

A gradient flow approach to large deviations for diffusion processes.

Max Fathi *

May 29, 2015

Abstract

In this work, we investigate links between the formulation of the flow of marginals of reversible diffusion processes as gradient flows in the space of probability measures and path wise large deviation principles for sequences of such processes. An equivalence between the LDP principle and Gamma-convergence for a sequence of functionals appearing in the gradient flow formulation is proved. As an application, we study large deviations from the hydrodynamic limit for two variants of the Ginzburg-Landau model endowed with Kawasaki dynamics.

Introduction

In this work, we are interested in the links between the gradient flow formulation of the flow of marginals of stochastic differential equations, and path wise large deviations for sequences of such processes.

Interest in gradient flows on the space of probability measures goes back to [20], where it was observed that the heat equation can be viewed as the gradient flow of the entropy

$$\text{Ent}(\rho) = \int \rho \log \rho dx$$

for the Wasserstein distance W_2 . Note that what we will call here entropy is the negative of the physical entropy. This was later developed into a notion of formal Riemannian structure on $\mathcal{P}(\mathbb{R}^n)$ by Otto in [32]. While a powerful tool to predict the behavior of certain partial differential equations, the point of view of Otto is formal, and we must rely on other tools for proofs.

Another point of view was developed by Ambrosio, Gigli and Savaré in [3], which uses the notion of ‘minimizing movement’ schemes, developed by De Giorgi and which first appeared in [8], to provide a rigorous framework to define gradient flows on spaces of probability measures. It is based on the idea that gradient flows on \mathbb{R}^n of the form

$$\dot{x}(t) = -\nabla F(x(t))$$

*LPMA, University Paris 6, France, max.fathi@etu.upmc.fr.

are the only solutions of

$$F(x(T)) - F(x(0)) + \frac{1}{2} \int_0^T |\nabla F(x(t))|^2 dt + \frac{1}{2} \int_0^T |\dot{x}(t)|^2 dt = 0. \quad (0.1)$$

While the usual gradient flow equation only makes sense in a Riemannian setting (at least in the classical sense), this alternative formulation can be given a meaning in a purely metric setting, as long as we can define a ‘length of the gradient’ functional $|\nabla F|$. Section 1.1 concerns this formulation in the setting of the space of probability measures on \mathbb{R}^n endowed with a Wasserstein distance W_2 , when the functional F is the relative entropy with respect to a nonnegative measure μ , that is

$$\text{Ent}_\mu(\nu) := \int f \log f d\mu$$

if $\nu = f\mu$, and $+\infty$ if ν is not absolutely continuous with respect to μ . This is the framework developed in the first sections of [3].

Several recent papers have been interested in using abstract gradient flow formulations to study convergence of sequences of solutions to partial differential equations. One method, tailored for the case of diffusion processes and based on the discrete approximation of gradient flows, has been devised in [4]. Another, more general, method has been presented in [38] (generalizing previous results of [37]). It consists in studying the asymptotic behavior of sequence of functionals of the form (0.1) for given functions F_n . Informally, it consists in showing that, if the sequence converges in a certain sense to a limiting functional F_∞ , we can directly identify limits of solutions of (0.1) as gradient flows for the limiting function F_∞ .

In the context of statistical physics, the method developed in [38] can be used to prove convergence to the hydrodynamic limit for some models of interacting diffusion processes, such as the Ginzburg-Landau model (see [19] or [18] for a presentation of the model, and its hydrodynamic limit). Such results consist in proving convergence in probability of the dynamics of some family of N -particle systems to a deterministic limiting object as the number of particles N goes to infinity. The limit generally appears as the solution to some partial differential equation. Gradient flows have also been used in [34] to study convergence to the hydrodynamic limit for microscopic models describing heat conduction.

Our aim here is to use the notion of gradient flows to study large deviations from the hydrodynamic limit for interacting spin systems. Such a result consists in proving that the probabilities of a significant deviation from the hydrodynamic limit decays exponentially fast in the system size. A standard textbook on the topic of large deviations is [11], and [22] contains a review of the literature in the context of large deviations from the hydrodynamic limit for particle systems.

In the work [1] (and its sequels [2] and [6]), links between gradient flows in spaces of probability measures and large deviations have been investigated for many examples of processes arising in statistical physics. The main contribution of [1] is to show that the gradient flow formulation for partial differential equations such as the heat equation can be deduced from the large deviation principle for N independent stochastic processes given by the stochastic differential equation whose flow of marginals is the solution to the PDE.

In this paper, we are interested in whether this link works both ways: can we use gradient flow structures to *prove* large deviation principles? We prove that process-level large deviations for sequences of diffusion processes are equivalent to the Gamma-convergence of a sequence of functionals that naturally appear in the gradient-flow formulation of these processes. (By process-level large deviations, we mean large deviations for the laws of trajectories of the processes, and not just the law at a given time.) This result generalizes a method used in [7] and [16] to obtain process-level large deviations for the empirical measure of independent Brownian motions to the case of interacting diffusion processes. Although these previous works do not discuss gradient flows or optimal transport, there are a lot of similarities between the formalism we use here and their framework, and the proof is based on a similar method.

The point of view developed here can be thought of in the following way: to study large deviations for sequences of reversible diffusion processes, it is enough (and even equivalent) to understand the asymptotic behavior of an underlying sequence of metric-measure spaces. In particular, the rate function arises as a limit of the interplay between the sequence of invariant (or equilibrium) measures and the sequence of (Riemannian) metric structures which is used in the gradient flow formulation.

As an application of this equivalence, we investigate the large deviations for two variants of the Ginzburg-Landau model endowed with Kawasaki dynamics, giving an alternative approach to obtaining the results of [10] and [35]. The first model is a random conductance model, and the second one is the non-gradient Ginzburg-Landau model of [39] and [35]. As far as the author knows, the large deviation principle for the random conductance model obtained here is new. The large deviations principle for the non-gradient model obtained here is not new, it was first proven in [35] (up to some extra technical assumptions), but is included to show that the method developed here also applies to non-gradient models.

Contents

1	Framework and Method	4
1.1	Gradient flows in $P_2(\mathbb{R}^n)$	4
1.2	Large deviations	8
1.3	Main result	8
1.4	Some questions	10
2	Proof of Theorem 1.10	11
2.1	Relative entropy and large deviations	12
2.2	Proof of Theorem 2.1	13
3	Large deviations for the Ginzburg-Landau model	17
3.1	The model	17
3.2	Some technical estimates	19
3.3	Large deviations for the GL model in a random environment	25

Notations and terminology

- If A is a symmetric positive matrix, then \sqrt{A} is the unique symmetric positive matrix whose square is A ;
- $\mathcal{P}_2(\mathbb{R}^d)$ is the space of probability measures on \mathbb{R}^d with finite second moment;
- $W_{2,G}$ is the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$ for the Riemannian distance d_G on \mathbb{R}^d endowed with the metric tensor $G : \mathbb{R}^d \rightarrow \mathcal{S}_{++}(\mathbb{R}^d)$. It is given by

$$W_{2,G}^2(\mu_0, \mu_1) := \inf_{\pi} \int d_G(x, y)^2 \pi(dx, dy)$$

where the infimum is taken over all coupling π of the probability measures μ_0 and μ_1 . Such a coupling is a probability measure on the product space whose first (resp. second) marginal is μ_0 (resp. μ_1). We refer to [40] and [41] for more information about optimal transport and Wasserstein distances;

- Z is a constant enforcing unit mass for a probability measure;
- C is a constant that may change from line to line, or even within a line;
- $C_b(X)$ is the space of real-valued, continuous bounded functions on the space X ;
- $\operatorname{div}(A)(x)$ is the vector of \mathbb{R}^d with coordinates $(\operatorname{div}(A)(x))_i := \sum_{j=1}^d \frac{\partial A_{ij}}{\partial x_j}(x)$, where $A : \mathbb{R}^d \rightarrow \mathcal{M}_d(\mathbb{R})$
- Given the law P of a random curve $x : [0, T] \rightarrow \mathbb{R}^d$, we call flow of marginals the family $(\nu_t)_{t \in [0, T]}$ the of probability measures on \mathbb{R}^d which describe the law of x_t for each time t ;
- We shall use both the notations $\operatorname{Ent}_{\mu}(\nu)$ and $H(\nu; \mu)$ to denote the relative entropy of a probability emasure ν with respect to the nonnegative measure μ . The notation Ent shall be used when dealing with the gradient flow side, while the notation H will be used when directly dealing with the probabilistic/large deviations side of the exposition.

1 Framework and Method

1.1 Gradient flows in $\mathcal{P}_2(\mathbb{R}^n)$

In this section, we endow \mathbb{R}^d with a Riemannian structure, given by a metric tensor $G(x)$, and a positive measure μ that is absolutely continuous with respect to the

Lebesgue measure, with density of the form $\mu(dx) = \exp(-H(x))dx$. We shall always assume that H is smooth, has a Hessian uniformly bounded from below, and grows at least quadratically at infinity (these are not optimal assumptions, but are more than sufficient for our purpose). We also consider the functional on the space of probability measures

$$\text{Ent}_\mu(\nu) := \int f \log f d\mu \quad (1.1)$$

when $d\nu = fd\mu$, and that takes value $+\infty$ for probability measures that are not absolutely continuous with respect to μ . Note that although we call this functional the entropy, it is the *negative* of the physical entropy. When $\mu = \exp(-H)$, this functional can be written as

$$\text{Ent}_\mu(\nu) = \text{Ent}_{dx}(\nu) + \int H(x)\nu(dx)$$

where

$$\text{Ent}_{dx}(\nu) = \int f e^{-H} \log(f e^{-H}) dx$$

is the relative entropy with respect to the Lebesgue measure, also known as the Boltzman entropy.

We can endow the space of probability measures with finite second moments with the Wasserstein distance associated to the Riemannian metric structure $W_{2,G}$.

In the sequel, we will consider curves $(\nu_t)_{t \in [0, T]}$ in $\mathcal{P}_2(\mathbb{R}^d)$ which are absolutely continuous, that is there exists a nonnegative function $\ell \in L^1([0, T])$ (which depends on $(\nu_t)_{t \in [0, T]}$) such that for any $s \leq t$ we have

$$W_{2,G}(\nu_s, \nu_t) \leq \int_s^t \ell(r) dr. \quad (1.2)$$

We will require the following technical assumptions on the metric tensor G (which are taken from [24]):

$$\frac{1}{c}|\xi|^2 \leq \langle G(x)\xi, \xi \rangle \leq c|\xi|^2, \forall x \in \mathbb{R}^n, \forall \xi \in \mathbb{R}^d \quad (1.3)$$

for some constant c , and

$$x \longrightarrow \langle G(x)\xi, \xi \rangle \quad \text{is lower semicontinuous } \forall \xi \in \mathbb{R}^d. \quad (1.4)$$

Since we now have a metric structure on $\mathcal{P}_2(\mathbb{R}^d)$, we can define the metric derivative of an absolutely continuous curve (ν_t) as

$$|\dot{\nu}|(t) := \limsup_{h \rightarrow 0^+} \frac{1}{h} W_{2,G}(\nu_t, \nu_{t+h}). \quad (1.5)$$

Definition 1.1. We denote by $\|h\|_\nu$ the H^1 norm of the smooth function h , defined by

$$\|h\|_\nu^2 := \int \langle A \nabla h, \nabla h \rangle d\nu \quad (1.6)$$

where $A(x)$ is the inverse of the matrix $G(x)$, and $\|\rho\|_{\nu,*}$ its dual norm, given by

$$\|\rho\|_{\nu,*}^2 := \sup_h 2 \int h \rho d\nu - \|h\|_\nu^2, \quad (1.7)$$

where the supremum runs over all smooth functions $h : \mathbb{R}^d \rightarrow \mathbb{R}$.

Let g be defined by

$$g(\nu) := \left(\int \left\langle \frac{A\nabla\nu}{\nu} + A\nabla H, \frac{\nabla\nu}{\nu} + \nabla H \right\rangle d\nu \right)^{1/2} \quad (1.8)$$

if ν is absolutely continuous with respect to the Lebesgue measure, and $+\infty$ if not. By $\nabla\nu$, we mean the gradient of the density of ν with respect to the Lebesgue measure. Recall that A is the inverse of G , and H is the opposite of the logarithm of the density of the reference measure $\mu(dx) = \exp(-H(x))dx$. This functional is known as the Fisher information with respect to μ , and will later also appear as the entropy production for the Fokker-Planck equation.

The following result explains how the functional g can be used to control the variation in relative entropy for absolutely continuous curves in the space of probability measures. Its proof in this context can be found in [24] (which generalizes previous results of [3]).

Proposition 1.2. *g is an upper gradient for Ent_μ , i.e. for every absolutely continuous curve $(\nu_t)_{0 \leq t \leq T}$ we have*

$$|\text{Ent}_\mu(\nu_t) - \text{Ent}_\mu(\nu_s)| \leq \int_s^t g(\nu_r) |\dot{\nu}|(r) dr$$

for every $0 \leq s \leq t \leq T$.

The setting of [24] only concerns curves satisfying a continuity equation $\dot{\nu}_t = \text{div}(Av_t\nu_t)$ in the weak sense, for some vector field v_t , but it is also proven in Theorem 2.4 of [24] that such a vector field automatically exists for the absolutely continuous curves we shall consider here.

Definition 1.3. *Let $(\nu_t)_{t \in [0, T]}$ be a time-dependent family of measures that is absolutely continuous. We say it is a gradient flow of the functional Ent_μ if*

$$\frac{d}{dt} \text{Ent}_\mu(\nu_t) = -\frac{1}{2}g(\nu_t)^2 - \frac{1}{2}|\dot{\nu}|(t)^2 \quad (1.9)$$

for almost every $t \in [0, T]$.

Gradient flows for the Wasserstein structure on $\mathcal{P}(\mathbb{R}^d)$ endowed with an Euclidean structure have been studied in [3], and it turns out they are related to the heat equation. Their results were then generalized to the case of a Riemannian structure in [24]:

Proposition 1.4. *$(\nu_t)_{t \in [0, T]}$ satisfies (1.9) iff the densities $f(t, \cdot) = \frac{d\nu_t}{d\mu}$ form a weak solution of the parabolic PDE*

$$\frac{\partial(f\mu)}{\partial t} = \text{div}(A(\nabla f)\mu), \quad (1.10)$$

where $A = G^{-1}$.

As a consequence of this result and Ito's formula, we also have a representation of a gradient flow as the flow of laws of the solution to a SDE:

Proposition 1.5. *If $(\nu_t)_t$ is a gradient flow of Ent_μ with $\mu = \exp(-H)dx$, then it is the flow of marginals of the law of a solution to the stochastic differential equation*

$$dX_t = -A(X_t)\nabla H(X_t)dt + \text{div}(A)(X_t)dt + \sqrt{2A(X_t)} dB_t \quad (1.11)$$

with initial condition X_0 that has law ν_0 . Here B_t is a standard Brownian motion on \mathbb{R}^d .

We refer the reader unfamiliar with stochastic differential equations and stochastic calculus to the textbook [36] for more information about these notions.

Remark 1.1. Diffusion processes that can be written in the form (1.11) are necessarily reversible.

We shall now give the definition of a key functional, which allows us to characterize gradient flows:

Proposition 1.6. *Let*

$$J((\nu_t)_t) := \text{Ent}_\mu(\nu_T) - \text{Ent}_\mu(\nu_0) + \frac{1}{2} \int_0^T g(\nu_t)^2 + |\dot{\nu}|(t)^2 dt. \quad (1.12)$$

Then $(\nu_t)_{t \in [0, T]}$ is a gradient flow of the functional Ent_μ iff $J((\nu_t)_t) = 0$.

Remark 1.2. In this setting, we have the following alternate formulation for the functional J :

$$J((\nu_t)_t) = \frac{1}{2} \int_0^T \|\dot{\nu} - \nabla \cdot (A(\nabla f)\mu)\|_{\nu, *}^2 dt, \quad (1.13)$$

at least for smooth functions. This formulation may seem more convenient, but in this context, (1.12) will be easier to manipulate.

Remark 1.3. We can make an interpretation of the notion of gradient flows in a statistical physics framework. It is a well-known principle in equilibrium statistical physics that steady states can be identified as minimizers of a thermodynamic functional, such as free energy, as a consequence of the second principle of thermodynamics. Here it is the relative entropy Ent_μ which plays the role of free energy (at least in isothermal situations, for non-isothermal situations it is the Massieu potential, see [28]) and indeed its minimizer is the equilibrium state μ . The gradient flow formulation identifies the correct trajectory as the minimizer of some action functional. This can be seen as an extension of the minimization principle to non-equilibrium statistical physics, with correct trajectories being those that decrease the free energy as fast as possible.

1.2 Large deviations

In this section, we introduce the notion of large deviations. This will only constitute a brief presentation of the topic, and we refer to the textbooks [11] and [15] for more information. The reader already familiar with this notion can easily skip this part. Lets start with the definition of large deviations:

Definition 1.7. *Let I be a lower semicontinuous, nonnegative function on a Polish space X and (a_n) a sequence of increasing, positive real numbers that goes to infinity. A sequence of probability measures \mathbb{P}_n on X is said to satisfy a large deviation principle with speed (a_n) and good rate function I iff*

- (i) *For any closed set F , $\limsup a_n^{-1} \log \mathbb{P}_n(F) \leq -\inf_{x \in F} I(x)$;*
- (ii) *For any open set O , $\liminf a_n^{-1} \log \mathbb{P}_n(O) \geq -\inf_{x \in O} I(x)$.*

Informally, we can think of a large deviations principle as making rigorous the statement that when n is large

$$\mathbb{P}_n(X_n \approx x) \approx \exp(-a_n I(x)).$$

In particular, if the rate function I cancels at a single point $x^* \in X$, the large deviations principle implies weak convergence of the probability measures to the Dirac mass δ_{x^*} . In such situations, the large deviations principle therefore *quantifies* the likelihood of observing a significant deviation from the typical asymptotic behavior of a sequence of random variables with laws \mathbb{P}_n , in the exponential scale.

There are many situations in which large deviations principles have been established. For example, Sanov's theorem states that empirical averages of n independent, identically distributed random variables satisfy a large deviations principle with speed n and rate function the relative entropy with respect to the law of a single random variable (see [11]).

A notion related to large deviations, and which we shall need in the sequel to state our results, is that of exponential tightness:

Definition 1.8 (Exponential tightness). *A sequence of probability measures (μ_n) on a topological space X is said to be exponentially tight with speed (a_n) if, for any $\alpha > 0$, there exists a compact set K_α such that*

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} \log \mu_n(K_\alpha^c) \leq -\alpha.$$

1.3 Main result

In this section, we shall state our main abstract result, relating the notions of gradient flows and of large deviations.

We consider a sequence of diffusion processes of the form (1.11). The parameters we allow to vary are the drift ∇H , the diffusion coefficient A , and the dimension of the underlying space d_n . We shall denote by Q_n the law of a trajectory on the time interval $[0, T]$ of the diffusion process

$$dX_t^n = -A_n(X_t^n) \nabla H_n(X_t^n) dt + \operatorname{div}(A_n)(X_t^n) + \sqrt{2A_n(X_t^n)} dB_t^n \quad (1.14)$$

on \mathbb{R}^{d_n} , with initial condition X_0^n distributed according to some law $Q_{0,n}$.

To be able to state a large deviation principle for the laws of these diffusion processes, we need to embed their trajectories into a single space. We therefore implicitly assume that all the spaces \mathbb{R}^{d_n} have been embedded into a single metric space X . In practice, the choice of the embedding will be suggested by the particular problem that is considered. We then endow the space $\mathcal{C}([0, T], X)$ with some topology that makes it a metric, separable space. A typical choice would be the supremum norm. Note that the choice of the topology affects the statement of the large deviations principle, since the open and closed sets for which the estimates are valid are specified by the topology we use.

To state the result, we must first define the notion of Γ -convergence of functionals:

Definition 1.9 (Gamma convergence). *Let X be a space endowed with a notion of convergence. A sequence (I_n) of functionals on X is said to Γ -converge to a functional I at point $x \in X$ if the two following conditions are met:*

(i) *For any sequence (x_n) that converges to x , we have $\liminf_{n \rightarrow \infty} I_n(x_n) \geq I(x)$;*

(ii) *There exists a sequence (x_n) that converges to x such that $\lim_{n \rightarrow \infty} I_n(x_n) = I(x)$.*

The sequence (I_n) is said to Γ -converge to I if it Γ -converges to I at every point.

We can now state the main result, relating large deviations and Γ -convergence of functionals of the form (1.12):

Theorem 1.10. *Let Q_n be the law of a trajectory on $[0, T]$ of a solution to a stochastic differential equation of the form (1.14), with $Q_{0,n}$ the law of the initial condition, and let J_n be the functional involved in the gradient flow formulation of the flow of marginals of (1.12), associated to the same SDE.*

Assume that the sequence $(Q_n)_{n \in \mathbb{N}}$ is exponentially tight, with speed (a_n) . Then it satisfies a large deviations principle with speed (a_n) and rate function I iff the functionals $(\nu_t)_{t \in [0, T]} \longrightarrow \frac{1}{a_n} (H(\nu_0; Q_{0,n}) + \frac{1}{2} J_n((\nu_t)))$ Γ -converge to I at the Dirac measure $\delta_{(x_t)}$, for every continuous trajectory $t \rightarrow x_t$.

Strictly speaking, the convergence used here is not a full Γ -convergence, but only Γ -convergence at every Dirac mass. For convenience, we shall still call this a Γ -convergence.

We expect the rate function I to be of the form $I(x_t) = I_0(x_0) + J(x_t)$, with J the function involved in the formulation (0.1) of a gradient flow in a certain metric space. This comes from the fact that often the rate function I will have a unique minimizer, which will be the deterministic limit of our sequence of processes, and in practice is defined as the solution to some PDE. So we can reformulate the actual limit as the unique minimizer of a function. But gradient flow formulations also characterize the solution to a PDE as the unique minimizer of a functional. The similarity between these two points of views makes us expect that they will be related, and this turns out to be the case in all the examples we are currently aware of. The idea that large deviations rate functionals often characterize a gradient flow structure has been explored in [31].

In the contributions [1], [2], [6] and [31], it is shown that the rate function for the large deviations principle for the empirical measure of independent copies of a diffusion

process on a small time interval (given by Sanov's theorem) asymptotically behaves (in the sense of Gamma-convergence) like the function J giving the gradient flow formulation of the law of a single copy of the diffusion process. These works show the strong connection between gradient flows and large deviations, and explain why some PDEs such as Fokker-Planck equations have a gradient flow structure. However they only concern the case of independent processes, and cannot be used to *prove* the validity of a large deviations principle, which is our main aim here.

An important element of the study of the functional J is the study of the Fisher information, or entropy-production functional. The importance of this functional can be understood in terms of statistical physics. It is a well-known principle in equilibrium statistical physics that equilibrium states can be obtained by optimizing some thermodynamic quantity, such as the free energy. This principle sometimes carries through to *non-equilibrium* statistical physics. Since the system seeks to increase the physical entropy (and therefore decrease the mathematical entropy), we can look at the entropy production functional, which we seek to optimize. Gamma-convergence corresponds to convergence of minimizers, so we can expect the most likely trajectories to be those that, in the limit, make the entropy production functional as small as possible.

Remark 1.4. For the lower bound in the Γ -convergence, it is enough to check the existence of a recovery sequence for every point y in a subset Y of X , such that given $x \in X$, there exists a sequence (y_k) of elements of Y that converges to x , and such that $I(y_k)$ converges to $I(x)$.

1.4 Some questions

- Is there a similar phenomenon for the large deviations of discrete systems, such as interacting particle systems? In the recent papers [26] and [29], Maas and Mielke independently showed that any reversible Markov chain on a finite space can be written as a gradient flow of a relative entropy for a well-chosen Riemannian structure on the space of probability measures. Can we exploit this structure to get the large deviations of systems such as a zero-range process, or exclusion processes?
- We apply in Section 3 this principle to get the large deviations for a system of diffusions with nearest neighbor interaction. It would also be interesting to look at mean-field models, where each diffusion interacts with all the others. In the case of smooth mean-field interactions, the question has been solved in [7], with a method that is very similar to the one we use here. A natural question is whether this extends to singular interactions. A case of particular interest would be that of Coulomb interactions.
- Another natural question is whether we can use this principle in a context of modeling. Say we wish to approximate a phenomenon characterized as the unique solution to a partial differential equation of the form $\partial_t \rho = H(\rho, \nabla \rho, \dots)$ with a system of N interacting diffusions, with N large. If we can find a sequence of diffusion processes on \mathbb{R}^N and a sequence of positive numbers a_N such that $\frac{1}{a_N} J_N$ Gamma-converges to a lower semicontinuous functional that

has the solution to the PDE as sole minimizer, then these diffusion processes form a good approximation. Can this idea be exploited in this context? This would be particularly interesting if we can extend our results to sequences of interacting particle systems.

- Our method works for reversible diffusion processes. Is there a similar method that works for non reversible processes, such as interacting diffusion processes with a boundary condition, or second-order diffusion processes? A possible candidate for a framework in which to work is the GENERIC framework studied in [9].

2 Proof of Theorem 1.10

The proof of Theorem 1.10 is essentially based on the following two arguments:

- The validity of a large deviations principle of a sequence of probability measures $(Q_n)_{n \in \mathbb{N}}$ is equivalent to Gamma-convergence of the relative entropy with respect to Q_n to the rate function. This is a result due to Mariani, which we shall explain in more details in Section 2.1.
- We can relate relative entropy with respect to Q_n and the functional J_n , in such a way that Gamma-convergence of the sequences are equivalent. More precisely, we have the following result:

Theorem 2.1. *Let Q be the law of a solution to a SDE of type (1.11) on a space \mathbb{R}^d , and P the law of a process with finite relative entropy with respect to Q , with flow of marginals (ν_t) . Then:*

(i) *We have the lower bound*

$$H(P; Q) \geq H(P_0; Q_0) + \frac{1}{2}J((\nu_t))$$

where P_0 and Q_0 are the laws of the initial conditions.

(ii) *There exists a process with law \tilde{P} that has the same flow of marginals as P , such that*

$$H(\tilde{P}; Q) = H(P_0; Q_0) + \frac{1}{2}J((\nu_t)).$$

This result implies that studying the Gamma-convergence of $a_n^{-1}H(\cdot; Q_n)$ and that of $a_n^{-1}(H(\cdot, Q_{n,0}) + J(\cdot)/2)$ is equivalent. An application of Theorem 2.3 then yields Theorem 1.10.

Of course, studying relative entropy to understand large deviations is a known technique (see [7] and [16]). Our contribution is to connect relative entropy to the framework of gradient flows in spaces of probability measures. Instead of studying relative entropy (which depends on the law of the whole trajectory), we can study the functional J , which only depends on the flow of marginals, and is easier to manipulate, at least in some cases of interest, due to its connection with optimal transport. It should be noted that some of the ideas we use here (relative entropy, variational formulations for rate functions) are reminiscent of those used in [7] to study large deviations for weakly-interacting mean-field models.

2.1 Relative entropy and large deviations

We would like to start by restating the two (equivalent) definitions of relative entropy:

Definition 2.2 (Relative entropy). *Given two probability measures P and Q on a Polish space X , the relative entropy of P with respect to Q is given by*

$$H(P; Q) := \sup_{f \in C_b(X)} \mathbb{E}_P(f) - \log \mathbb{E}_Q(e^f).$$

Equivalently, we have

$$H(P; Q) = \mathbb{E}_P \left[\ln \left(\frac{dP}{dQ} \right) \right]$$

if P is absolutely continuous with respect to Q , and $H(P; Q) = +\infty$ if not.

This equivalence is well-known, and is obtained by a computation of the Legendre transform (see Lemma 6.2.13 in [11]).

Mariani's result relating entropy and large deviations can be stated as follows :

Theorem 2.3 ([27], 2012). *Let (μ_n) be a sequence of probability measures on a Polish space X , (a_n) a sequence of positive real numbers such that $\lim_{n \rightarrow \infty} a_n = +\infty$ and $I : X \rightarrow [0, +\infty]$ a measurable, lower semicontinuous functional. We endow the space of probability measures with the topology of weak convergence.*

(i) The sequence (μ_n) satisfies a large deviations upper bound with speed (a_n) and rate function I iff it is exponentially tight with speed (a_n) and if for any sequence (ν_n) of probability measures on X that weakly converges to a Dirac measure δ_x , we have

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} H(\nu_n; \mu_n) \geq I(x);$$

(ii) The sequence (μ_n) satisfies a large deviations upper bound with speed (a_n) and rate function I iff for any point x , there exists a sequence (ν_n) of probability measures on X that weakly converges to δ_x such that

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} H(\nu_n; \mu_n) \leq I(x).$$

A heuristic explanation of Theorem 2.3 can be made in terms of the Bryc-Varadhan theorem (see Sections 4.3 and 4.4 in [11]). One can relate the relative entropy functional and exponential moments of functions by the relation

$$\int \exp(f) d\mu = \sup_{\nu \in \mathcal{P}(X)} \int f d\nu - \text{Ent}_\mu(\nu).$$

Since the Bryc-Varadhan lemma states that we can understand the large deviations for sequences of measures by looking at $\frac{1}{a_n} \log \int \exp(a_n f) d\mu_n$ for bounded continuous functions, the above relation translates the problem to investigating the behavior of the sequence of relative entropy functionals.

2.2 Proof of Theorem 2.1

The proof is a generalization of ideas coming from Part II of [16]. It will consist in three steps : first we shall use Girsanov's theorem to give a representation of the law of processes that are absolutely continuous with respect to the law of the gradient flow. In a second step, we shall give a representation of the relative entropy of such a process, and finally we shall use this representation to obtain the lower bound of our theorem.

The following result is a direct application of Girsanov's theorem (see for example Theorem 2.4 in [25]):

Proposition 2.4. *Let Q be the law of the solution of (1.11) on $[0, T]$, and P the law of a process that is absolutely continuous with respect to Q . Then there exists an adapted process (b_t) valued in \mathbb{R}^d such that*

$$\frac{dP}{dQ}((X_t)_{0 \leq t \leq T}) = 1_{\frac{dP}{dQ} > 0} \frac{dP_0}{dQ_0}(X_0) \exp \left(\int_0^T b_t \cdot \sqrt{2A(X_t)} dB_t^P + \int_0^T \langle A(X_t) b_t, b_t \rangle dt \right). \quad (2.1)$$

In this equation, B^P is a P -Brownian motion, that is a local martingale under P which P -almost surely has quadratic variation equal to t . Moreover, P can be viewed as the law of a solution to the SDE

$$dX_t = A(X_t)(2b_t - \nabla H(X_t))dt + \operatorname{div}(A)(X_t)dt + \sqrt{2A(X_t)}dB_t^P \quad (2.2)$$

As a consequence, the relative entropy is given by

$$H(P; Q) = H(P_0; Q_0) + \mathbb{E}_P \left[\int_0^T \langle A(X_t) b_t, b_t \rangle dt \right]. \quad (2.3)$$

Note that the relative entropy does not only depend on the flow of marginals (P_t) , but on the law of the whole trajectory, unlike the functional J .

Lemma 2.5 (Markov version of the process). *Let $b_t = b(t, (X_s)_{0 \leq s \leq t})$ be the adapted process associated to a law P . Define*

$$\tilde{b}_t(x) := \mathbb{E}_P[b(t, (X_s)_{0 \leq s \leq t}) | X_t = x]. \quad (2.4)$$

Then the process defined by

$$dX_t = A(X_t)(2\tilde{b}_t(X_t) - \nabla H(X_t))dt + \operatorname{div}(A)(X_t)dt + \sqrt{2A(X_t)}dB_t \quad (2.5)$$

with initial condition $X_0 \sim P_0$ is a Markov process, and its law has the same flow of marginals as P .

Proof. The fact that this process is a Markov process is a classic result on SDEs, so we shall concentrate on proving that both processes have same marginals. Let g be a smooth function, X_t be a solution of (2.2) and \tilde{X}_t a solution of (2.5). Using Ito's

formula, we have

$$\begin{aligned}
\mathbb{E}[g(\tilde{X}_t)] &= \int_0^t \mathbb{E} \left[A(\tilde{X}_s) \nabla g(\tilde{X}_s) \cdot (2\tilde{b}_s(X_s) - \nabla H(X_s)) \right] ds \\
&\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(\tilde{X}_s) \nabla g(\tilde{X}_s)) \right] ds \\
&= \int_0^t \mathbb{E} \left[A(\tilde{X}_s) \nabla g(\tilde{X}_s) \cdot (2\mathbb{E}_P[b(s, (X_r)_{0 \leq r \leq s}) | X_s] - \nabla H(X_s)) \right] ds \\
&\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(\tilde{X}_s) \nabla g(\tilde{X}_s)) \right] ds
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[g(X_t)] &= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2b(s, (X_r)_{0 \leq r \leq s}) - \nabla H(X_s)) \right] ds \\
&\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds \\
&= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2\mathbb{E}_P[b(s, (X_r)_{0 \leq r \leq s}) | X_s] - \nabla H(X_s)) \right] ds \\
&\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds \\
&= \int_0^t \mathbb{E} \left[A(X_s) \nabla g(X_s) \cdot (2\tilde{b}(X_s) - \nabla H(X_s)) \right] ds \\
&\quad + \int_0^t \mathbb{E} \left[\operatorname{div}(A(X_s) \nabla g(X_s)) \right] ds.
\end{aligned}$$

This shows that the marginals of the laws of X and \tilde{X} satisfy the same parabolic PDE

$$\frac{\partial f}{\partial t} = \operatorname{div}(A \nabla f) + \operatorname{div}(A(2\tilde{b} - \nabla H)f).$$

Since they have the same initial condition, and since solutions to such PDEs are unique, they are the same. \square

Lemma 2.6. *Let \tilde{P} be the law of the solution of (2.4).*

We have

$$H(P; Q) \geq H(\tilde{P}; Q) \tag{2.6}$$

Proof. We already know that

$$H(\tilde{P}; Q) = H(P_0; Q_0) + \mathbb{E}_{\tilde{P}} \left[\int_0^T \langle A(X_t) \tilde{b}_t, \tilde{b}_t \rangle dt \right]. \tag{2.7}$$

An application of Jensen's inequality and the definition of \tilde{b} yields

$$\begin{aligned}
H(P; Q) - H(P_0; Q_0) &= \mathbb{E}_P \left[\int_0^T \langle A(X_t) b_t, b_t \rangle dt \right] \\
&= \int_0^T \mathbb{E}_P [\langle A(X_t) b_t, b_t \rangle] dt \\
&= \int_0^T \mathbb{E}_{P_t} [\mathbb{E}_P [\langle A(X_t) b_t, b_t \rangle | X_t]] dt \\
&\geq \int_0^T \mathbb{E}_{P_t} [\langle A(X_t) \tilde{b}(X_t), \tilde{b}(X_t) \rangle] dt \\
&= H(\tilde{P}; Q) - H(P_0; Q_0), \tag{2.8}
\end{aligned}$$

which is the desired lower bound. \square

Lemma 2.7 (Entropy of the Markov process). *The entropy of the Markov version of the process satisfies*

$$H(\tilde{P}; Q) = H(\nu_0; Q_0) + \frac{1}{2} J((\nu_t))$$

where (ν_t) is the flow of marginals of the process P .

Proof. Let g be a smooth, compactly supported function. Itô's formula applied to the SDE (2.5) yields

$$\begin{aligned}
\mathbb{E}[g(X_t)] &= \mathbb{E}[g(X_0)] + \int_0^t \mathbb{E}[A \nabla g(X_s) \cdot (2\tilde{b}_s(X_s) - \nabla H(X_s))] ds \\
&\quad + \int_0^t \mathbb{E}[(\nabla \cdot A(X_s) \nabla g(X_s))] ds. \tag{2.9}
\end{aligned}$$

It is easy to deduce from the Ito formulation (2.9) that the flow of marginals (ν_t) solves (in a weak sense) the PDE

$$\dot{\nu}_t = -\operatorname{div} \left(2A\tilde{b}_t \nu_t - \nu_t A \nabla H - A \nabla \nu_t \right). \tag{2.10}$$

Therefore the variation of the entropy of the marginals is given by

$$\begin{aligned}
\operatorname{Ent}_\mu(\nu_T) - \operatorname{Ent}_\mu(\nu_0) &= \int_0^T \int A \nabla \nu_t(x) \cdot (2\tilde{b}_t(x) - \nabla H(x)) dx dt \\
&\quad - \int_0^T \int \frac{\langle A(x) \nabla \nu_t(x), \nabla \nu_t(x) \rangle}{\nu_t(x)} dx dt \\
&\quad + \int_0^T \int A(2\tilde{b}_t - \nabla H) \cdot \nabla H d\nu_t dt - \int_0^T \int A \nabla H \cdot \nabla \nu_t dx dt \\
&= \int_0^T \int 2A\tilde{b}_t \cdot \nabla \nu_t + \nu_t \nabla H dx dt - \int \frac{A(\nabla \nu_t + \nu_t \nabla H) \cdot (\nabla \nu_t + \nu_t \nabla H)}{\nu_t} dx \tag{2.11}
\end{aligned}$$

From the Benamou-Brenier formula for $W_{2,G}$ (see for example [40]):

$$W_{2,G}^2(\nu_0, \nu_1) = \inf \left\{ \int_0^1 \int \langle Gv, v \rangle d\nu_t dt; \quad \dot{\nu} + \operatorname{div}(v\nu_t) = 0 \right\}$$

we deduce

$$\begin{aligned} & \frac{1}{2} \int_0^T |\dot{\nu}_t|^2 dt \\ &= \frac{1}{2} \int_0^T \left\langle A \left(2\tilde{b}_t - \nabla H - \frac{\nabla \nu_t}{\nu_t} \right), \left(2\tilde{b}_t - \nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt \\ &= 2 \int_0^T \int \langle A\tilde{b}_t, \tilde{b}_t \rangle d\nu_t dt + \frac{1}{2} \int_0^T \int \left\langle A \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right), \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt \\ &\quad - \int_0^T \int \left\langle 2A\tilde{b}_t, \left(\nabla H + \frac{\nabla \nu_t}{\nu_t} \right) \right\rangle d\nu_t dt. \end{aligned} \tag{2.12}$$

By the definition (1.8) of the upper gradient g , we have

$$\frac{1}{2} \int_0^T g(\nu_t)^2 dt = \frac{1}{2} \int_0^T \int \left\langle \frac{A\nabla \nu_t}{\nu_t} + A\nabla H, \frac{\nabla \nu_t}{\nu_t} + \nabla H \right\rangle d\nu_t dt. \tag{2.13}$$

Summing (2.11), (2.12) and (2.13), we get

$$J((\nu_t)) = 2 \int_0^T \int \langle A\tilde{b}_t, \tilde{b}_t \rangle d\nu_t dt, \tag{2.14}$$

and then the lemma immediately follows from (2.7). \square

The combination of Lemmas 2.6 and 2.7 then immediately yields Theorem 2.1. To deduce Corollary 1.10 from Theorem 2.1, the only thing we still have to prove is that, if (ν_t) is a flow of marginals such that $J((\nu_t))$ is finite, there exists a process whose law is absolutely continuous with respect to Q , and with flow of marginals (ν_t) . Let (ν_t) be an absolutely continuous flow of marginals such that $J((\nu_t))$ is finite. From Theorem 2.4 in [24], we know that there exists a vector field (v_t) such that the continuity equation

$$\dot{\nu}_t = \nabla \cdot (Av_t \nu_t) \tag{2.15}$$

is satisfied. On the other hand, we know that (ν_t) is the flow of marginals of the solution to an SDE of type

$$dX_t = 2A(X_t)b_t(X_t)dt + \sqrt{2A(X_t)}dB_t,$$

whose law would then be absolutely continuous with respect to Q , if the flow solves in a weak sense the PDE

$$\dot{\nu}_t = \operatorname{div}(A(\nabla \nu_t + 2b_t \nu_t)).$$

Since $J((\nu_t))$ is finite, the upper gradient $g(\nu_t)$ is finite for almost every t , and $\nabla \nu_t$ exists. We therefore only have to take $2b_t(x) = v_t(x) - \frac{\nabla \nu_t}{\nu_t}(x)$ to see that the flow given by (2.15) solves the above PDE.

3 Large deviations for the Ginzburg-Landau model

3.1 The model

The (classical) Ginzburg-Landau model equipped with Kawasaki dynamics is a sequence of N diffusions, interacting according to the SDE

$$dX_t^i = N^2(\psi'(X_t^{i+1}) + \psi'(X_t^{i-1}) - 2\psi'(X_t^i))dt + \sqrt{2}N(dB_t^{i+1} - dB_t^i).$$

This model can be viewed as an overdamped limit of chains of interacting anharmonic oscillators. It can also be seen as a continuous-spin version of the Ising model endowed with Kawasaki dynamics. Beyond its relevance as a physical model, it is also widely used in the litterature as a prototype of dynamics governed by differential operators, on which we can test techniques to study large-scale behavior of interacting systems. We refer to [19], [42] and the references contained therein for more information on the history of the model.

The drift is given by the derivative of a potential $\psi : \mathbb{R} \rightarrow \mathbb{R}$, which we assume to be C^2 . This diffusion is not ergodic on the whole space \mathbb{R}^N , since it preserves the mean spin $m = N^{-1} \sum X_t^i$, but it is ergodic when restricted to a hyperplane

$$X_{N,m} := \left\{ x \in \mathbb{R}^N; \quad \frac{1}{N} \sum_{i=1}^N x_i = m \right\}. \quad (3.1)$$

It then has an invariant measure

$$\mu_{N,m}(dx) := \frac{1}{Z} \exp\left(-\frac{1}{N} \sum_{i=1}^N \psi(x_i)\right) \mathbb{1}_{x \in X_{N,m}} \mathcal{L}^{N-1}(dx), \quad (3.2)$$

where \mathcal{L}^{N-1} is the Lebesgue measure on the $(N-1)$ -dimensional hyperplane $X_{N,m}$. Equivalently, its law is given by the solution of the PDE

$$\frac{\partial f \mu_{N,m}}{\partial t} = \operatorname{div}(A_0 \nabla f \mu_{N,m}), \quad (3.3)$$

where f is the density with respect to $\mu_{N,m}$, and A_0 is the discrete Laplacian scaled by N^2 , that is

$$(A_0)_{i,j} := N^2(\delta_{i,j+1} + \delta_{i,j-1} - 2\delta_{i,j}).$$

To embed all the spaces $X_{N,m}$ in a single space, and conveniently formulate limit results, we associate to an element $x = (x_1, \dots, x_N) \in X_{N,m}$ a step function on the torus \bar{x} , given by

$$\bar{x}(\theta) = x_i \quad \forall \theta \in \left(\frac{i-1}{N}, \frac{i}{N} \right].$$

Note that $\int \bar{x}(\theta) d\theta = m$. Through this procedure, we can embed all the spaces $X_{N,m}$ in the space of functions on the torus, which we endow with the usual H^{-1} norm, and which we denote by $H^{-1}(\mathbb{T})$.

The notion of convergence we shall use in the sequel is the following

Definition 3.1. Let ν_N be a probability measure on $X_{N,m}$. A sequence (ν_N) is said to weakly converge to a deterministic profile $\rho \in L^2(\mathbb{T})$ if, for all smooth test functions $J : \mathbb{T} \rightarrow \mathbb{R}$, we have

$$\int_{X_{N,m}} \int_{\mathbb{T}} \bar{x}(\theta) J(\theta) d\theta \nu_N(dx) \longrightarrow \int_{\mathbb{T}} \rho(\theta) J(\theta) d\theta.$$

If we denote by \bar{x}_N a random step function on the torus with law ν_N , this notion of convergence corresponds to convergence in law of the sequence \bar{x}_N to ρ , in the space $H^{-1}(\mathbb{T})$.

When N goes to infinity, if the initial condition behaves deterministically in the limit, the (properly rescaled) solutions at time t weakly converge to a deterministic profile $\rho(t, \cdot) \in L^2(\mathbb{T})$, called the hydrodynamic limit, which has been studied in [19]. Large deviations from this hydrodynamic limit have been studied in [10]. The scaling limit is given by the PDE

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\rho),$$

where φ is the Cramer transform of ψ , given by

$$\varphi(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \log \int \exp(\sigma x - \psi(x)) dx \right).$$

For technical reasons, we shall assume that the initial data follows a local Gibbs state, that is

$$f_0(x) = \frac{1}{Z} \exp \left(\sum x_i \varphi'(\rho_0(i/N)) \right) \quad (3.4)$$

for some continuous function ρ_0 . We will later see that this initial data concentrates around the deterministic profile ρ_0 . It can be shown that, for initial data that behaves deterministically in the limit, solutions at any positive time are close (in the sense of relative entropy) to such a local Gibbs state. See [21] or [12] for a proof.

We will also assume that the single-site potential ψ is of the form

$$\psi(x) = \frac{1}{p} x^p + \delta\psi(x) \quad (3.5)$$

for some $p \geq 2$ and a perturbation $\delta\psi$ that is C^2 , bounded and with bounded first and second derivative.

It is likely that the results we shall obtain hold for more general functions ψ , but such a result would require more general technical tools than those developed in the next section. For example, if ψ does not grow at least as $|x|^2$ as x goes to infinity, then the logarithmic Sobolev inequality does not hold. In [10], the LDP is proved for the case where ψ is only superquadratic, and $\psi' = o(\psi)$.

Although the method can be used to reprove the result of [10], rather than work on the classical Ginzburg-Landau model we just described, we shall focus on two variants of the model, with the extra input of a non-constant conductance. The first model will involve random conductances, and the second will be the model of [35], with conductances that depend on the configuration of spins. Our method will rely on

Theorem 1.10, and reduce the problem to the study of the behavior of the functional J_N associated with the gradient flow formulation of these dynamics.

There are several technical elements of our approach that are in common with the one of [10], such as tightness estimates and gradient replacement estimates. In particular, it turns out that the proof of the upper bound in the convergence of J_N is the same as the proof of [35] of the lower bound for the LDP. We shall therefore only sketch the proofs of the upper bounds, and concentrate on the lower bounds in the Γ -convergence.

3.2 Some technical estimates

In this section, we give a few technical results, collected from various sources, which we shall use in the proofs of the large deviation principles. Most of them are classical results in the study of hydrodynamic limits, and we will often only give a brief sketch of the proofs, or simply refer to the original source.

We will use logarithmic Sobolev inequalities, which we now define:

Definition 3.2 (Logarithmic Sobolev inequality). *Let X be a Riemannian manifold. A probability measure μ on X is said to satisfy a logarithmic Sobolev inequality (which we shall abbreviate into LSI) with constant $\rho > 0$ if, for any locally Lipschitz, nonnegative function $f \in L^1(\mu)$,*

$$\int f \log(f) d\mu - \left(\int f d\mu \right) \log \left(\int f d\mu \right) \leq \frac{1}{\rho} \int \frac{|\nabla f|^2}{2f} d\mu.$$

The following result was proven in [30]:

Theorem 3.3 ([30]). *Under the assumption (3.5), the measures $\mu_{N,m}$ defined by (3.2) satisfy the logarithmic Sobolev inequality*

$$\text{Ent}_{\mu_{N,m}}(g) \leq C \int \frac{|\nabla g|^2}{g} d\mu_N$$

for any nonnegative, locally Lipschitz function g , with constant C independent of the dimension N and the mean spin m . This implies the inequality

$$\text{Ent}_{\mu_{N,m}}(g) \leq C \int \frac{\langle A_0 \nabla g, \nabla g \rangle}{g} d\mu_N$$

for some constant C that is independent of the dimension and the mean spin. As previously, A_0 is the (scaled) discrete Laplacian.

As a consequence of this result and of Lemma 26 in [18], we have the following result:

Lemma 3.4. *Let f_N be a sequence of probability densities with respect to $\mu_{N,m}$ such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f_N, \nabla f_N \rangle}{f} d\mu_{N,m} < +\infty.$$

Then we also have

$$\sup_N \frac{1}{N} \int \sum |x_i|^2 f_N(x) \mu_{N,m}(dx) < +\infty.$$

This result still holds if we replace $\mu_{N,m}$ by another sequence of measures with bounded second moment and which satisfy a LSI with uniform constant.

This can be generalized to the following result:

Lemma 3.5. *Assume that ψ is of the form $\frac{1}{p}|x|^p + \delta\psi(x)$. Let f_N be a sequence of probability densities with respect to $\mu_{N,m}$ such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f_N, \nabla f_N \rangle}{f} d\mu_{N,m} < +\infty.$$

Then we also have

$$\sup_N \frac{1}{N} \int \sum |x_i|^p f_N(x) \mu_{N,m}(dx) < +\infty.$$

Proof. It has been shown in [13] that, under our assumptions on ψ , $\mu_{N,m}$ satisfies the following transport-entropy inequality: for any probability measure ν_N ,

$$W_p^p(\nu_N, \mu_N) \leq C \text{Ent}_{\mu_{N,m}}(\nu_N)$$

for some constant $C > 0$ that does not depend on N , and W_p is the L^p Wasserstein distance

$$W_p^p(\nu, \mu) := \inf_{\pi} \int \sum |x_i - y_i|^p \pi(dx, dy).$$

From the W_p -Lipschitz continuity of p-moments (see Proposition 7.29 in [41]), we know that

$$\left(\int \sum |x_i|^p \nu_N(dx) \right)^{1/p} - \left(\int \sum |x_i|^p \mu_{N,m}(dx) \right)^{1/p} \leq W_p(\mu_{N,m}, \nu_N)$$

so that

$$\int \sum |x_i|^p \nu_N(dx) \leq C \text{Ent}_{\mu_{N,m}}(\nu_N) + C \int \sum |x_i|^p \mu_{N,m}(dx)$$

Since $\mu_{N,m}$ satisfies a logarithmic Sobolev inequality, $\text{Ent}_{\mu_{N,m}}(\nu_N) \leq CN$, and it is also easy to see that

$$\int \sum |x_i|^p \mu_{N,m}(dx) \leq CN,$$

which concludes the proof. \square

We now give a version the version of the local Cramér theorem we shall use:

Theorem 3.6. *Let (a_i) be some sequence of real numbers. We define*

$$\psi_K(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp\left(\sum a_i x_i + \psi(x_i)\right) dx$$

and

$$\varphi_K(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \sum_{i=1}^K \log \int_{\mathbb{R}} \exp((\sigma + a_i)x - \psi(x)) dx \right).$$

We then have, for any $L > 0$ and any compact subset E of \mathbb{R} ,

$$\lim_{K \rightarrow \infty} \sup_{a_1 \dots a_K \in [-L, L]} \|\psi_K - \varphi_K\|_{\infty, E} = 0.$$

In particular, if $a_i = \lambda(i/K)$ for some smooth function λ , then ψ_K converges to

$$\varphi_\lambda(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \int_0^1 \log \int_{\mathbb{R}} \exp((\sigma + \lambda(\theta))x - \psi(x)) dx d\theta \right).$$

uniformly on compact sets.

A proof of this result can be found in [14] or Appendix A in [21]. Roughly speaking, it says that local averages of a large number K of spins behave like random variables satisfy a large deviation principle of speed K and rate function φ .

The following proposition is a consequence of Theorem 4.1 in [19].

Proposition 3.7. *Let f_N be a sequence of probability densities with respect to μ_N which weakly converges to a deterministic profile ρ . Assume that*

$$\frac{1}{N} \int \frac{\langle A_0 \nabla f_N, \nabla f_N \rangle}{f_N} d\mu_{N,m} \leq C.$$

Then, for any smooth function $J : \mathbb{T} \rightarrow \mathbb{R}$ and bounded continuous function $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$, we have

$$\frac{1}{N} \int \sum J(i/N) F(x_{i-k}, \dots, x_{i+k}) f_N(x) \mu_N(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \tilde{F}(\rho(\theta)) d\theta$$

where

$$\tilde{F}(y) := \int F(x_1, \dots, x_{2k+1}) \mu^{\lambda, \otimes 2k+1}(dx), \quad (3.6)$$

with $\mu^\lambda(dx) = \frac{1}{Z} \exp(\lambda x - \psi(x)) dx$ and $\lambda = \varphi'(y)$.

Note that, in [19], it was also required that there exists a superlinear function ω such that $\int \sum \omega(x_i) f(x) \mu_N(dx) \leq CN$. However, under our assumptions on ψ , the bound on $\int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_N$ implies that $\int \sum |x_i|^2 f(x) \mu_N(dx) \leq CN$, as we have seen in Lemma 3.4.

Proposition 3.8. *Let ρ be a smooth function on the torus, and define the probability density with respect to μ*

$$G_N(x) = \frac{1}{Z} \exp \left(\sum_{i=1}^N \varphi'(\rho(i/N)) x_i \right).$$

Then

- (i) *The measures $G_N \mu_{N,m}$ weakly converge to the deterministic profile ρ ;*
- (ii) *They satisfy a logarithmic Sobolev inequality, with a constant that only depends on ρ and ψ , but which is uniform in N ;*

(iii) For any sequence of probability measures ν_N on \mathbb{R}^N , if

$$\frac{1}{N} \text{Ent}_{G_N \mu_{N,m}}(\nu_N) \longrightarrow 0,$$

then the sequence weakly converges to the deterministic profile ρ . Moreover, we then have

$$\frac{1}{N} \text{Ent}_{\mu_{N,m}}(\nu_N) \longrightarrow \int \varphi(\rho) d\theta - \varphi \left(\int \rho d\theta \right).$$

Proof. (i) is a classic hydrodynamic limit result. See for example [42]. (ii) was proven in [14]. (iii) is a consequence of these two results, and we can prove it as follows.

Since the measures $G_N \mu_N$ satisfy a logarithmic Sobolev inequality, they also satisfy a transport entropy inequality, that is

$$W_{2,A_0^{-1}}(\nu_N, G_N \mu_{N,m})^2 \leq C \text{Ent}_{G_N \mu_{N,m}}(\nu_N).$$

The fact that we can use the inner product given by A_0 rather than the usual inner product follows from the discrete Poincaré inequality. Therefore, we have

$$\frac{1}{N} W_{2,A_0^{-1}}(\nu_N, G_N \mu_{N,m})^2 \longrightarrow 0.$$

The result then follows from the fact that $(G_N \mu_{N,m})$ weakly converges to ρ , and that $\frac{1}{N} \langle A_0 x, x \rangle \leq C \|\bar{x}\|_{H^{-1}}^2$.

The second part is a consequence of the identity

$$\frac{1}{N} \text{Ent}_{\mu_{N,m}}(\nu_N) = \frac{1}{N} \text{Ent}_{G_N \mu_{N,m}}(\nu_N) + \frac{1}{N} \int \log G_N d\nu_N$$

and the convergence

$$\begin{aligned} \frac{1}{N} \int \log G_N d\nu_N &= \frac{1}{N} \int \sum_{i=1}^N \varphi'(\rho_i) x_i \nu_N(dx) \\ &- \frac{1}{N} \log \int \exp \left(\sum_{i=1}^N \varphi'(\rho(i/N)) x_i - \psi(x_i) \right) dx + \frac{1}{N} \log \int \exp \left(\sum_{i=1}^N -\psi(x_i) \right) dx \\ &\longrightarrow \int \varphi(\rho(\theta)) d\theta - \varphi \left(\int \rho(\theta) d\theta \right). \end{aligned} \quad (3.7)$$

This last convergence follows from the convergence of (ν_N) to the deterministic profile ρ and Theorem 3.6. A complete proof is given in Lemma 7.1 of [21]. \square

Proposition 3.9. *Let f_N be a sequence of probability densities with respect to $\mu_{N,m} N$ which weakly converges to a deterministic profile ρ . Assume that*

$$\frac{1}{N} \int \frac{\langle A_0 \nabla f, \nabla f \rangle}{f} d\mu_{N,m} \leq C$$

and that

$$\frac{1}{N} \int \sum |x_i|^2 f_N(x) \mu_{N,m}(dx) \leq C.$$

Then, for any sequence (J^N) of step functions on the torus that are constant on the intervals $((i-1)/N, i/N]$ and which converges in H^1 to a function J , we have

$$\frac{1}{N} \int \sum J^N(i/N) x_i f_N \mu_{N,m}(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \rho(\theta) d\theta \quad (3.8)$$

and

$$\frac{1}{N} \int \sum J^N(i/N) \psi'(x_i) f_N \mu_{N,m}(dx) \longrightarrow \int_{\mathbb{T}} J(\theta) \varphi'(\rho(\theta)) d\theta. \quad (3.9)$$

Proof. For the first part, notice that

$$\begin{aligned} & \left| \frac{1}{N} \int \sum J^N(i/N) x_i f_N \mu_{N,m}(dx) - \int \int_{\mathbb{T}} J(\theta) \bar{x}(\theta) d\theta f_N(x) \mu_{N,m}(dx) \right| \\ & \leq \|J^N - J\|_{H^1} \left(\int \|\bar{x}\|_{H^{-1}}^2 f_N(x) \mu_{N,m}(dx) \right)^{1/2} \longrightarrow 0 \end{aligned}$$

and

$$\begin{aligned} & \left| \int \int_{\mathbb{T}} J(\theta) \bar{x}(\theta) d\theta f_N(x) \mu_{N,m}(dx) - \int_{\mathbb{T}} J(\theta) \rho(\theta) d\theta \right| \\ & \leq \|J\|_{H^1} \left(\int \|\bar{x} - \rho\|_{H^{-1}} f_N(x) \mu_{N,m}(dx) \right) \\ & \leq \|J\|_{H^1} \left(\int \|\bar{x} - \rho\|_{L^2} f_N(x) \mu_{N,m}(dx) \right) \end{aligned}$$

so we just have to show that

$$\int \|\bar{x} - \rho\|_{L^2} f_N(x) \mu_{N,m}(dx) \longrightarrow 0.$$

This quantity is the Wasserstein distance W_1 between $f_N \mu_N$ and δ_ρ for the L^2 distance. Since we already have weak convergence, to show that there is convergence for W_1 , according to Theorem 7.12 of [40], we just have to prove the following tightness estimate

$$\lim_{R \rightarrow \infty} \limsup_{N \rightarrow \infty} \int_{\sum |x_i|^2 \geq NR^2} \sqrt{\frac{1}{N} \sum |x_i|^2} f_N(x) \mu_{N,m}(dx) = 0.$$

This estimate automatically follows from the bound

$$\sup_N \frac{1}{N} \int \sum |x_i|^2 f_N(x) \mu_{N,m}(dx) < +\infty$$

that was given by Lemma 3.4.

For the second part, we give a very brief sketch of the method of proof that was used in [19]. Let ψ_ℓ be a cutoff of ψ' at level $\ell > 0$, that is

$$\psi_\ell(x) = \psi'(x) \text{ if } |\psi'(x)| \leq \ell, \psi_\ell(x) = \pm \ell \text{ if not.}$$

From the bound of Lemma 3.5, we can deduce

$$\frac{1}{N} \int \sum J^N(i/N) \psi'_\ell(x_i) f_N \mu_{N,m}(dx) \lim_{\ell \rightarrow \infty} \frac{1}{N} \int \sum J^N(i/N) \psi'(x_i) f_N \mu_{N,m}(dx) \quad (3.10)$$

uniformly in N , since ψ goes to infinity faster than $|\psi'|$. Moreover, from Proposition 3.7, we obtain

$$\frac{1}{N} \int \sum J^N(i/N) \psi'_\ell(x_i) f_N \mu_{N,m}(dx) \lim_{N \rightarrow \infty} \int_{\mathbb{T}} J(\theta) \tilde{\psi}_\ell(\rho(\theta)) d\theta, \quad (3.11)$$

so all we need to do is show that $\tilde{\psi}_\ell$ converges to φ' , which was done in Lemma 6.4 of [19]. \square

Remark 3.1. Similarly, under the assumption that $\int_0^T \int \frac{\langle A_0 \nabla f_N, \nabla f_N \rangle}{f_N} f_N \mu_{N,m} dt \leq CN$ uniformly in N , then (3.8) and (3.9) hold in a time-integrated sense.

We also give a priori estimates on weak limits of sequences of probability measures, obtained as direct consequences of [19], Lemmas 6.3 and 6.6:

Lemma 3.10. *Under our assumptions on ψ , for any sequence of probabilities with density f_N with respect to $\mu_{N,m}$ that weakly converges to a deterministic trajectory ρ , such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla f_N, \nabla f_N \rangle}{f} d\mu_{N,m} \leq C$$

we have

$$\int_{\mathbb{T}} \varphi(\rho(\theta)) d\theta \leq C$$

and

$$\int_{\mathbb{T}} (\partial_\theta \varphi(\rho)(\theta))^2 d\theta \leq C.$$

Finally, we shall need the following lower bound on the slope of absolutely continuous curves.

Lemma 3.11. *Let (ν_t) be an absolutely continuous curve of probability measures on \mathbb{R}^n , which is equipped with a Riemannian tensor $(A^{-1}(x))$ satisfying the assumptions (1.3) and (1.4). Then we have, for any smooth function $V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$,*

$$\begin{aligned} \int |\dot{\nu}_t|^2 dt &\geq 2 \int V(T, x) \nu_T(dx) - 2 \int V(0, x) \nu_0(dx) \\ &\quad - 2 \int_0^T \int \frac{\partial V}{\partial t}(t, x) \nu_t(dx) dt - \int_0^T \int \langle A(x) \nabla V, \nabla V \rangle \nu_t(dx) dt \end{aligned} \quad (3.12)$$

Proof. From Theorem 2.4 of [24], we know that there exists a vector field v_t such that

$$\dot{\nu}_t + \operatorname{div}(v_t \nu_t) = 0 \quad (3.13)$$

and

$$\int_0^T |\dot{\nu}_t|^2 dt = \int_0^T \int \langle A^{-1}(x) v_t(x), v_t(x) \rangle \nu_t(dx) dt. \quad (3.14)$$

Since we have

$$\langle A^{-1}(x)v_t(x), v_t(x) \rangle \geq 2\langle v_t(x), \nabla V(t, x) \rangle - \langle A(x)\nabla V(t, x), \nabla V(t, x) \rangle$$

for any t and x , we get

$$\begin{aligned} \int_0^T |\dot{v}_t|^2 dt &\geq 2 \int_0^T \int \langle v_t(x), \nabla V(t, x) \rangle \nu_t(dx) dt \\ &\quad - \int_0^T \int \langle A(x)\nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt \end{aligned}$$

Using (3.13) to do an integration by parts on the first term, the result immediately follows. \square

3.3 Large deviations for the GL model in a random environment

In this section, we shall be interested in the large deviations for a version of the process (3.3) in a random environment, where the operator A is replaced by a realization of the symmetric random matrix

$$A_{i,j}(\omega) := N^2 a_{i+1}(\omega)(\delta_{i,j-1} - \delta_{i,j}) - N^2 a_i(\omega)(\delta_{i,j} - \delta_{i,j+1}) \quad (3.15)$$

where the a_i are iid random variables defined on a probability space Ω , and we assume there exists a constant $c > 0$ such that we almost surely have

$$c \geq a_i \geq 1/c. \quad (3.16)$$

This assumption corresponds to an ellipticity assumption on the operator A that is uniform in the realization of the random field. Therefore, for any x and any realization of the random field, we have

$$\frac{1}{c} \langle A_0 x, x \rangle \leq \langle A x, x \rangle \leq c \langle A_0 x, x \rangle. \quad (3.17)$$

Under these assumptions, the quantity

$$\bar{a} := \mathbb{E} \left[\frac{1}{a_1} \right] \quad (3.18)$$

is well defined and finite.

The associated SDE is

$$dX_t^i = N^2 (a_i(\psi(X_t^{i+1}) - \psi(X_t^i)) - a_{i-1}(\psi(X_t^i) - \psi(X_t^{i-1}))) + \sqrt{2a_i} dB_t^i - \sqrt{2a_{i-1}} dB_t^{i-1}. \quad (3.19)$$

Given a realization of the random environment, we denote by $L_{a,N}$ the generator of this diffusion.

The following hydrodynamic limit result for the random environment model has been proven in [17]:

Theorem 3.12 ([17]). *Assume that the sequence of initial data $f_{N,0}\mu_{N,m}$ weakly converges to a deterministic profile $\rho_0 \in H^1(\mathbb{T})$. Then, for any time $t > 0$ the sequence $f_{N,t}\mu_{N,m}$ giving the distributions of the systems at time t weakly converges to a deterministic profile $\rho(t, \theta)$, which is given by the unique solution of the PDE*

$$\frac{\partial \rho}{\partial t} = \bar{a} \Delta \varphi'(\rho)$$

with initial condition ρ_0 . This convergence holds for almost every realization of the random field.

We are interested in the following quenched large deviation principle for this model, using the gradient flow approach we developed in section 1.

Theorem 3.13. *Assume that the sequence of initial data is of the form (3.4) for some smooth initial profile m_0 . Then, for almost every realization of the random field, the sequence of laws of solutions to (3.19) on $[0, T]$ satisfies a LDP in $L_t^\infty(H_\theta^{-1})$ with speed N and rate function*

$$I(\rho) := \int \varphi(\rho(0, \theta)) - \varphi(m_0(\theta)) - \varphi'(m_0(\theta))(\rho(\theta) - m_0(\theta)) d\theta \\ + \frac{1}{4\bar{a}} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \bar{a} \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) \right\|_{H^{-1}}^2 dt.$$

This generalizes the large deviations principle for the classical Ginzburg-Landau model of [10] to the case of a random environment.

In terms of gradient flows, this result follows from two facts :

- The relative entropy with respect to the invariant measure, divided by N , Γ -converges to $\rho \mapsto \int \varphi(\rho) - \varphi(\int \rho)$. This corresponds to a large deviations principle for the sequence of invariant measures $\mu_{N,m}$;
- The sequence of metrics given by $A^{-1}(w)$ almost surely converges to the H^{-1} norm, divided by a factor \bar{a} .

As a technical tool, we shall need the following convergence result, which will be used to formalize the convergence of the discrete norms.

Lemma 3.14. *Let $(a_q)_{q \in \mathbb{Q}}$ be a sequence of positive, bounded, iid random variables, and let $a_i^N := a_{i/N}$. With probability one, for any sequence (h^N) of step functions on \mathbb{T} that converges to a function h in L^1 , such that h^N is constant on $(\frac{i-1}{N}, \frac{i}{N}]$, and denoting by h_i^N the value of h^N on such an interval, we have*

$$\lim \frac{1}{N} \sum_{i=1}^N a_i^N h_i^N = \mathbb{E}(a) \int h(\theta) d\theta.$$

Proof. Let M be an integer. The strong law of large numbers implies that, with probability 1, for any step function h that is constant on the intervals $(\frac{i-1}{M}, \frac{i}{M}]$, we have

$$\lim \frac{1}{N} \sum_{i=1}^N a_i^N h(i/N) = \mathbb{E}(a) \int h(\theta) d\theta.$$

This then remains true simultaneously for every integer M , still with probability 1. An approximation argument in $L^1(\mathbb{T})$ then yields the desired result. \square

Proof of Theorem 3.13. Given a realization of our random environment, the functional J_N is given by

$$J_N(\nu_t) = \text{Ent}_{\mu_{N,m}}(\nu_T) - \text{Ent}_{\mu_{N,m}}(\nu_0) + \frac{1}{2} \int_0^T |\dot{\nu}_t|_{A^{-1}}^2 dt + \frac{1}{2} \int_0^T \int \frac{\langle A \nabla g_t, \nabla g_t \rangle}{g_t} d\mu_{N,m} dt, \quad (3.20)$$

where g_t is the density of ν_t with respect to $\mu_{N,m}$. We add the subscript $|\cdot|_{A^{-1}}$ to emphasize the fact that we used the Euclidean norm associated to A^{-1} as the quadratic cost for the Wasserstein distance.

Lemma 3.15. *The functional $\frac{1}{N} \text{Ent}_{\mu_N}$ Γ -converges at every Dirac mass to $\rho \rightarrow \int \varphi(\rho) d\theta - \varphi(\int \rho d\theta)$.*

Proof. Let ν_N be a sequence that weakly converges to a deterministic profile $\rho : \mathbb{T} \rightarrow \mathbb{R}$, and let

$$G_N(x) := \frac{1}{Z_N} \exp \left(\sum_{i=1}^N \varphi'(\rho_{i,N}) x_i \right) \quad (3.21)$$

be a local Gibbs profile with respect to $\mu_{N,m}$, where $\rho_{i,N} = \int_{(i-1)/N}^{i/N} \rho(\theta) d\theta$ and Z_N is the normalization constant such that $G_N \mu_{N,m}$ is a probability measure.

We then have the decomposition

$$\text{Ent}_{G_N \mu_{N,m}}(\nu_N) = \text{Ent}_{\mu_{N,m}}(\nu_N) - \int \log G_N d\nu_N. \quad (3.22)$$

By definition of G_N , since ν_N weakly converges to the deterministic profile ρ , as we have seen in the proof of Proposition 3.8, $\frac{1}{N} \int \log G_N d\nu_N$ converges to $\int \varphi(\rho) d\theta - \varphi(\int \rho d\theta)$. Since the relative entropy is nonnegative, we can deduce from (3.22) the inequality

$$\liminf \frac{1}{N} \text{Ent}_{\mu_{N,m}}(\nu_N) \geq \int \varphi(\rho) d\theta - \varphi\left(\int \rho d\theta\right).$$

Moreover, the measures $G_N \mu_{N,m}$ weakly converge to the deterministic profile ρ , so that they provide the recovery sequence for this Gamma-convergence result. \square

Lemma 3.16. *Let (ν_N) be a sequence of probability measures that converges to a profile ρ , such that, for any N , ν_N is absolutely continuous with respect to $f_{N,0} \mu_{N,m}$. Then*

$$\frac{1}{N} \text{Ent}_{f_{N,0} \mu_{N,m}}(\nu_N) - \frac{1}{N} \text{Ent}_{\mu_{N,m}}(\nu_N) \longrightarrow \int_{\mathbb{T}} \varphi'(m_0)(m_0 - \rho) - \varphi(m_0) d\theta.$$

Proof. Since ν_N is absolutely continuous with respect to $f_{N,0}\mu_{N,m}$ (and therefore to $\mu_{N,m}$), we have

$$\begin{aligned}
& \frac{1}{N} \text{Ent}_{f_{N,0}\mu_{N,m}}(\nu_N) - \frac{1}{N} \text{Ent}_{\mu_{N,m}}(\nu_N) \\
&= -\frac{1}{N} \int \log f_{N,0} d\nu_N \\
&= -\frac{1}{N} \int \sum \varphi'(m_0(i/N))x_i \nu_N(dx) + \frac{1}{N} \log \int \exp\left(\sum \varphi'(m_0(i/N))x_i\right) \mu_{N,m}(dx) \\
&= -\frac{1}{N} \int \sum \varphi'(m_0(i/N))x_i \nu_N(dx) + \frac{1}{N} \log \int \exp\left(\sum \varphi'(m_0(i/N))x_i - \psi(x_i)\right) dx \\
&\quad - \frac{1}{N} \log \int \exp\left(\sum -\psi(x_i)\right) dx \\
&\rightarrow -\int \varphi'(m_0(\theta))\rho(\theta)d\theta + \int \varphi'(m_0(\theta))m_0(\theta)d\theta \\
&\quad - \int \varphi(m_0(\theta))d\theta
\end{aligned}$$

since (ν_N) has asymptotic profile ρ , and applying Theorem 3.6. \square

We will now investigate the behavior of the slope:

Lemma 3.17 (Lower bound for the time-derivative). *For any sequence $(\nu_{N,t})$ weakly converging to some ρ and such that*

$$\sup_N \frac{1}{N} J_N((\nu_{N,t})) < +\infty,$$

we have

$$\liminf \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{A^{-1}}^2 dt \geq \frac{1}{a} \int_0^T \|\partial\rho/\partial t\|_{H^{-1}}^2 dt.$$

Proof. Let $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ be a smooth function. Applying Lemma 3.11 with

$$J_N(t, x) := \frac{1}{N} \sum_i J(t, i/N)x_i + \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(t, j/N)b_j x_i \text{ and } b_i = \frac{\bar{a}}{a_i} - 1, \text{ we have}$$

$$\begin{aligned}
& \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{A^{-1}}^2 dt \geq \frac{2}{N} \int \sum J(T, i/N)x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(T, j/N)b_j x_i \nu_{N,T}(dx) \\
&\quad - \frac{2}{N} \int \sum J(0, i/N)x_i \nu_{N,T}(dx) - \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(0, j/N)b_j x_i \nu_{N,0}(dx) \\
&\quad - \frac{2}{N} \int_0^T \int \sum \frac{\partial J}{\partial t}(t, i/N)x_i \nu_{N,t}(dx) dt - \frac{2}{N^2} \int_0^T \int \sum_i \sum_{j=1}^{i-1} \frac{\partial^2 J}{\partial t \partial \theta}(t, j/N)b_j x_i \nu_{N,t}(dx) dt \\
&\quad - \frac{1}{N} \int_0^T \sum a_i \left(NJ\left(t, \frac{i+1}{N}\right) - NJ\left(t, \frac{i}{N}\right) + b_i \frac{\partial J}{\partial \theta}\left(t, \frac{i}{N}\right) \right)^2 dt.
\end{aligned}$$

Taking the limit $N \rightarrow +\infty$ and using the second-moment bounds of Lemma 3.4, we get

$$\begin{aligned} & \frac{2}{N} \int \sum J(T, i/N) x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(T, j/N) b_j x_i \nu_{N,T}(dx) \\ &= \frac{2}{N} \int \sum J(T, i/N) x_i \nu_{N,T}(dx) + O\left(\frac{1}{N}\right) \\ &\rightarrow \int_{\mathbb{T}} J(T, \theta) \rho(T, \theta) d\theta. \end{aligned}$$

In the same way,

$$\begin{aligned} & \frac{2}{N} \int \sum J(0, i/N) x_i \nu_{N,T}(dx) + \frac{2}{N^2} \int \sum_i \sum_{j=1}^{i-1} \frac{\partial J}{\partial \theta}(0, j/N) b_j x_i \nu_{N,0}(dx) \\ &\rightarrow \int_{\mathbb{T}} J(0, \theta) \rho(0, \theta) d\theta \end{aligned}$$

and

$$\begin{aligned} & \frac{2}{N} \int_0^T \int \sum \frac{\partial J}{\partial t}(t, i/N) x_i \nu_{N,t}(dx) dt + \frac{2}{N^2} \int_0^T \int \sum_i \sum_{j=1}^{i-1} \frac{\partial^2 J}{\partial t \partial \theta}(t, j/N) b_j x_i \nu_{N,t}(dx) dt \\ &\rightarrow \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt. \end{aligned}$$

Finally, using Lemma 3.14, we get

$$\begin{aligned} & \frac{1}{N} \int_0^T \sum a_i \left(N J(t, \frac{i+1}{N}) - N J(t, \frac{i}{N}) + b_i \frac{\partial J}{\partial \theta}(t, i/N) \right)^2 dt \\ &\rightarrow \int_0^T \int_{\mathbb{T}} \bar{a} \left(\frac{\partial J}{\partial \theta}(t, \theta) \right)^2 d\theta dt. \end{aligned}$$

Combining these lower bounds, we get

$$\begin{aligned} \liminf \frac{1}{N} \int_0^T |\dot{\nu}_{N,t}|_{\nu_{N,t}}^2 dt &\geq 2 \int J(T, \theta) \rho(T, \theta) d\theta - 2 \int_{\mathbb{T}} J(0, \theta) \rho(0, \theta) d\theta \\ &\quad - 2 \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt - \int_0^T \int_{\mathbb{T}} \bar{a} \left(\frac{\partial J}{\partial \theta}(t, \theta) \right)^2 d\theta dt. \end{aligned}$$

Taking the supremum over all smooth functions J yields the result. \square

Lemma 3.18 (Lower bound for the upper gradient). *For any time t and subsequence $(\nu_{N,t})$ with densities $g_{N,t}$ with respect to $\mu_{N,m}$ and such that*

$$\sup_N \frac{1}{N} \int \frac{\langle A_0 \nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_{N,m} < +\infty,$$

we almost surely have

$$\liminf_N \frac{1}{N} \int \frac{\langle A \nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_{N,m} \geq \bar{a} \int_{\mathbb{T}} (\partial_\theta \varphi'(\rho(t, \theta)))^2 d\theta.$$

Proof. From Lemma 3.10, we know that, under these assumptions, $\varphi'(\rho)$ lies in $H^1(\mathbb{T})$. Let $J : \mathbb{T} \rightarrow \mathbb{R}$ be a smooth function, and define

$$G_N(x) := \exp \left(\sum_{i=1}^N \left(\sum_{j=1}^i \frac{\bar{a}}{a_j} J(i/N) \right) x_i \right). \quad (3.23)$$

Since the upper gradient takes value $+\infty$ when ν_N is not absolutely continuous with respect to $\mu_{N,m}$, and we are looking for a lower bound, we can assume without loss of generality that ν_N is absolutely continuous with respect to $\mu_{N,m}$, and therefore to $G_N \mu_{N,m}$. Let $g_{N,t}$ be the density of $\nu_{N,t}$ with respect to $\mu_{N,m}$.

We consider the quantity $\int \frac{\langle A \nabla \left(\frac{g_N}{G_N} \right), \nabla \left(\frac{g_N}{G_N} \right) \rangle}{g_{N,t}/G_N} G_N d\mu_{N,m}$, which is nonnegative. We have

$$\begin{aligned} & \int \frac{\langle A \nabla \left(\frac{g_N}{G_N} \right), \nabla \left(\frac{g_N}{G_N} \right) \rangle}{g_{N,t}/G_N} G_N d\mu_{N,m} \\ &= \int \frac{\langle A \nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_{N,m} - 2 \int \frac{\langle A \nabla g_{N,t}, \nabla G_N \rangle}{G_N} d\mu_{N,m} \\ & \quad + \int \frac{\langle A \nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_{N,m} \\ &= \int \frac{\langle A \nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_{N,m} - 2 \int \frac{\langle A \nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_{N,m} \\ & \quad + \int \frac{\langle A \nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_{N,m}. \end{aligned}$$

Therefore, we have

$$\int \frac{\langle A \nabla g_{N,t}, \nabla g_{N,t} \rangle}{g_{N,t}} d\mu_{N,m} \geq 2 \int \frac{\langle A \nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_{N,m} - \int \frac{\langle A \nabla G_N, \nabla G_N \rangle}{G_N^2} d\mu_{N,m} \quad (3.24)$$

for any realization of the random field, any N and any t .

Applying Lemma 3.14, we have

$$\begin{aligned} \frac{1}{N} \int \frac{\langle A \nabla G_{N,t}, \nabla G_{N,t} \rangle}{G_{N,t}^2} d\mu_{N,m} &= \frac{1}{N} \sum_{i=1}^N \frac{\bar{a}^2}{a_i} J(i/N)^2 \\ &\rightarrow \bar{a} \int_{\mathbb{T}} J(\theta)^2 d\theta. \end{aligned} \quad (3.25)$$

We also have

$$\begin{aligned}
\frac{1}{N} \int \frac{\langle A \nabla H, \nabla G_N \rangle}{G_N} g_{N,t} d\mu_{N,m} &= \frac{\bar{a}}{N} \int \sum (\psi'(x_{i+1}) - \psi'(x_i)) J(i/N) g_N(x) \mu_{N,m}(dx) \\
&= \frac{\bar{a}}{N} \int \sum \psi'(x_i) (J((i-1)/N) - J(i/N)) g_N(x) \mu_{N,m}(dx) \\
&= \frac{\bar{a}}{N} \int \sum \psi'(x_i) J'(i/N) g_N(x) \mu_{N,m}(dx) \\
&\quad + O\left(\frac{1}{N^2} \int \sum |\psi'(x_i)| g_N(x) \mu_{N,m}(dx)\right) \\
&\longrightarrow \bar{a} \int_{\mathbb{T}} \varphi'(\rho(\theta)) J'(\theta) d\theta. \tag{3.26}
\end{aligned}$$

Combining these two lower bounds and taking the supremum over smooth functions J , we get the lower bound of our Lemma. \square

From the previous Lemma and Fatou's Lemma, we can then deduce that, for a sequence that converges to a Dirac mass, and such that $J_N(f_N) \leq CN$, we have

$$\liminf \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \geq \int_0^T \bar{a} \int_{\mathbb{T}} (\partial_\theta \varphi'(\rho(t, \theta)))^2 d\theta dt,$$

which was the last element we needed for the lower bound of the Γ -convergence.

We now turn to the recovery sequence. Given a profile $\rho(t, \theta)$ that is weakly continuous in time, and such that $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho) - \partial_t \rho\|_{H^{-1}}^2 dt$ is finite, there exists a sequence of smooth profiles ρ_k that converge to ρ , and such that $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho_k) - \partial_t \rho_k\|_{H^{-1}}^2 dt$ converges to $\int_0^T \|\bar{a} \partial_\theta^2 \varphi'(\rho) - \partial_t \rho\|_{H^{-1}}^2 dt$. Therefore, in view of Remark 1.4, we only have to prove the existence of a recovery sequence for profiles ρ that are smooth.

Given such a smooth profile ρ , there exists a continuous function $h : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ such that

$$\frac{\partial \rho}{\partial t} = \bar{a} \frac{\partial}{\partial \theta} \left(h(t, \theta) + \frac{\partial}{\partial \theta} \varphi'(\rho) \right). \tag{3.27}$$

We now consider a dynamic with law given by the time-dependent generator

$$\tilde{L}_{a,N} = L_{a,N} + N \sum_{i=1}^N \frac{\bar{a}}{a_i} h(t, i/N) \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) \tag{3.28}$$

and initial condition given by the local Gibbs state associated to $\rho(0, \cdot)$.

We need to prove two things: that the solutions to such dynamics converge to the deterministic profile ρ , and that $N^{-1} J_N((\nu_{N,t}))$ has the correct limit (where $\nu_{N,t}$ are the marginals of the law of the solution to the modified dynamic). The first part can be done in the same way as in Section 3 of [35], so we concentrate on the second part.

We have

$$\begin{aligned}
\frac{1}{N}J_N(f_N) &= \frac{1}{2} \int_0^T \left\| \frac{\partial f_N}{\partial t} - L_{a,N}f_N \right\|_{H^{-1}(A)}^2 dt \\
&= \frac{1}{2} \int_0^T \left\| \tilde{L}_{a,N}f_N - L_{a,N}f_N \right\|_{H^{-1}(A)}^2 dt \\
&= \frac{1}{2} \int_0^T \sum_{i=1}^N \frac{\bar{a}^2}{a_i} h(t, i/N)^2 dt \\
&\rightarrow \frac{1}{2} \int_0^T \int_{\mathbb{T}} \bar{a}h(t, \theta)^2 d\theta dt
\end{aligned} \tag{3.29}$$

and, using (3.27), it is easy to see that this is equal to $\frac{1}{2\bar{a}} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \bar{a} \frac{\partial^2}{\partial \theta^2} \varphi'(\rho) \right\|_{H^{-1}}^2 dt$, which was what we needed to prove.

We still have to prove exponential tightness for the laws of solutions to (3.19). It is given by the following two results:

Lemma 3.19. *Let \mathbb{P}_N be the law of a solution to the SDE (3.19) with initial condition X_0 having a distribution $f_0\mu_N$ that satisfies $\text{Ent}_{\mu_N}(f_0) \leq CN$. Then*

$$\lim_{\ell \rightarrow +\infty} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) = -\infty.$$

Lemma 3.20. *Under the same assumptions as the previous lemma, for any $\epsilon > 0$ and any smooth function on the torus J , we have*

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq s \leq t \leq T, |s-t| \leq \delta} \left| \frac{1}{N} \sum J(i/N)(X_t^i - X_s^i) \right| \geq \epsilon \right) = -\infty.$$

Proof of Lemma 3.19. This proof is exactly the same as in [10], we give a brief sketch to show that the random field (a_i) does not make any difference.

Let $\tilde{P}^{eq,N}$ be the law of a solution to the SDE starting from the equilibrium measure $\mu_{N,m}$. From Lemma 1.12 in [23], we know that, for any symmetric function g on \mathbb{R}^N , we have

$$\tilde{P}^{eq,N} \left(\sup_{0 \leq t \leq T} g(X_t) \geq \ell \right) \leq \frac{3}{\ell} \sqrt{a + Tb}$$

with $a = \int g^2 d\mu_{N,m}$ and $b = \int \frac{\langle A \nabla g, \nabla g \rangle}{g} d\mu_{N,m}$. When $g(x) = \exp(\sum |x_i|)$, there exists $C > 0$ such that $a \leq C^N$ and $b \leq N^2 C^N$. Using the Tchebychev inequality, we obtain

$$\tilde{P}^{eq,N} \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) \leq \sqrt{C^N(1 + N^2)} e^{-N\ell} \leq C' e^{-C''N\ell}.$$

We can then use the basic entropy inequality

$$\mathbb{P}_N(A) \leq \frac{\log(2) + H_N}{\log(1 + 1/\tilde{P}^{eq,N}(A))},$$

where H_N is the relative entropy of the non equilibrium process with respect to the equilibrium process, which satisfies the bound $H_N \leq CN$, to get the result. \square

In the same way, we refer the reader interested in the proof of Lemma 3.20 to Lemma 2.8 in [10]. The proof is exactly the same, since the random variables a_i are bounded. \square

3.4 Large deviations for the non-gradient Ginzburg-Landau model

We consider the SDE given by

$$dX_t^i = N^2(W_{i,i+1} - W_{i-1,i})dt + N\sqrt{2a(X_t^i, X_t^{i+1})}dB_t^{i+1} - N\sqrt{2a(X_t^{i-1}, X_t^i)}dB_t^i \quad (3.30)$$

where

$$W_{i,i+1} = a(X_t^i, X_t^{i+1})(\psi'(X_t^i) - \psi'(X_t^{i+1})) - \frac{\partial a}{\partial x}(X_t^i, X_t^{i+1}) + \frac{\partial a}{\partial y}(X_t^i, X_t^{i+1}). \quad (3.31)$$

This model is known as the non-gradient Ginzburg-Landau model. This terminology is unrelated to gradient flow structures (since this dynamics has a gradient flow structure), but is used to emphasize the fact that in this model, currents cannot be written as discrete gradients of some function of the spin configuration.

The marginals of the law of a solution (3.30) solve the PDE

$$\frac{\partial f_N \mu_{N,m}}{\partial t} = \nabla \cdot (A(x) \nabla f_N \mu_{N,m}) \quad (3.32)$$

where the matrix $A(x)$ is given by

$$A(x)_{i,j} := a(x_{i-1}, x_i)(\delta_{i,j} - \delta_{i,j+1}) + a(x_i, x_{i+1})(\delta_{i,j} - \delta_{i,j-1}).$$

Once more, we will assume that the initial data $f_{N,0}$ is of the form (3.4).

The generator of this dynamics is given by

$$L_N f = N^2 \sum e^H \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) e^{-H} a(x_i, x_{i+1}) \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) f \quad (3.33)$$

It has been shown in [39] that trajectories of such a dynamic concentrate around the solution to the PDE

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \frac{\partial}{\partial \theta} \varphi'(\rho) \right) \quad (3.34)$$

where φ is the same function as in the previous section, and \hat{a} is a bounded continuous function, which we shall now define.

Let $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ be a bounded, smooth function of a finite number of variables.

The function $\xi : x \in \mathbb{R}^\infty \rightarrow \sum_{i=-\infty}^{+\infty} F(x_{i-k}, \dots, x_{i+k})$ is not well-defined, but its partial derivatives are. We can therefore define

$$a_F(y) := \int a(x_0, x_1) \left(1 - \frac{\partial \xi}{\partial x_1} + \frac{\partial \xi}{\partial x_0} \right)^2 \mu^{\infty,y}(dx) \quad (3.35)$$

where $\mu^{\infty,y}$ is the product measure on \mathbb{R}^∞ with every one-dimensional marginal having density $Z^{-1} \exp(\lambda x - \psi(x))$, with λ the unique real number such that this measure has expectation y .

The function \hat{a} is then given by

$$\hat{a}(y) := \inf_F a_F(y) \quad (3.36)$$

with the infimum running over the set of all smooth, bounded functions of a finite number of variables.

In the next proposition, we summarize a few properties of the function \hat{a} :

Proposition 3.21. (i) \hat{a} is a bounded, continuous function.

(ii) For any $\epsilon > 0$ and $C < +\infty$, there exists a smooth real-valued function $g(x_k, \dots, x_k, y)$ on \mathbb{R}^{2k+2} with bounded first derivatives such that

$$\sup_{|y| \leq C} (a_{g(\cdot, y)}(y) - \hat{a}(y)) < \epsilon$$

and

$$\sup_{y \in \mathbb{R}} (a_{g(\cdot, y)}(y) - \hat{a}(y)) \leq \|a\|_\infty.$$

Part (i) of this Proposition comes from [39], and part (ii) from [35].

Our aim is to prove the following large deviations result:

Theorem 3.22. Assume that the sequence of initial data is of the form (3.4) for some smooth initial profile m_0 . The sequence of random functions satisfies a LDP in $L^\infty(H^{-1})$ with speed N and rate function

$$\begin{aligned} I(\rho) := & \int \varphi(\rho(0, \theta)) - \varphi(m_0(\theta)) - \varphi'(m_0(\theta))(\rho(\theta) - m_0(\theta)) d\theta \\ & + \frac{1}{4} \int_0^T \left\| \frac{\partial \rho}{\partial t} - \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \frac{\partial}{\partial \theta} \varphi'(\rho) \right) \right\|_{H^{-1}(\hat{a}(\rho(t, \cdot)))}^2 dt. \end{aligned}$$

In the rate function, the norm is defined as

$$\|u\|_{H^{-1}(\hat{a}(\rho(t, \cdot)))}^2 := \sup_{v \in H^1(\mathbb{T})} 2 \int_{\mathbb{T}} u(\theta) v(\theta) d\theta - \int \hat{a}(\rho(t, \theta)) \left(\frac{\partial v}{\partial \theta} \right)^2 d\theta. \quad (3.37)$$

This result was already proved in [35], under the assumption that the single site potential ψ is uniformly convex, and that its second derivative is bounded from above. Our assumptions are a priori more general, since they allow for superquadratic potentials, but it seems likely that, using the logarithmic Sobolev inequality proved in [30], the method of [35] could be extended for such functions. The main aim of this section is not to prove a new result, but to show that the ideas developed here can be used for non-gradient models.

The following result is the key technical estimate to prove large deviations for non-gradient models. It has been proven in [39].

Let R_N be the law of the stationary solution to the SDE (3.30), with the initial condition X_0 having law μ_N . Let $\rho(t, \theta)$ be a deterministic profile, which we assume to be in $L^\infty(H^1)$.

Denote by $\rho_\theta^c(t) := \int_{\theta-c}^{\theta+c} \rho(t, s) ds$. For a given smooth function $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ and a function g as in part (ii) of Proposition 3.21, we define

$$\begin{aligned} V(t) := & \sum_{i=1}^N J(t, i/N) \left[W_{i, i+1} - \frac{1}{N^2} Lg \left(X_t^{i-k}, \dots, X_t^{i+k}, \rho_{i/N}^{\ell/N}(t) \right) \right] \\ & + \frac{1}{N} \sum_{i=1}^N J(t, i/N) \hat{a}(\rho_{i/N}^{\epsilon_1}(t)) \left(\frac{\varphi'(\rho_{i/N+\epsilon_2}^{\epsilon_1}(t)) - \varphi'(\rho_{i/N-\epsilon_2}^{\epsilon_1}(t))}{2\epsilon_2} \right) \\ & - \frac{\alpha}{N} \sum_{i=1}^N J(t, i/N)^2 (a_{g(\cdot, \rho_{i/N}^{\epsilon_1}(t))}(\rho_{i/N}^{\epsilon_1}(t)) - \hat{a}(\rho_{i/N}^{\epsilon_1}(t))) \end{aligned} \quad (3.38)$$

where α , ϵ_1 and ϵ_2 are positive numbers, and ℓ is a positive integer. Under these notations, we have the following exponential estimate, which is due to Varadhan [39]:

Theorem 3.23. *For any profile ρ , any $\alpha > 0$, any J and g , we have*

$$\lim_{\epsilon_2 \rightarrow 0} \limsup_{\epsilon_1 \rightarrow 0} \limsup_{\ell \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}^{R_N} \left[\exp \left(\alpha N \int_0^T V(t) dt \right) \right] \leq 0.$$

Using the result, we obtain what is known in the hydrodynamic literature as the gradient replacement estimate.

Corollary 3.24. *Let $(\nu_{N,t})$ be a sequence of flows of time-marginals of a a sequence of processes that weakly converge to a deterministic flow $\rho(t, \theta)$. Assume moreover that $J_N(\nu_{N,t}) \leq CN$ for some $C > 0$.*

Then for any smooth, bounded functions $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^{2k+2} \rightarrow \mathbb{R}$, we have

$$\begin{aligned} & \lim_{\ell \rightarrow \infty} \lim_N \int_0^T \int \sum J(t, i/N) (W_{i, i+1}(x) - \frac{1}{N^2} (Lg)(x_{i-k}, \dots, x_{i+k}, \rho_{i/N}^{\ell/N})) \nu_{N,t}(dx) dt \\ & = \int_0^T \int_{\mathbb{T}} J(t, \theta) \hat{a}(\rho(t, \theta)) \partial_\theta \varphi'(\rho(t, \theta)) d\theta dt. \end{aligned}$$

Proof. First, let us show that we can build a diffusion process of law P with time-marginals $(\nu_{N,t})$ such that $H(P_N, R_N) = O(N)$. Since the equilibrium process is a solution of (3.30) with initial condition μ_N , we know that there exists a process with marginals $(\nu_{N,t})$ such that

$$\begin{aligned} H(P_N, R_N) &= \frac{1}{2} \text{Ent}_{\mu_N}(\nu_{0,N}) + \frac{1}{2} \text{Ent}_{\mu_N}(\nu_{N,T}) + \frac{1}{4} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\ &+ \frac{1}{4} \int_0^T |\dot{\nu}_t|_{A^{-1}}^2 dt \\ &= J_N(\nu_{N,t}) + \text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) \end{aligned}$$

so we only have to get an upper bound on $\text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0})$. Moreover, the bound on $J_N(\nu_{N,t})$ implies that $\text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) \leq CN$.

Denoting $\nu_{N,0} = \rho_N \mu_N$, we have

$$\begin{aligned}
\text{Ent}_{\mu_N}(\nu_{N,0}) - \text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) &= \int \rho \log f_{0,N} d\mu_N \\
&= \int \sum \varphi'(m_0(i/N)) x_i \nu_{N,0}(dx) - \log Z_N \\
&\leq CN \times \sqrt{\frac{1}{N} \int \sum |x_i|^2 \nu_{N,0}(dx)} + CN \\
&\leq CN
\end{aligned} \tag{3.39}$$

where the last bound follows from $\text{Ent}_{f_{0,N}\mu_N}(\nu_{N,0}) \leq CN$, Lemma 3.4 and the fact that the measures $f_{0,N}\mu_N$ satisfy a LSI with uniform constant, and have uniformly bounded second moments.

Consequently, using the entropy inequality, we have

$$\mathbb{E}^{P_N} \left[\int_0^T V(t) dt \right] \leq \frac{C}{\alpha} + \frac{1}{\alpha N} \log \mathbb{E}^{R_N} \left[\exp \left(\alpha N \int_0^T V(t) dt \right) \right] \tag{3.40}$$

and the result immediately follows from Theorem 3.23 and using the fact that the inequality is valid for both J and $-J$. \square

We also recall the exponential tightness estimates that have been proven in [39], and which we need to apply Corollary 1.10:

Lemma 3.25. *Let \mathbb{P}_N be the law of a solution to the SDE 3.30 with initial condition X_0 having a distribution $f_0 \mu_N$ that satisfies $\text{Ent}_{\mu_N}(f_0) \leq CN$. Then*

$$\lim_{\ell \rightarrow +\infty} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq t \leq T} \frac{1}{N} \sum |X_t^i| \geq \ell \right) = -\infty$$

and for any $\epsilon > 0$ and any smooth function on the torus J , we have

$$\lim_{\delta \rightarrow 0} \lim_{N \rightarrow +\infty} \frac{1}{N} \log \mathbb{P}_N \left(\sup_{0 \leq s \leq t \leq T, |s-t| \leq \delta} \left| \frac{1}{N} \sum J(i/N)(X_t^i - X_s^i) \right| \geq \epsilon \right) = -\infty.$$

The proof of these estimates is exactly the same as for the gradient case studied in [10]. Once more, the use of a variable function $a(x_i, x_{i+1})$ does not make a difference as long as it is bounded.

Lemmas 3.15 and 3.16 remain valid, so that, to prove Theorem 3.22, we only have to study the behavior of the slopes.

Lemma 3.26.

$$\liminf_N \frac{1}{N} \int_0^T |\dot{\nu}_t|_{A^{-1}}^2 dt \geq \int_0^T \|\partial_t \rho\|_{H^{-1}(\hat{a})}^2 dt.$$

Proof. Let $J : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ be a smooth function, and $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ be a smooth, bounded function. Applying Lemma 3.11 with $V(t, x) = \sum J(t, i/N) x_i +$

$\frac{1}{N}J'(t, i/N)F(x_{i-k}, \dots, x_{i+k})$, we get

$$\begin{aligned} \frac{1}{N} \int_0^T |\dot{v}_t|_{A^{-1}}^2 dt &\geq 2 \frac{1}{N} \int V(T, x) \nu_T(dx) - 2 \frac{1}{N} \int V(0, x) \nu_0(dx) \\ &\quad - 2 \frac{1}{N} \int_0^T \int \frac{\partial V}{\partial t}(x) \nu_t(dx) dt \\ &\quad - \frac{1}{N} \int_0^T \int \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt. \end{aligned} \quad (3.41)$$

We have

$$\begin{aligned} \frac{1}{N} \int V(T, x) \nu_T(dx) &= \frac{1}{N} \int J(T, i/N) x_i \nu_T(dx) + \frac{1}{N^2} \int J'(T, i/N) F(x_{i-k}, \dots, x_{i+k}) \nu_T(dx) \\ &= \frac{1}{N} \int J(T, i/N) x_i \nu_T(dx) + O\left(\frac{1}{N}\right) \\ &\longrightarrow \int J(T, \theta) \rho(T, \theta) d\theta. \end{aligned} \quad (3.42)$$

In the same way,

$$\frac{1}{N} \int V(0, x) \nu_0(dx) \longrightarrow \int J(0, \theta) \rho(0, \theta) d\theta \quad (3.43)$$

and

$$\frac{1}{N} \int_0^T \int \frac{\partial V}{\partial t}(x) \nu_t(dx) dt \longrightarrow \int_0^T \int_{\mathbb{T}} \frac{\partial J}{\partial t}(t, \theta) \rho(t, \theta) d\theta dt. \quad (3.44)$$

For the last term, we have

$$\begin{aligned} &\frac{1}{N} \int_0^T \int \langle A(x) \nabla V(t, x), \nabla V(t, x) \rangle \nu_t(dx) dt \\ &= \frac{1}{N} \int_0^T \int \sum_i N^2 a(x_i, x_{i+1}) \left(J(t, (i+1)/N) - J(t, i/N) \right. \\ &\quad \left. + \frac{1}{N} \sum_{j=i-k}^{i+k} \frac{\partial}{\partial x_{i+1}} F(x_{j-k}, \dots, x_{j+k}) - \frac{\partial}{\partial x_i} F(x_{j-k}, \dots, x_{j+k}) \right)^2 \nu_t(dx) dt \\ &= \frac{1}{N} \int_0^T \int \sum_i a(x_i, x_{i+1}) J' \left(t, \frac{k}{N} \right)^2 \\ &\quad \times \left(1 - \left(\frac{\partial}{\partial x_{i+1}} - \frac{\partial}{\partial x_i} \right) \sum_{j=i-k}^{i+k} F(x_{j-k}, \dots, x_{j+k}) + O\left(\frac{k}{N}\right) \right)^2 \nu_t(dx) dt \\ &\longrightarrow \int_0^T \int_{\mathbb{T}} a_F(\rho(t, \theta)) \left(\frac{\partial J}{\partial \theta} \right)^2 d\theta dt. \end{aligned} \quad (3.45)$$

We combine these convergence estimates, and then optimize in F and J to get the desired result. \square

Lemma 3.27. *For a flow of marginals $\nu_{N,t}$ that weakly converges to ρ , and such that $\frac{1}{N}J_N((\nu_{N,t})_t)$ is bounded, we have*

$$\liminf_N \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \geq \int_0^T \int_{\mathbb{T}} \hat{a}(\rho(t, \theta)) (\partial_\theta \varphi'(\rho))^2 d\theta dt.$$

Proof. Let $J(t, \theta)$ be a smooth function, $F : \mathbb{R}^{2k+1} \rightarrow \mathbb{R}$ a smooth function and let $\xi(x) = \sum F(x_{i-k}, \dots, x_{i+k})$. We define $\vec{J}_N(t, x)$ the element of \mathbb{R}^N given by $\vec{J}_N(t, x)_i := \sum_{j=1}^{i-1}$

$$J'(t, j/N) + \frac{\partial}{\partial x_i} \left[\frac{1}{N} \sum_{j=1}^N J'(t, j/N) F(x_{j-k}, \dots, x_{j+k}) \right]. \text{ We have}$$

$$\begin{aligned} & \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\ & \geq \frac{2}{N} \int_0^T \int \langle A(\nabla \nu_t + \nu_t \nabla H), \vec{J} \rangle dt - \frac{1}{N} \int_0^T \int \langle A \vec{J}, \vec{J} \rangle \nu_t(dx) dt \\ & = 2 \int_0^T \int \sum J'(t, i/N) W_{i,i+1}(x) \nu_t(dx) dt \\ & \quad - 2 \int_0^T \int \sum \frac{J'(t, i/N)}{N^2} (LF)(x_{i-k}, \dots, x_{i+k}) \nu_t(dx) dt \\ & \quad - \int_0^T \int \sum a(x_i, x_{i+1}) J'(t, i/N)^2 \left(1 - \frac{\partial \xi}{\partial x_{i+1}} + \frac{\partial \xi}{\partial x_i}\right)^2 \nu_t(dx) dt + o(1) \end{aligned} \tag{3.46}$$

We then just have to use Corollary 3.24 and optimize in F to obtain

$$\begin{aligned} & \liminf_N \frac{1}{N} \int_0^T \int \frac{\langle A(\nabla \nu_t + \nu_t \nabla H), (\nabla \nu_t + \nu_t \nabla H) \rangle}{\nu_t} dt \\ & \geq 2 \int_0^T \int_{\mathbb{T}} \varphi'(\rho(t, \theta)) J'(t, \theta) d\theta dt - \int_0^T \int_{\mathbb{T}} \hat{a}(\rho) J(t, \theta)^2 d\theta dt. \end{aligned}$$

Taking the supremum over all smooth functions J then yields our Lemma.

For the Gamma-convergence upper bound, the method we use is pretty much the same as the proof of the LDP lower bound in [35], so we only give a rough sketch. We fix a smooth profile ρ , for which there exists a continuous function $h : [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ such that

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \theta} \left(\hat{a}(\rho) \left(\frac{\partial}{\partial \theta} \varphi'(\rho) + h \right) \right). \tag{3.47}$$

We consider an evolution given by the generator

$$\tilde{L}_N f := L_N f + N \sum_{i=1}^N h(t, i/N) a(x_i, x_{i+1}) \left(1 + \frac{\partial}{\partial x_{i+1}} \xi - \frac{\partial}{\partial x_i} \xi \right) \left(\frac{\partial}{\partial x_{i+1}} f - \frac{\partial}{\partial x_i} f \right).$$

and initial condition the local Gibbs state associated to $\rho(0, \cdot)$. It is shown in Section 3 of [35] that the solutions converge to the deterministic profile ρ . We write $f_{N,F}$ the

law of the solution. We have

$$\begin{aligned} \frac{1}{N} J_N(f_N) &= \frac{1}{2N} \int_0^T \int a(x_i, x_{i+1}) h(t, i/N)^2 \left(1 + \frac{\partial}{\partial x_{i+1}} \xi - \frac{\partial}{\partial x_i} \xi \right)^2 f_N(dx) dt \\ &\rightarrow \frac{1}{2} \int_0^T \int_{\mathbb{T}} a_F(\rho) \left(\frac{\partial h}{\partial \theta} \right)^2 d\theta dt \end{aligned} \quad (3.48)$$

If, instead of using a fixed function F , we use a sequence (F_N) such that $a_{F_N}(\rho)$ uniformly converges to $\hat{a}(\rho)$ on compact sets (which is possible, see Proposition 3.21), we obtain the upper bound

$$\limsup \frac{1}{N} J_N(f_{N, F_N}) \leq \frac{1}{2} \int_0^T \int_{\mathbb{T}} \hat{a}(\rho) \left(\frac{\partial h}{\partial \theta} \right)^2 d\theta dt$$

which, by representation (3.47), is the one we needed to prove the upper bound in the Gamma convergence. \square

Acknowledgments: I would like to thank Christian Léonard for having pointed out to me the use of Girsanov's theorem to understand relative entropy for diffusion processes. I would also like to thank Thierry Bodineau, Georg Menz, Felix Otto, S.R.S. Varadhan and Cédric Villani for discussions about gradient flows and hydrodynamic limits, as well as the referees for their advice and comments.

References

- [1] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Comm. Math. Phys.*, 307:791–815, 2011
- [2] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. Large deviations and gradient flows. *Philosophical Transactions of the Royal Society A* (2013).
- [3] L. Ambrosio, N. Gigli, and G. Savaré, Gradient Flows In Metric Spaces And In The Space Of Probability Measures, Second edition. Lectures in Mathematics ETH Zurich. Birkhauser Verlag, Basel, 2008. x+334 pp. ISBN: 978-3-7643-8721-1
- [4] L. Ambrosio, G. Savaré and L. Zambotti, Existence and Stability for Fokker-Planck equations with log-concave reference measure, *PTRF*, vol. 145, 517-564 (2009)
- [5] L. Bertini, C. Landim and M. Mourragui, Dynamical large deviations for the boundary driven weakly asymmetric exclusion process. *Ann. Probab.* 37 (2009), no. 6, 2357–2403.
- [6] Manh Hong Duong, Vaios Laschos, and D.R.M. Renger. Wasserstein gradient flows from large deviations of thermodynamic limits. *ESAIM: COCV*, 19(4), 1166-1188, 2013.
- [7] D. A. Dawson and J. Gärtner Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* 20 (1987), no. 4, 247–308.

- [8] E. De Giorgi, A. Marino and M. Tosques, Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) 68 (1980), no. 3, 180–187.
- [9] M. H. Duong, M. Peletier and J. Zimmer, GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to Large Deviation Principle. *Nonlinearity* 26 (2013) 2951-2971.
- [10] M. Donsker and S.R.S. Varadhan, Large deviations from a hydrodynamic scaling limit. *Comm. Pure Appl. Math.*, Number 42 (1989), 243-270.
- [11] A. Dembo and O. Zeitouni, Large deviations techniques and applications. *Stochastic Modelling and Applied Probability*, 38. Springer-Verlag, Berlin, 2010. xvi+396 pp.
- [12] M. Fathi A two-scale approach to the hydrodynamic limit, part II : local Gibbs behavior. *ALEA Lat. Am. J. Probab. Math. Stat.* 10 (2013), no. 2, 625651.
- [13] M. Fathi, Modified logarithmic Sobolev inequalities for canonical ensembles. *to appear in ESAIM: Probab. Stat.*, <http://arxiv.org/abs/1306.1484>
- [14] M. Fathi and G. Menz, Hydrodynamic limit for conservative spin systems with super-quadratic, partially inhomogeneous single-site potential. *preprint*, <http://arxiv.org/abs/1405.3327>
- [15] J. Feng and Thomas G. Kurtz. Large deviations for stochastic processes, volume 131 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2006
- [16] H. Föllmer, Random fields and diffusion processes, Ecole d'Eté de Probabilités de Saint-Flour XV-XVII, 1985-87, Lecture Notes in Math., 1362, Springer, 1988. (101-203).
- [17] J. Fritz, Hydrodynamics in a symmetric random medium. *Comm. Math. Phys.*, 125, 13-25, (1989).
- [18] N. Grunewald, F. Otto, C. Villani and M. G. Westdickenberg, A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. H. Poincaré Probab. Statist.* 45 (2009), 2, 302–351.
- [19] M.Z. Guo, G.C. Papanicolaou and S.R.S. Varadhan, Nonlinear Diffusion Limit for a System with Nearest Neighbor Interactions, *Commun. Math. Phys.* 118, 31-59 (1988)
- [20] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, **29(1)**:1-17, (1998).
- [21] E. Kosygina, The Behavior of the Specific Entropy in the Hydrodynamic Scaling Limit for the Ginzburg-Landau Model, *Markov Processes and Related Fields*, **7**, 3 (2001), pp. 383-417.
- [22] C. Kipnis and C. Landim, Scaling limits of interacting particle systems. *Grundlehren der Mathematischen Wissenschaften*, 320. Springer-Verlag, Berlin, 1999. xvi+442 pp. ISBN: 3-540-64913-1
- [23] C. Kipnis and S.R.S. Varadhan, Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.* 104, 1-19 (1986)
- [24] S. Lisini, Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces, *ESAIM: COCV* 15 (2009) 712-740
- [25] C. Leonard, Girsanov Theory under a finite entropy condition, Séminaire de probabilités de Strasbourg, vol. XLIV. Lecture Notes in Mathematics 2046, Springer-Verlag, 2012, 429-465.

- [26] J. Maas, Gradient flows of the entropy for finite Markov chains *J. Funct. Anal.* 261 (8) (2011), 2250-2292.
- [27] M. Mariani, A Gamma-convergence approach to large deviations, <http://arxiv.org/abs/1204.0640>
- [28] A. Mielke, Formulation of thermoelastic dissipative material behavior using GENERIC. *Contin. Mech. Thermodyn.*, 23 (3): 233-256, 2011.
- [29] A. Mielke, Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. PDE*, 48 (1-2), 1–31, 2013.
- [30] G. Menz and F. Otto, Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.* 41 (3B), 2182-2224 (2013)
- [31] A. Mielke, M. A. Peletier and M. Renger, On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Poten. Anal.*, 41(4): 1293-1327, 2014.
- [32] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, **26**(1-2):101-174, (2001).
- [33] F. Otto and C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, **173**(2):361-400, (2000).
- [34] M. A. Peletier, F. Redig and K. Vafayi, Large Deviations in Stochastic Heat-Conduction Processes Provide a Gradient-Flow Structure for Heat Conduction [arxiv:1403.4994](https://arxiv.org/abs/1403.4994) (2014)
- [35] J. Quastel, Large deviations from a hydrodynamic scaling limit for a nongradient system. *Ann. Probab.* 23 (1995), no. 2, 724-742.
- [36] D. Revuz and M. Yor, Continuous Martingales and Brownian Motion, *Grundlehren der Mathematischen Wissenschaften*, 293. Springer-Verlag, Berlin, 1999. 602 pp. ISBN: 978-3-540-64325-8
- [37] E. Sandier and S. Serfaty, Gamma-convergence of gradient flows with applications to Ginzburg-Landau, *Comm. Pure Appl. Math.*, 57, No 12, (2004), 1627-1672.
- [38] S. Serfaty, Gamma-convergence of gradient flows on Hilbert and metric spaces and applications, *Disc. Cont. Dyn. Systems*, A, 31, No 4, (2011), 1427-1451
- [39] S.R.S. Varadhan, Nonlinear diffusion limit for a system with nearest neighbor interactions II, in *Asymptotic problems in probability theory : stochastic models and diffusions on fractals*, Pitman Research Notes in Mathematics Series, **283**.
- [40] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, Vol. **58**, American Mathematical Society.
- [41] C. Villani, Optimal Transport, Old and New. *Grundlehren der mathematischen Wissenschaften*, Vol. 338, Springer-Verlag, 2009.
- [42] H.T. Yau, Relative Entropy and Hydrodynamics of Ginzburg-Landau Models, *Lett. Math. Phys.*, **22** (1991), 63-80.