal{C}([0, T]; \mathbb{R}^d)$  equipped with the uniform topology and

$$\liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log [\mathbb{P} \{X^\sigma \in \mathbb{G}\}] \geq - \inf_{\varphi \in \mathbb{G}} I(\varphi)$$

for any open subset  $\mathbb{G} \subset \mathcal{C}([0, T]; \mathbb{R}^d)$ .  $I$  is a *good* rate function if its level sets are compact subsets of  $\mathcal{C}([0, T]; \mathbb{R}^d)$ .

We now focus our attention to stochastic differential equations driven by a Brownian motion. Schilder's theorem deals with the LDP of  $X^\sigma := (\sigma B_t)_{t \in [0, T]}$ , where  $B$  is a standard  $d$ -dimensional Wiener process (see Theorem 5.2.3 in [7]). The associated good rate function is given by

$$I_0(\varphi) := \frac{1}{4} \int_0^T \|\dot{\varphi}(t)\|^2 dt,$$

if  $\varphi$  belongs to the set of absolutely continuous functions starting in 0, denoted by  $\mathcal{H}_0$ . If  $\varphi \notin \mathcal{H}_0$ , we set  $I_0(\varphi) := +\infty$ . Here  $\|\cdot\|$  stands for the euclidean norm in  $\mathbb{R}^d$ . The study elaborated by Schilder permits us to go further in the description of LDP for diffusions as presented by Freidlin and Wentzell. If  $X^\sigma$  satisfies the stochastic differential equation:

$$X_t^\sigma = x + \sigma B_t - \int_0^t b(s, X_s^\sigma) ds,$$

where the drift term  $b(t, x)$  is a continuous function with respect to the time variable and locally Lipschitz with respect to the space variable; then the family  $(X^\sigma)_{\sigma>0}$  admits a LDP with the good rate function

$$I_b(\varphi) := \frac{1}{4} \int_0^T \|\dot{\varphi}(t) + b(t, \varphi(t))\|^2 dt$$

for  $\varphi \in \mathcal{H}_x$  (the set of absolutely continuous functions starting in  $x$ ). For  $\varphi \notin \mathcal{H}_x$ ,  $I_b(\varphi) := +\infty$ . Let us focus our attention to the *typical paths* of such a diffusion. In fact, in the particular case of a deterministic equation

$$\Psi_t(x) = x - \int_0^t b(s, \Psi_s(x)) ds \quad (1.1)$$

admitting a unique solution, the diffusion  $X^\sigma$  starting in  $x$  converges in probability towards the deterministic trajectory  $\Psi(x)$  in the small-noise limit. The Freidlin-Wentzell LDP estimates the rate of convergence: introducing

$$\mathbb{F} := \{\varphi \in \mathcal{C}([0, T]; \mathbb{R}^d) : \|\varphi - \Psi(x)\|_\infty \geq \delta\},$$

where  $\|\cdot\|_\infty$  stands for the uniform norm, we obtain

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(X^\sigma \in \mathbb{F}) \leq - \inf_{\varphi \in \mathbb{F}} I_b(\varphi) < 0.$$

Let us finally note that the precise description of the deviation phenomenon permits us to deal with the small-noise asymptotics of exit times  $\tau_{\mathcal{D}}$  from a domain of attraction  $\mathcal{D}$ . Namely if the drift term of the diffusion is in the so-called gradient case, that is  $b(t, x) = \nabla V(x)$ , if moreover  $V$  reaches a local minimum for  $x = a$  and  $\mathcal{D}$  is a bounded domain of attraction associated to  $a$ , then a Kramers' type law can be observed. A weak version of this result is the following asymptotic expression:

$$\lim_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{E}[\tau_{\mathcal{D}}] = \inf_{T>0} \inf_{\substack{\varphi(0)=a, \\ \varphi(T) \in \partial \mathcal{D}}} I_{\nabla V}(\varphi) = \inf_{y \in \partial \mathcal{D}} V(y) - V(a)$$

In other words, not only is the rate function a key tool for the description of the diffusion deviation from typical trajectories (linked to a study on a fixed time interval  $[0, T]$ ), but it is also involved in the description of exit times from a domain (a study developed on the whole time axis).

The aim of our paper, therefore, is to describe some nice properties of the rate function, not in the classical diffusion case just described above, but for self-stabilizing diffusions of the McKean-Vlasov type, diffusions attracted by their own law. Let us finally note that for other applications of large deviations to communication, optic and biology, we refer the reader to [7].

**1.2. The self-stabilizing model.** From now on, we restrict the study to the McKean-Vlasov model: for  $x \in \mathbb{R}^d$ , the process satisfies the following stochastic differential equation:

$$\begin{cases} X_t^\sigma = x + \sigma B_t - \int_0^t \nabla W_s^\sigma(X_s^\sigma) ds, \\ W_t^\sigma := V + F * u_t^\sigma. \end{cases} \quad (1.2)$$

The  $*$  symbol stands for the convolution product and  $u_t^\sigma$  denotes the density of the probability distribution  $\mathbb{P}_{X_t^\sigma}$ . Since the law of the solution plays an important role in the structure of the drift term, this equation is nonlinear, in the sense of McKean, see for instance [21, 20]. Three terms contribute to the infinitesimal dynamics.

- The first one is the noise generated by the  $d$ -dimensional Brownian motion  $(B_t, t \geq 0)$ .
- The second force is related to the attraction between any fixed trajectory  $t \mapsto X_t^\sigma(\omega_0)$ ,  $\omega_0 \in \Omega$ , and the entire set of trajectories. Indeed, we observe:

$$\nabla F * u_t^\sigma(X_t^\sigma(\omega_0)) = \int_{\omega \in \Omega} \nabla F(X_t^\sigma(\omega_0) - X_t^\sigma(\omega)) d\mathbb{P}(\omega) .$$

Consequently,  $F$  is called the *interaction potential*. The interaction only depends on the difference  $X_t^\sigma(\omega_0) - X_t^\sigma(\omega)$  and therefore can be associated to the convolution product. Let us note that other dependences have been studied e.g. the quantile case: the drift is then a continuous function of the quantile of the distribution  $\mathbb{P}_{X_t^\sigma}$ , see [18].

- The last term corresponds to the function  $V$ , the so-called *confining potential*. The solution  $X_t^\sigma$  roughly represents the motion of a Brownian particle living in a landscape  $V$  and whose inertia is characterized by  $F$  (if  $F$  is large then the solution moves with difficulty since it is strongly attracted by its law). Therefore it is easy to imagine that the minimizers of the potential  $V$  attract the diffusion if  $F(0) = 0$ . Indeed  $X_t^\sigma$  converges in probability towards one of them as  $\sigma$  tends to 0.

Let us now present the hypotheses concerning the functions  $F$  and  $V$ . The confining potential  $V$  satisfies:

(V1)  $V$  is a  $\mathcal{C}^2$ -continuous function.

(V2)  $\nabla^2 V(x) \geq 0$  for all  $x \notin K$  where  $K$  is a compact subset of  $\mathbb{R}^d$ .

Combining (V1) and (V2) ensures the existence of a solution to (1.2). The interaction function satisfies:

(F1) There exists a function  $G : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $F(x) = G(\|x\|)$ .

(F2)  $G$  is an even polynomial function with  $\deg(G) \geq 2$  and  $G(0) = 0$ .

(F3) The following asymptotic property holds  $\lim_{r \rightarrow +\infty} G(r) = +\infty$ .

Let us now complete the description of this McKean-Vlasov model by briefly recalling several already known results concerning (1.2).

- *Probabilistic interpretation of PDEs.* In fact, the self-stabilizing diffusion corresponds to the probabilistic interpretation of the granular media equation. The probability density function of  $X_t^\sigma$ , starting at  $x$ , is represented by  $(t, x) \mapsto u_t^\sigma(x)$  and satisfies the following partial differential equation

$$\frac{\partial}{\partial t} u_t^\sigma = \operatorname{div} \left\{ \frac{\sigma^2}{2} \nabla_x u_t^\sigma + u_t^\sigma (V + F * u_t^\sigma) \right\} . \quad (1.3)$$

This equation is strongly nonlinear since it contains a *quadratic* term of the form  $u_t^\sigma (F * u_t^\sigma)$ . This link between the granular media equation (1.3)

and the McKean-Vlasov diffusion (1.2) permits us to study the PDE by probabilistic methods [4, 19, 12].

- *Existence and uniqueness.* The existence and the uniqueness of a strong solution  $X^\sigma$  to (1.2) defined on  $\mathbb{R}_+$  has been proven in [14] (Theorem 2.13). Moreover, the long-time asymptotic behaviour of the probability distribution  $\mathbb{P}_{X_t^\sigma}$  has been studied in [4, 2] (for convex functions  $V$ ) and in [25, 26] for the non-convex case. In this second case, the key of the proofs essentially consists in using the results of [15, 16, 17] about the non-uniqueness of the invariant measures (that means in particular that there exist several positive stationary solutions of the granular media equation (1.3) which have a total mass equal to 1).
- *Large deviation principle.* The noise intensity appearing in the equation (1.2) is parametrized by  $\sigma$ . The aim of the large deviation principle is to describe precisely the behaviour of the paths in the small-noise limit. For any  $T > 0$ , we can prove, see [14], that the family of processes  $(X^\sigma)_{\sigma>0}$  satisfies a large deviation principle with the following good rate function  $J_\infty$ :

$$J_\infty(f) := \frac{1}{4} \int_0^T \|\dot{f}(t) + \nabla V(f(t)) + \nabla F(f(t)) - \Psi_\infty^x(t)\|^2 dt, \quad (1.4)$$

if  $f \in \mathcal{H}_x$  and otherwise  $J_\infty(f) := +\infty$ . Here the function  $\Psi_\infty^x$  is independent of  $F$  and satisfies the following ordinary differential equation:

$$\Psi_\infty^x(t) = x - \int_0^t \nabla V(\Psi_\infty^x(s)) ds, \quad x \in \mathbb{R}^d. \quad (1.5)$$

In other words, the diffusion process  $(X_t^\sigma, t \geq 0)$  converges exponentially fast towards the deterministic solution  $\Psi_\infty^x$  as  $\sigma$  tends to 0. The limit function for a classical diffusion  $Y_t^\sigma$  defined by

$$Y_t^\sigma = x + \sigma B_t - \int_0^t \nabla V(Y_s^\sigma) ds$$

is exactly the same: the self-stabilizing phenomenon does not change the limit, it only changes the speed of convergence. Indeed the rate function  $J_\infty$  clearly depends on  $F$ . If the function  $F$  is convex, the trajectories of the McKean-Vlasov diffusion  $X^\sigma$  are closer to  $\Psi_\infty^x$  than the ones of the diffusion  $Y^\sigma$ .

Since the asymptotic behaviour has been described on a fixed-time interval  $[0, T]$ , the next step is to describe the asymptotic behaviour on the whole time interval and namely the study of exit problems: the first time the self-stabilizing diffusion exits from a given bounded domain. This problem has already been solved if both  $V$  and  $F$  are uniformly strictly convex functions, see [14, 24] by the use of large deviation techniques. In [24], the method is based on the exit problem for an associated mean-field system of particles.

**1.3. An interacting particle system.** The McKean-Vlasov diffusion  $X^\sigma$ , described in the previous section, corresponds to the movement of a particle in a

continuous mean-field system in the so-called hydrodynamical limit, that is, as the number of particles tends to infinity. The mean-field system associated to the self-stabilizing process (1.2) is a  $N$  dimensional random dynamical system  $(X^{i,N,\sigma})_{1 \leq i \leq N} \in \otimes^N \mathcal{C}([0, T]; \mathbb{R}^d)$  satisfying

$$dX_t^{i,N,\sigma} = \sigma dB_t^i - \nabla V(X_t^{i,N,\sigma}) dt - \frac{1}{N} \sum_{j=1}^N \nabla F(X_t^{i,N,\sigma} - X_t^{j,N,\sigma}) dt \quad (1.6)$$

and  $X_0^{i,N,\sigma} := x$ . Here,  $(B_t^i)_{t \in \mathbb{R}_+}$  is a family of  $N$  independent  $d$ -dimensional Brownian motions. We also assume  $B^1 = B$ , in other words, both diffusions  $X^{1,N,\sigma}$  and  $X^\sigma$  (see (1.2)) are defined with respect to the same Wiener process (this is possible due to the existence of a strong solution).

The *propagation of chaos* then permits us to link (1.2) and (1.6). It is essentially based on the following intuitive remark. The larger  $N$  is, the less influence a given particle  $X^{j,N,\sigma}$  has on the first particle  $X^{1,N,\sigma}$ . Consequently, it is reasonable to consider that the particles become less and less dependent as the number of particles becomes large. The empirical measure  $\frac{1}{N} \sum_{j=1}^N \delta_{X_t^{j,N,\sigma}}$  therefore converges towards a measure  $\mu_t^\sigma$  which corresponds to the own distribution of  $X_t^{1,\infty,\sigma}$ . In fact this law corresponds to  $\mathbb{P}_{X_t^\sigma}$ . For a rigorous proof of this statement, see [23, 22]. A large deviations principle which permits us to describe the weak convergence speed of the empirical measure towards  $\mu_t^\sigma$  (in the McKean-Vlasov limit) can be obtained from a generalization of the theory of Freidlin and Wentzell, see [5]. It is also possible to adapt a coupling result developed for instance in [1] in order to obtain the following convergence:

$$\lim_{N \rightarrow \infty} \mathbb{E} \left\{ \sup_{0 \leq t \leq T} \|X_t^{1,N,\sigma} - X_t^\sigma\|^2 \right\} = 0.$$

*Large deviation principle.* For  $N$  large, the diffusion  $X^\sigma$  defined in (1.2) is close to the diffusion  $X^{1,N,\sigma}$  defined in (1.6). Then it is of particular interest to know if these two diffusions have the same small-noise asymptotic behaviour. The large deviations associated with (1.6) are quite classical since the system of particles is a Kolmogorov diffusion of the form

$$dX_t^{i,N,\sigma} = \sigma dB_t^i - N \times \nabla_{x_i} \Upsilon^N(X_t^{1,N,\sigma}, \dots, X_t^{N,N,\sigma}) dt, \quad X_0^{i,N,\sigma} = x, \quad (1.7)$$

with the following potential:

$$\Upsilon^N(z_1, \dots, z_N) := \int_{\mathbb{R}^d} V(x) \mu^N(dx) + \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} F(x-y) \mu^N(dx) \mu^N(dy).$$

Here  $\mu^N := \frac{1}{N} \sum_{j=1}^N \delta_{z_j}$ . This approach is directly linked to the particular form of the interaction function  $F$  which only depends on the norm, see hypothesis (F1) and (F2). The good rate function, associated with the uniform topology, is a functional defined by

$$I^N(\Phi) := \frac{1}{4} \int_0^T \|\dot{\Phi}(t) + N \nabla \Upsilon^N(\Phi(t))\|^2 dt,$$

if  $\Phi : [0, T] \rightarrow (\mathbb{R}^d)^N$  is an absolutely continuous function with the initial condition  $\Phi(0) = \bar{x} := (x, \dots, x)$  and  $I^N(\Phi) := +\infty$  otherwise. The rate function  $I^N$  can therefore be rewritten in this way: if  $\Phi := (f_1, \dots, f_N)$ , we obtain

$$I^N(\Phi) = \frac{1}{4} \sum_{i=1}^N \int_0^T \|\dot{f}_i(t) + \nabla V(f_i(t)) + \frac{1}{N} \sum_{k=1}^N \nabla F(f_i(t) - f_k(t))\|^2 dt \quad (1.8)$$

if  $f_i \in \mathcal{H}_x$  for any  $1 \leq i \leq N$ . If one function of the family  $(f_i)_{1 \leq i \leq N}$  does not belong to  $\mathcal{H}_x$ , then we set  $I^N(f_1, \dots, f_N) := +\infty$ . Let us just note that this LDP leads to the description of the exit problem for the McKean-Vlasov system [24]. Since a LDP holds for the whole particle system, a LDP in particular holds for the first particle  $(X^{1,N,\sigma})$  with the good rate function  $J_N$  obtained by projection:

$$J_N(f) := \inf_{f_2, \dots, f_N \in \mathcal{H}_x} I^N(f, f_2, \dots, f_N). \quad (1.9)$$

Since  $X^{1,N,\sigma}$  is close to the self-stabilizing process  $X^\sigma$ , solution of (1.2), ‘our objective is to establish that the functional  $J_N$  converges towards  $J_\infty$ , the entropy function of the mean-field diffusion, as  $N$  becomes large.

In other words, we show that it is possible to interchange the limiting operations concerning the asymptotic small noise  $\sigma$  and the number of particles  $N$  i.e. the hydrodynamic limit.

## 2. Convergence of the Rate Functions

In this section, we emphasize the main result of this study. We prove that the large deviation rate function  $J_N$  associated to the first particle in the huge McKean-Vlasov system of particles in interaction is close to the rate function of the self-stabilizing (nonlinear) diffusion.

**Theorem 2.1.** *Let  $x \in \mathbb{R}^d$ . Under Hypotheses (V1)–(V2) and (F1)–(F3), the rate function  $J_N$  defined by (1.9) converges towards  $J_\infty$  defined by (1.4), as  $N$  tends to infinity. Let  $f$  be an absolutely continuous function from  $[0, T]$  to  $\mathbb{R}^d$  such that  $f(0) = x$ , then*

$$\lim_{N \rightarrow +\infty} J_N(f) = J_\infty(f).$$

*Moreover, the convergence is uniform with respect to  $f$  on any compact subset of  $\mathcal{C}([0, T], \mathbb{R}^d)$  endowed with the uniform topology.*

*Proof. Step 1.* Let us first prove (easiest part) the upper-bound

$$\limsup_{N \rightarrow +\infty} J_N(f) \leq J_\infty(f). \quad (2.1)$$

By definition,  $J_N(f) \leq I^N(f, f_2, \dots, f_N)$  for any  $f_2, \dots, f_N \in \mathcal{H}_x$  where  $I^N$  is defined by (1.8). Hence, we can choose  $f_k := \Psi_\infty^x$  for all  $2 \leq k \leq N$ . Let us remind

the reader that  $\Psi_\infty^x$  is the solution of (1.5). Thus we obtain:

$$\begin{aligned} J_N(f) &\leq I^N(f, \Psi_\infty^x, \dots, \Psi_\infty^x) \\ &\leq \frac{1}{4} \int_0^T \|\dot{f}(t) + \nabla V(f(t)) + \left(1 - \frac{1}{N}\right) \nabla F(f(t) - \Psi_\infty^x(t))\|^2 dt \\ &\quad + \frac{N-1}{4} \int_0^T \|\dot{\Psi}_\infty^x(t) + \nabla V(\Psi_\infty^x(t)) + \frac{1}{N} \nabla F(\Psi_\infty^x(t) - f(t))\|^2 dt. \end{aligned}$$

By definition of  $\Psi_\infty^x$ , we have  $\dot{\Psi}_\infty^x + \nabla V(\Psi_\infty^x) = 0$ . The previous inequality yields:

$$\begin{aligned} J_N(f) &\leq \frac{1}{4} \int_0^T \|\dot{f}(t) + \nabla V(f(t)) + \left(1 - \frac{1}{N}\right) \nabla F(f(t) - \Psi_\infty^x(t))\|^2 dt \\ &\quad + \frac{1}{4N} \int_0^T \|\nabla F(\Psi_\infty^x(t) - f(t))\|^2 dt. \end{aligned}$$

Taking the limit as  $N$  goes to infinity in the previous inequality leads to the announced upper-bound (2.1). Let us just note that, due to the local Lipschitz property of the interaction function  $\nabla F$ , this convergence is uniform with respect to  $f$  on any compact set for the uniform topology.

**Step 2.** Let us focus our attention to the lower bound:

$$\liminf_{N \rightarrow +\infty} J_N(f) \geq J_\infty(f). \quad (2.2)$$

**Step 2.1** Let us recall that  $J_N$  is defined as a minimum (1.9) and let us prove that it is reached: there exists  $(f_2^*, \dots, f_N^*) \in \mathcal{H}_x^{N-1}$  such that

$$J_N(f) = I^N(f, f_2^*, \dots, f_N^*).$$

If we consider a function  $g \in \mathcal{H}_x$  then the function  $\bar{g}$  defined by  $\bar{g}(t) := g(t) - x$  for all  $t \in [0, T]$  belongs to  $\mathcal{H}_0$  which is an Hilbert space endowed with the usual norm  $\|\bar{g}\|_H^2 := \int_0^T \|\dot{g}(t)\|^2 dt$ . Let us introduce now the Hilbert space  $\mathcal{H}_0^{N-1}$  with the norm  $\|(g_2, \dots, g_N)\|^2 := \sum_{k=2}^N \|g_k\|_H^2$ . Due to the regularity of both the interaction potential  $F$  and the confining potential  $V$ , it is quite simple to prove that

$$(\bar{g}_2, \dots, \bar{g}_N) \mapsto I^N(f, g_2, \dots, g_N)$$

is a continuous function in the Hilbert space  $\mathcal{H}_0^{N-1}$  (the details are left to the reader).

Since  $J_N(f)$  is the minimum, then for any  $\epsilon > 0$ , there exist  $f_2^\epsilon, \dots, f_N^\epsilon$  belonging to  $\mathcal{H}_x$  such that

$$I^N(f, f_2^\epsilon, \dots, f_N^\epsilon) \leq J_N(f) + \epsilon.$$

Let us consider the set  $\mathcal{S}_f^x \subset \mathcal{H}_0^{N-1}$  of functions  $(\bar{g}_2, \dots, \bar{g}_N)$  satisfying

$$I^N(f, g_2, \dots, g_N) \leq 2J_\infty(f). \quad (2.3)$$

By (2.1), for  $N$  large enough and  $\epsilon$  small enough, we obtain that  $(\bar{f}_2^\epsilon, \dots, \bar{f}_N^\epsilon) \in \mathcal{S}_f^x$ . Moreover, let us prove that the subset  $\mathcal{S}_f^x$  is included in a closed ball of  $\mathcal{H}_0^{N-1}$ .



Indeed the following inequality holds

$$I^N(f, g_2, \dots, g_N) \geq \frac{1}{4} \sum_{i=1}^N \int_0^T \|\dot{g}_i(t)\|^2 dt + R_1(f, g, N) + R_2(f, g, N) \quad (2.4)$$

with

$$\begin{cases} R_1(f, g, N) := \frac{1}{2} \sum_{i=1}^N \int_0^T \langle \dot{g}_i(t), \nabla V(g_i(t)) \rangle dt, \\ R_2(f, g, N) := \frac{1}{2N} \sum_{i=1}^N \sum_{k=1}^N \int_0^T \langle \dot{g}_i(t), \nabla F(g_i(t) - g_k(t)) \rangle dt, \end{cases}$$

and the convention  $g_1 := f$ . Here  $\langle \cdot, \cdot \rangle$  stands for the Euclidean scalar product in  $\mathbb{R}^d$ . We first observe that Hypothesis (F1) leads to

$$\begin{aligned} R_2(f, g, N) &= \frac{1}{4N} \sum_{i=1}^N \sum_{k=1}^N \int_0^T \langle \dot{g}_i(t) - \dot{g}_k(t), \nabla F(g_i(t) - g_k(t)) \rangle dt \\ &= \frac{1}{4N} \sum_{i=1}^N \sum_{k=1}^N F(g_i(T) - g_k(T)) \geq \frac{N}{4} \inf_{z \in \mathbb{R}^d} F(z). \end{aligned}$$

Due to the hypothesis on the interaction function  $F$  the right hand side of the previous inequality is finite. With similar arguments, we get

$$R_1(f, g, N) = \frac{1}{2} \sum_{i=1}^N (V(g_i(T)) - V(x)) \geq \frac{N}{2} \inf_{z \in \mathbb{R}^d} V(z) - \frac{N}{2} V(x).$$

These two previous inequalities combined with (2.4) and (2.3) permit to prove the existence of a constant  $C(N, x, f)$  only depending on  $f$ ,  $x$  and  $N$  such that

$$\sum_{i=2}^N \int_0^T \|\dot{g}_i(t)\|^2 dt \leq C(N, x, f). \quad (2.5)$$

We immediately deduce that the subset  $\mathcal{S}_f^x$  is included in a closed ball of the Hilbert space  $\mathcal{H}_0^{N-1}$ . Since  $(\overline{f_2^\epsilon}, \dots, \overline{f_N^\epsilon}) \in \mathcal{S}_f^x$ , it is possible to extract a subsequence  $(\overline{f_2^{\epsilon_n}}, \dots, \overline{f_N^{\epsilon_n}})_{n \geq 0}$  which converges in the weak topology towards a limiting function  $(\overline{f_2^*}, \dots, \overline{f_N^*}) \in \mathcal{H}_0^{N-1}$ . Finally, due to the lower semicontinuity of the function  $I_N$  we deduce that:

$$I^N(f, \overline{f_2^*}, \dots, \overline{f_N^*}) = \lim_{n \rightarrow \infty} I^N(f, \overline{f_2^{\epsilon_n}}, \dots, \overline{f_N^{\epsilon_n}}) \leq J_N(f) + \lim_{n \rightarrow \infty} \epsilon_n = J_N(f).$$

The minimum  $J_N$  is then reached for  $(\overline{f_2^*}, \dots, \overline{f_N^*}) \in \mathcal{H}_x^{N-1}$ .

**Step 2.2** In order to compute  $J_N$ , let us point out particular properties of the functions  $(\overline{f_2^*}, \dots, \overline{f_N^*})$ . Since the minimum is reached, we are going to compute the derivative of  $I^N$  with respect to each coordinate. Since we restrict ourselves to the functional space  $\mathcal{H}_x$ , we take an absolutely continuous function  $g \in \mathcal{H}_0$  and we consider the following limit

$$D_2 I^N(f, \overline{f_2^*}, \dots, \overline{f_N^*})(g) := \lim_{\delta \rightarrow 0} \frac{I^N(f, \overline{f_2^*} + \delta g, \overline{f_3^*}, \dots, \overline{f_N^*}) - I^N(f, \overline{f_2^*}, \dots, \overline{f_N^*})}{\delta}.$$

That defines the derivative of  $I^N$  with respect to the second argument. In a similar way, we can define  $D_i I^N$  for any  $2 \leq i \leq N$ . Let us compute explicitly  $D_2 I^N$ . For any  $1 \leq i \leq N$ , we set

$$\xi_i := \dot{f}_i + \nabla V(f_i) + \frac{1}{N} \sum_{k=1}^N \nabla F(f_i - f_k), \quad (2.6)$$

with  $f_1 := f$  by convention. So we note that

$$I^N(f, f_2, \dots, f_N) = \frac{1}{4} \sum_{i=1}^N \int_0^T \|\xi_i(t)\|^2 dt. \quad (2.7)$$

We now observe the derivative of the quantity  $\xi_i$  with respect to the second function. In other words, introducing  $f_2^\delta := f_2 + \delta g$ , and defining  $\xi_i^{2,\delta}$  like  $\xi_i$  just by replacing  $f_2$  by  $f_2^\delta$  in (2.6), we get

$$\lim_{\delta \rightarrow 0} \frac{\xi_i^{2,\delta} - \xi_i}{\delta} = -\frac{1}{N} H(F)(f_i - f_2)g, \quad \text{for } i \neq 2,$$

and, for  $i = 2$ :

$$\lim_{\delta \rightarrow 0} \frac{\xi_2^{2,\delta} - \xi_2}{\delta} = \dot{g} + H(V)(f_2)g + \frac{1}{N} \sum_{\substack{1 \leq k \leq N \\ k \neq 2}} H(F)(f_2 - f_k)g.$$

Here  $H(F)(x)$  represents the Hessian matrix of the function  $F$  at the point  $x \in \mathbb{R}^d$ . From now on, we simplify the notation  $D_2 I^N(f, f_2, \dots, f_N)(g)$  and replace it by  $D_2 I^N$ . By (2.7) and the polarization identity, we obtain

$$\begin{aligned} D_2 I^N &= \lim_{\delta \rightarrow 0} \frac{1}{4\delta} \sum_{i=1}^N \int_0^T \|\xi_i^{2,\delta}(t)\|^2 - \|\xi_i(t)\|^2 dt \\ &= \lim_{\delta \rightarrow 0} \frac{1}{4\delta} \sum_{i=1}^N \int_0^T \left( \|\xi_i^{2,\delta}(t) - \xi_i(t)\|^2 + 2\langle \xi_i(t), \xi_i^{2,\delta}(t) - \xi_i(t) \rangle \right) dt \\ &= -\frac{1}{2N} \sum_{i=1}^N \int_0^T \langle \xi_i(t), H(F)(f_i - f_2)g(t) \rangle dt \\ &\quad + \frac{1}{2} \int_0^T \langle \xi_2(t), \dot{g}(t) + \mathcal{R}_2^f g(t) \rangle dt, \end{aligned} \quad (2.8)$$

with

$$\mathcal{R}_2^f := H(V)(f_2) + \frac{1}{N} \sum_{i=1}^N H(F)(f_2 - f_i). \quad (2.9)$$

Let us assume that  $\xi_2$  is regular (let us say continuously differentiable). Then by an integration by parts and since  $g(0) = 0$ , we obtain

$$D_2 I^N = -\frac{1}{2} \int_0^T \langle \mathcal{E}_2^f(\xi_2, \dots, \xi_N)(t), g(t) \rangle dt + \frac{1}{2} \langle \xi_2(T), g(T) \rangle,$$

where the function  $\mathcal{E}_2^f$  is defined by

$$\mathcal{E}_2^f(\xi_2, \dots, \xi_N)(t) := \dot{\xi}_2(t) - \mathcal{R}_2^f \xi_2(t) + \frac{1}{N} \sum_{i=1}^N H(F)(f_2 - f_i) \xi_i(t). \quad (2.10)$$

We proceed in the same way for  $D_j I^N(f_1, f_2, \dots, f_N)(g)$  for any  $2 \leq j \leq N$ , just replacing 2 by  $j$  in (2.9) and (2.10). Since the minimum  $J_N(f)$  is reached, that is  $J_N(f) = I^N(f, f_2^*, \dots, f_N^*)$  for some functions  $(f_2^*, \dots, f_N^*) \in \mathcal{H}_x^{N-1}$  (see Step 2.1), the following expression vanishes:  $D_j I^N(f, f_2^*, \dots, f_N^*)(g) = 0$  for  $2 \leq j \leq N$  and any function  $g$ , and therefore  $(\xi_2^*, \dots, \xi_N^*)$  is solution to the system

$$\begin{cases} \mathcal{E}_j^{f^*}(\xi_2^*, \dots, \xi_N^*) = 0, \\ \xi_j^*(T) = 0 \quad \text{for any } 2 \leq j \leq N, \end{cases} \quad (2.11)$$

where the functions  $\mathcal{E}_j^{f^*}$  are defined like  $\mathcal{E}_j^f$  in (2.10) (respectively  $\xi_j^*$  like  $\xi_j$  in (2.6)), we need just to replace  $(f, f_2, \dots, f_N)$  by  $(f, f_1^*, \dots, f_N^*)$ .

In fact we do not know if  $\xi_j^*$ ,  $2 \leq j \leq N$ , are regular functions as assumed. We deduce therefore that  $(\xi_2^*, \dots, \xi_N^*)$  is a *generalized solution* of the system (2.11).

**Step 2.3** In Step 2.1, we proved that  $(f_2^*, \dots, f_N^*)$  minimizes the function  $I^N$  and, by Step 2.2, that  $(\xi_2^*, \dots, \xi_N^*)$  which can be expressed as a function of  $(f_2^*, \dots, f_N^*)$  – see (2.6) – satisfies a particular system of differential equations, in a generalized sense, namely (2.11). Of course  $(\xi_2^*(t), \dots, \xi_N^*(t)) = (0, \dots, 0)$  is a classical solution (and consequently a generalized solution) of (2.11). By uniqueness of the *generalized solutions*, we shall obtain that

$$\xi_j^*(t) = 0 \quad \text{for a.e. } 0 \leq t \leq T \quad \text{and for } 2 \leq j \leq N.$$

Let us prove this uniqueness property. Let us consider a function  $h = (h_2, \dots, h_N)$  belonging to  $\otimes^{N-1} \mathcal{C}([0, T], \mathbb{R}^d)$ , then using the Cauchy-Lipschitz theorem (see for instance Theorem 3.1 in [13]), there exists a  $\mathcal{C}^1$ -solution  $(g_2, \dots, g_N)$  of the following system of equations:

$$\dot{g}_j(t) = -\mathcal{R}_j^{f^*} g_j(t) - \frac{1}{N} \sum_{k=1}^N H(F)(f_j^* - f_k^*) g_k(t) + h_j(t), \quad 2 \leq j \leq N, \quad (2.12)$$

with the initial condition  $g_j(0) = 0$  for any  $2 \leq j \leq N$ . Here  $f_1^*$  stands for  $f$  for notational convenience. In particular  $g \in \mathcal{H}_0^{N-1}$ . Using results developed in Step 2.2, the function  $I^N$  reaches its minimal value for the arguments  $(f_2^*, \dots, f_N^*)$  and consequently:

$$\sum_{j=2}^N D_j I^N(f, f_2^*, \dots, f_N^*)(g_j) = 0,$$

where  $g_j$  are solutions of (2.12). Combining the expression (2.8) and (2.12) leads to

$$\sum_{j=2}^N \int_0^T \langle \xi_j^*(t), h_j(t) \rangle dt = 0, \quad (2.13)$$

for any continuous functions  $(h_j, 2 \leq j \leq N)$ . Since  $f^*$  is in the function space  $\mathcal{H}_x^{N-1}$  and since  $\xi_j^*$  is related to  $f^*$  by (2.6), we know that  $\xi_j$  is square integrable

function. Using Carleson's theorem (see for instance Theorem 1.9 in [10]) and (2.13) we deduce that  $\xi_j^*(t) = 0$  for a.e.  $t \in [0, T]$  and for any  $2 \leq j \leq N$ .

**Step 2.4** Using the definition of  $\xi_j^*$ , the previous step permits to obtain that, for any  $2 \leq j \leq N$ ,

$$\dot{f}_j^*(t) + \nabla V(f_j^*)(t) + \frac{1}{N} \sum_{k=1}^N \nabla F(f_j^* - f_k^*)(t) = 0 \quad \text{for a.e. } t \in [0, T], \quad (2.14)$$

with the boundary condition  $f_j^*(0) = x$ . Applying once again the arguments presented in Step 2.3 leads to the uniqueness of the solutions for (2.14). Since the system is symmetric, we get the existence of a  $\mathcal{C}^1$ -function  $\Psi_N^f$  satisfying

$$f_2^*(t) = \dots = f_N^*(t) = \Psi_N^f(t) \quad \text{for a.e. } t \in [0, T],$$

and, on the time interval  $[0, T]$ ,

$$\dot{\Psi}_N^f(t) + \nabla V(\Psi_N^f)(t) + \frac{1}{N} \nabla F(\Psi_N^f - f)(t) = 0, \quad \text{with } \Psi_N^f(0) = x. \quad (2.15)$$

We just recall that  $f_1^* = f$  for notational convenience. Using the definition of  $J_N(f)$ , we get

$$\begin{aligned} J_N(f) &= I^N(f, f_2^*, \dots, f_N^*) \\ &= \int_0^T \|\dot{f}(t) + \nabla V(f(t)) + \left(1 - \frac{1}{N}\right) \nabla F(f(t) - \Psi_N^f(t))\|^2 dt. \end{aligned} \quad (2.16)$$

Since the first order differential equation (2.15) can be associated to a Lipschitz constant which does not depend on the parameter  $1/N$ , the unique solution  $\Psi_N^f(t)$  depends continuously on both the parameter  $1/N$  and the time variable (see, for instance, Theorem 3.2 p. 20 in [13]). Here we consider that  $(t, 1/N)$  belongs to the compact set  $[0, T] \times [0, 1]$ , consequently  $(t, 1/N) \mapsto \Psi_N^f(t)$  is uniformly continuous. Moreover,  $\Psi_\infty^f = \Psi_\infty^x$ . Hence (2.16) implies

$$\lim_{N \rightarrow \infty} J_N(f) = J_\infty(f),$$

where  $J_\infty$  is defined by (1.4) and (1.5). The proof of the lower-bound (2.2) is then achieved. It is quite easy to prove that the convergence is uniform with respect to the function  $f$  on any compact set for the uniform topology.  $\square$

We could provide the precise rate of convergence because we give better than a simple lower-bound in the previous proof. Indeed, we have obtained the exact expression of  $J_N(f)$ .

### 3. Immediate Consequences and Further Results

Theorem 2.1 emphasizes the link between the large deviation rate function of the self-stabilizing diffusion (1.2) and the rate function associated to the mean-field system (1.6). This result is of particular interest since one of the diffusion is nonlinear whereas the second one is linear and therefore well-known. In this section, we present a coupling result concerning these two diffusions and extend Theorem 2.1 to a more general nonlinear model.

Let us first recall the large deviation principle already presented in the introduction. A family of continuous stochastic processes  $(X^\sigma)_{\sigma>0}$  is said to satisfy a *large deviation principle* for the uniform topology with good rate function  $I$  if the level sets of  $I$  are compact subsets of  $\mathcal{C}([0, T]; \mathbb{R}^d)$  and if

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(X^\sigma \in \mathbb{F}) \leq - \inf_{\varphi \in \mathbb{F}} I(\varphi)$$

for any closed subset  $\mathbb{F} \subset \mathcal{C}([0, T]; \mathbb{R}^d)$  equipped with the uniform topology and

$$\liminf_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(X^\sigma \in \mathbb{G}) \geq - \inf_{\varphi \in \mathbb{G}} I(\varphi)$$

for any open subset  $\mathbb{G} \subset \mathcal{C}([0, T]; \mathbb{R}^d)$ .

Theorem 2.1 ensures the convergence of  $J_N(f)$  to  $J_\infty(f)$  for any continuous function  $f$ . In order to point out the large deviation principle, we need to precise this convergence on open and closed subsets of continuous functions.

**Corollary 3.1.** *For any open or closed subset  $\mathbb{O} \subset \mathcal{C}([0, T]; \mathbb{R}^d)$  (for the uniform topology) the following convergence holds*

$$\lim_{N \rightarrow \infty} \inf_{\varphi \in \mathbb{O}} J_N(\varphi) = \inf_{\varphi \in \mathbb{O}} J_\infty(\varphi). \quad (3.1)$$

*Proof.* Let  $\mathbb{O}$  be a non empty set. If  $\inf_{\varphi \in \mathbb{O}} J_\infty(\varphi) = +\infty$  then  $\mathbb{O} \subset \mathcal{H}^c$  where  $\mathcal{H}^c$  is the complementary of the set of absolutely continuous functions. Therefore  $J_N(\varphi) = +\infty$  for any function  $\varphi \in \mathbb{O}$  and (3.1) is obviously satisfied. Let us assume now that  $\inf_{\varphi \in \mathbb{O}} J_\infty(\varphi) = \alpha < \infty$ . Denoting by

$$\kappa_\lambda = \{\varphi \in \mathcal{C}([0, T], \mathbb{R}^d) : J_\infty(\varphi) \leq \lambda\}$$

which is a compact set since  $J_\infty$  is a *good* rate function (see [14]), we obtain

$$\inf_{\varphi \in \mathbb{O}} J_\infty(\varphi) = \inf_{\varphi \in \mathbb{O} \cap \kappa_{2\alpha}} J_\infty(\varphi).$$

In order to conclude the proof, it suffices to apply the convergence of  $J_N$  towards  $J_\infty$  developed in Theorem 2.1, which is in fact uniform with respect to  $\varphi$  on any compact subset, in particular on the subset  $\kappa_{2\alpha}$ .  $\square$

This first corollary concerns the rate functions. Let us now focus our attention on the associated processes. A nice coupling property can be obtained describing the link between the self-stabilizing diffusion  $(X_t^\sigma, t \geq 0)$  defined by (1.2) and the linear diffusion  $(X_t^{1,N,\sigma}, t \geq 0)$  defined by (1.6). Since there exists a unique strong solution to each of these two equations, we can construct  $X^\sigma$  and  $X^{1,N,\sigma}$  on the same probability space  $(\Omega, \mathcal{B}, \mathbb{P}_x)$ .

**Corollary 3.2.** *Under Hypotheses (V1)–(V2) and (F1)–(F3), for any  $x \in \mathbb{R}^d$ , each element of the family  $(X^{1,N,\sigma})_N$  converges in probability towards the diffusion  $X^\sigma$  as  $\sigma \rightarrow 0$ , uniformly with respect to the parameter  $N$ . In particular let  $\delta > 0$ , then for  $N$  sufficiently large (resp.  $\sigma$  small), there exists a constant  $K_\delta(T) > 0$  such that*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \|X_t^{1,N,\sigma} - X_t^\sigma\| \geq \delta\right) \leq e^{-\frac{K_\delta(T)}{\sigma^2}}.$$

*Remark 3.3.* Let us just note that this result implies the convergence in distribution of the first particle – first coordinate – of the linear mean-field system (1.6) towards the self-stabilizing process. Combining Corollary 3.1 and Corollary 3.2 leads to the following statement: for any closed set  $\mathbb{F}$ ,

$$\lim_{N \rightarrow \infty} \limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log \mathbb{P}(X^{1,N,\sigma} \in \mathbb{F}) \quad \text{and} \quad \limsup_{\sigma \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\sigma^2}{2} \log \mathbb{P}(X^{1,N,\sigma} \in \mathbb{F})$$

have the same upper bound namely:  $-\inf_{\varphi \in \mathbb{F}} J_\infty(\varphi)$ . A similar result holds for open sets  $\mathbb{G}$ , replacing both the limit inferior by the limit superior and the upper bound by a lower one.

*Remark 3.4.* Large deviation principles can sometimes be proven directly by the use of coupling bounds. Nevertheless the coupling bound developed in Corollary 3.2 is not strong enough for such implications. Indeed, the family  $(X^{1,N,\sigma})_N$  is called an *exponentially good approximation* of  $X^\sigma$  if, for any  $\delta > 0$ ,

$$\lim_{N \rightarrow \infty} \limsup_{\sigma \rightarrow 0} \sigma^2 \log \mathbb{P}_x \left( \sup_{0 \leq t \leq T} \|X_t^{1,N,\sigma} - X_t^\sigma\| \geq \delta \right) = -\infty \quad (3.2)$$

(such a large deviation notion has been for instance developed in [7], Definition 4.2.14). For such approximations, the rate function of the limiting process (as  $N$  tends to  $\infty$ ) can be obtain as follows:

$$J_\infty(\varphi) := \sup_{\delta > 0} \liminf_{N \rightarrow \infty} \inf_{z \in B(\varphi, \delta)} J_N(z),$$

where  $B(\varphi, \delta) = \{z : \sup_{0 \leq t \leq T} \|z(t) - \varphi(t)\| < \delta\}$ . In practice, (3.2) is quite difficult to obtain. Such techniques were used in order to prove the Freidlin-Wentzell large deviation result for classical diffusions: the process is approximated by an other stochastic process with piecewise constant diffusion and drift terms (see Theorem 5.6.7 in [7]). For the large deviation principle associated with the self-stabilizing diffusion developed in [14], an argument of exponentially good approximation is used but it does not concern the approximation of the non linear process by the first particle of the mean field linear system and therefore it does not use (3.2).

*Proof.* By definition, the family of processes  $(X^\sigma)_{\sigma > 0}$  satisfies a large deviation principle associated with the good rate function  $J_\infty$ . So, for any closed subset  $\mathbb{F}$  of  $\mathcal{H}_x$ , we have on one hand

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log [\mathbb{P} \{X^\sigma \in \mathbb{F}\}] \leq -\inf_{\varphi \in \mathbb{F}} J_\infty(\varphi).$$

On the second hand, by Corollary 3.1 and for  $N$  large enough, we obtain

$$\limsup_{\sigma \rightarrow 0} \frac{\sigma^2}{2} \log [\mathbb{P} \{X^{1,N,\sigma} \in \mathbb{F}\}] \leq -\inf_{\varphi \in \mathbb{F}} J_N(\varphi) \leq -\frac{3}{4} \inf_{\varphi \in \mathbb{F}} J_\infty(\varphi).$$

Introducing the particular subset:

$$\mathbb{F} := \left\{ \varphi \in \mathcal{H}_x : \exists t_0 \in [0, T] \text{ s.t. } \|\varphi(t_0) - \Psi_\infty^x(t_0)\| \geq \frac{\delta}{4} \right\},$$

where  $\Psi_\infty^x$  is defined in (1.5) that is  $\Psi_\infty^x(t) := x - \int_0^t \nabla V(\Psi_\infty^x(s)) ds$ , we observe that, for  $N$  large and  $\sigma$  small,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \|X_t^{1,N,\sigma} - X_t^\sigma\| \geq \delta\right) \leq \mathbb{P}(X^{1,N,\sigma} \in \mathbb{F}) + \mathbb{P}(X^\sigma \in \mathbb{F}) \leq e^{-\frac{K_\delta(T)}{\sigma^2}},$$

where  $K_\delta(T) := C \inf_{\varphi \in \mathbb{F}} J_\infty(\varphi) > 0$  with  $0 < C < 3/4$ .  $\square$

In Theorem 2.1, we only deal with the *gradient case* of the so-called McKean-Vlasov diffusion starting from initial position  $x$ . Let us now discuss a more general setting by considering the following nonlinear diffusion:

$$Y_t^\sigma = x + \sigma B_t - \int_0^t \nabla V(Y_s^\sigma) ds - \int_0^t \int_{\mathbb{R}^d} \mathcal{A}(Y_s^\sigma, y) \nu_s^\sigma(dy) ds - l(t).$$

Here  $\mathcal{A}$  is a general two variables  $\mathbb{R}^d$ -valued function being a vector flow, non necessary gradient and  $l$  is a  $\mathcal{C}^1$ -continuous function from  $\mathbb{R}_+$  to  $\mathbb{R}^d$ . Finally the probability measure  $\nu_s^\sigma$  stands for the distribution  $\mathbb{P}_{Y_s^\sigma}$ . The aim of this discussion does not concern the existence and uniqueness of such equation, so we assume that  $V$ ,  $\mathcal{A}$  and  $l$  satisfy suitable conditions for the unique solution to exist. Then, it is possible to adapt the arguments developed in [14] in order to prove that  $(Y^\sigma)_{\sigma>0}$  satisfies a large deviation principle with the associated rate function:

$$J_\infty(f) := \frac{1}{4} \int_0^T \|\dot{f}(t) + \nabla V(f(t)) + \mathcal{A}(f(t), \Psi^x(t)) + \dot{l}(t)\|^2 dt \quad (3.3)$$

for any function  $f \in \mathcal{H}_x$  and  $J_\infty(f) := +\infty$  otherwise. Here, the function  $\Psi^x$  is defined as the unique solution of the ordinary differential equation:

$$\Psi^x(t) = x - \int_0^t \nabla V(\Psi^x(s)) ds - \int_0^t \mathcal{A}(\Psi^x(s), \Psi^x(s)) ds - l(t).$$

The stochastic model  $(Y_t^\sigma)$  can also be approximated by a system of interacting particles. In this context, we can develop a statement similar to Theorem 2.1. The functional  $J_\infty$  is effectively the limit as  $N$  goes to infinity of the functional  $J_N(f)$  defined by

$$\inf_{f_2, \dots, f_N \in \mathcal{H}_x} \frac{1}{4} \sum_{i=1}^N \int_0^T \|\dot{f}_i(t) + \nabla V(f_i(t)) + \frac{1}{N} \sum_{j=1}^N \mathcal{A}(f_i(t), f_j(t)) + \dot{l}(t)\|^2 dt \quad (3.4)$$

with the convention  $f_1 = f$ . Such a result can be proven under suitable assumptions:

- the confining potential  $V$  satisfies Hypotheses (V1)–(V2).
- there exists a *lower bounded*  $\mathcal{C}^\infty$ -function  $\mathbb{A} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that

$$\mathcal{A}(x, y) = \nabla_x \mathbb{A}(x, y) \quad \text{and} \quad \inf_{(x,y) \in \mathbb{R}^d \times \mathbb{R}^d} \mathbb{A}(x, y) > -\infty.$$

- $\mathcal{A}$  satisfies a symmetry property:  $\mathcal{A}(x, y) = -\mathcal{A}(y, x)$

The details of the proof are left to the reader, it suffices to apply the same arguments. Let us just note that the assumptions, just formulated, concerning  $\mathcal{A}$  are sufficient in order to get the upper-bound (2.5), a crucial step for proving the claimed statement.

We end this study pointing out an example of such a diffusion:

$$dX_t = \sigma dB_t - (\nabla W(X_t) - \overline{W}_t) dt - \dot{l}(t)dt,$$

where  $\overline{W}_t := \mathbb{E} \{ \nabla W(X_t) \}$  and  $W$  is such that the required conditions are satisfied. This equation actually corresponds to the hydrodynamic limit of an equation characterizing the charge and the discharge of the cathode in a lithium battery (see [8, 9]). In such a framework,  $\mathcal{A}(x, y) := \nabla W(x) - \nabla W(y)$  and  $\mathbb{A}(x, y) := W(x) - \langle x, \nabla W(y) \rangle$ . Therefore, the rate function can be explicitly computed:

$$J_\infty(f) = \frac{1}{4} \int_0^T \|\dot{f}(t) + \nabla W(f(t)) - \nabla W(x + l(t)) + \dot{l}(t)\|^2 dt$$

and is obtained, as announced, as the limit for large  $N$  of the rate function:

$$\inf_{f_2, \dots, f_N \in \mathcal{H}_x} \frac{1}{4} \sum_{i=1}^N \int_0^T \|\dot{f}_i(t) + \nabla W(f_i(t)) - \frac{1}{N} \sum_{j=1}^N \nabla W(f_j(t)) + \dot{l}(t)\|^2 dt.$$

**Acknowledgment.** Samuel Herrmann would like to thank Dierk Peithmann for interesting discussions concerning this study. Julian Tugaut would like to thank the Institut de Mathématiques de Bourgogne for the time spent there which permits the idea of this article to emerge. J. T. also thanks Bielefeld University and especially Barbara Gentz. Velika hvala Marini za sve. Également, un très grand merci à Manue et à Sandra pour tout.

## References

1. Benachour, S., Roynette, B., Talay, D., and Vallois, P.: Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos, *Stochastic Process. Appl.*, **75** (1998), no. 2, 173–201.
2. Benachour, S., Roynette, B., and Vallois, P.: Nonlinear self-stabilizing processes. II. Convergence to invariant probability, *Stochastic Process. Appl.*, **75** (1998), no. 2, 203–224.
3. Bryc, W.: A remark on the connection between the large deviation principle and the central limit theorem, *Statist. Probab. Lett.*, **18** (1993), no. 4, 253–256.
4. Cattiaux, P., Guillin, A., and Malrieu, F.: Probabilistic approach for granular media equations in the non-uniformly convex case, *Probab. Theory Related Fields*, **140** (2008), no. 1–2, 19–40.
5. Dawson, A. D. and Gärtner, J.: Large deviations from the McKean-Vlasov limit for weakly interacting diffusions, *Stochastics*, **20** (1987), no. 4, 247–308.
6. Del Moral, P. and Tugaut, J.: Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions. available on <http://hal.archives-ouvertes.fr/hal-00798813>, 2013
7. Dembo, A. and Zeitouni, O.: *Large Deviations Techniques and Applications*, volume 38 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2010.
8. Dreyer, W., Gaberscek, M., Guhlke, C., Huth, R., and Jamnik, J.: Phase transition and hysteresis in a rechargeable lithium battery revisited, *European J. Appl. Math.*, **22** (2011), 267–290.
9. Dreyer, W., Guhlke, C., and Huth, R.: The behavior of a many particle cathode in a lithium-ion battery *Phys. D*, **240** (2011), 1008–1019.
10. Duoandikoetxea, J.: *Fourier Analysis*, Graduate Studies in Mathematics, **29**, American Mathematical Society, 2001.



11. Freidlin, M. I. and Wentzell, A. D.: *Random Perturbations of Dynamical Systems*, volume 260 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, New York, second edition, 1998.
12. Funaki, T.: A certain class of diffusion processes associated with nonlinear parabolic equations, *Z. Wahrsch. Verw. Gebiete*, **67** (1984), no. 3, 331–348.
13. Hale, J. K.: *Ordinary Differential Equations*, Pure and Applied Mathematics, Vol. XXI, Wiley-Interscience, 1969.
14. Herrmann, S., Imkeller, P., and Peithmann, D.: Large deviations and a Kramers' type law for self-stabilizing diffusions, *Ann. Appl. Probab.*, **18** (2008), no. 4, 1379–1423.
15. Herrmann, S. and Tugaut, J.: Non-uniqueness of stationary measures for self-stabilizing processes, *Stochastic Process. Appl.*, **120** (2010), no. 7, 1215–1246.
16. Herrmann, S. and Tugaut, J.: Stationary measures for self-stabilizing processes: asymptotic analysis in the small noise limit. *Electron. J. Probab.*, **15** (2010), 2087–2116.
17. Herrmann, S. and Tugaut, J.: Self-stabilizing processes: uniqueness problem for stationary measures and convergence rate in the small noise limit. *ESAIM Probab. Stat.*, **16** (2012), 277–305.
18. Kolokoltsov, V.: Nonlinear diffusions and stable-like processes with coefficients depending on the median or VaR. *Appl. Math. Optim.*, **68** (2013), no. 1, 85–98.
19. Malrieu, F.: Convergence to equilibrium for granular media equations and their Euler schemes, *Ann. Appl. Probab.*, **1** (2003), no. 2, 540–560.
20. McKean, H. P.: A class of Markov processes associated with nonlinear parabolic equations, *Proc. Nat. Acad. Sci. U.S.A.*, **56** (1966), 1907–1911.
21. McKean, H. P.: Propagation of chaos for a class of nonlinear parabolic equations, in: *Stochastic Differential Equations, Lecture Series in Differential Equations*, **7** (1967), 41–57. Catholic Univ. of America, Washington. D.C.
22. Méléard, S.: Asymptotic behaviour of some interacting particle systems, McKean-Vlasov and Boltzmann models, in: *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math.* **1627** (1996), 42–95, Springer, Berlin.
23. Sznitman, A.-S.: Topics in propagation of chaos, in: *École d'Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math.*, **1964** (1991) 165–251, Springer, Berlin.
24. Tugaut, J.: Exit problem of McKean-Vlasov diffusions in convex landscapes, *Electronic Journal of Probability*, **17** (2012), no. 76, 1–26.
25. Tugaut, J.: Convergence to the equilibria for self-stabilizing processes in double-well landscape, *Ann. Probab.*, **41** (2013), no. 3A, 1427–1460.
26. Tugaut, J.: Self-stabilizing processes in multi-wells landscape in  $\mathbb{R}^d$  - Convergence, *Stochastic Process. Appl.*, **123** (2013), no. 5, 1780–1801.

SAMUEL HERRMANN: INSTITUT DE MATHÉMATIQUES DE BOURGOGNE, IMB UMR5584, CNRS, UNIV. BOURGOGNE FRANCHE-COMTÉ, F-21000 DIJON, FRANCE  
*E-mail address:* [Samuel.Herrmann@u-bourgogne.fr](mailto:Samuel.Herrmann@u-bourgogne.fr)

JULIAN TUGAUT: UNIVERSITÉ JEAN MONNET, INSTITUT CAMILLE JORDAN, 23, RUE DU DOCTEUR PAUL MICHELON, CS 82301, 42023 SAINT-ÉTIENNE CEDEX 2, FRANCE  
*E-mail address:* [tugaut@math.cnrs.fr](mailto:tugaut@math.cnrs.fr)