

ruhr.paD

UA Ruhr Zentrum für
partielle Differentialgleichungen

Variational approach to coarse-graining of generalized gradient flows

M. H. Duong, A. Lamacz, M. A. Peletier and U. Sharma

Preprint 2015-29

Variational approach to coarse-graining of generalized gradient flows

Manh Hong Duong, Agnes Lamacz, Mark A. Peletier and Upanshu Sharma

August 4, 2015

Abstract

In this paper we present a variational technique that handles coarse-graining and passing to a limit in a unified manner. The technique is based on a duality structure, which is present in many gradient flows and other variational evolutions, and which often arises from a large-deviations principle. It has three main features: (A) a natural interaction between the duality structure and the coarse-graining, (B) application to systems with non-dissipative effects, and (C) application to coarse-graining of approximate solutions which solve the equation only to some error. As examples, we use this technique to solve three limit problems, the overdamped limit of the Vlasov-Fokker-Planck equation and the small-noise limit of randomly perturbed Hamiltonian systems with one and with many degrees of freedom.

Contents

1	Introduction	2
1.1	Variational approach—an outline	2
1.2	Origin of the functional I^ε : large deviations of a stochastic particle system	4
1.3	Concrete Problems	6
1.3.1	Overdamped limit of the Vlasov-Fokker-Planck equation	6
1.3.2	Small-noise limit of a randomly perturbed Hamiltonian system with one degree of freedom	6
1.3.3	Small-noise limit of a randomly perturbed Hamiltonian system with d degrees of freedom	8
1.4	Comparison with other work	8
1.5	Outline of the article	9
1.6	Summary of notation	9
2	Overdamped Limit of the VFP equation	9
2.1	Setup of the system	9
2.2	A priori bounds	10
2.3	Coarse-graining and compactness	13
2.4	Local equilibrium	14
2.5	Liminf inequality	15
2.6	Discussion	17
3	Diffusion on a Graph, $d = 1$	17
3.1	Construction of the graph Γ	18
3.2	Adding noise: diffusion on the graph	19
3.3	Compactness	19
3.4	Local equilibrium	20
3.5	Continuity of ρ and $\hat{\rho}$	21
3.6	Liminf inequality	22
3.7	Study of the limit problem	24
3.8	Conclusion and Discussion	27
4	Diffusion on a Graph, $d > 1$	27
5	Conclusion and discussion	29
A	Proof of Lemma 2.1	30
B	Proof of Theorem 2.3	32

1 Introduction

Coarse-graining is the procedure of approximating a system by a simpler or lower-dimensional one, often in some limiting regime. It arises naturally in various fields such as thermodynamics, quantum mechanics, and molecular dynamics, just to name a few. Typically coarse-graining requires a separation of temporal and/or spatial scales, i.e. the presence of fast and slow variables. As the ratio of ‘fast’ to ‘slow’ increases, some form of averaging or homogenization should allow one to remove the fast scales, and obtain a limiting system that focuses on the slow ones.

Coarse-graining limits are by nature *singular limits*, since information is lost in the coarse-graining procedure; therefore rigorous proofs of such limits are always non-trivial. Although the literature abounds with cases that have been treated successfully, and some fields can even be called well-developed—singular limits in ODEs and homogenization theory, to name just two—many more cases seem out of reach, such as coarse-graining in materials [dPC07], climate prediction [SATS07], and complex systems [FR07, NN12].

All proofs of singular limits hinge on using certain *special structure* of the equations; well-known examples are compensated compactness [Tar79, Mur87], the theories of viscosity solutions [CIL92] and entropy solutions [Kru70, Smo94], and the methods of periodic unfolding [CDG02, CDG08] and two-scale convergence [All92]. *Variational-evolution structure*, such as in the case of gradient flows and variational rate-independent systems, also facilitates limits [SS04, Ste08, MRS08, DS10, Ser11, MRS12, Mie14].

In this paper we introduce and study such a structure, which arises from the theory of *large deviations* for stochastic processes. In recent years we have discovered that many gradient flows, and also many ‘generalized’ gradient systems, can be matched one-to-one to the large-deviation characterization of some stochastic process [ADPZ11, ADPZ13, DPZ14, DPZ13, DLZ12, MPR14]. The large-deviation rate functional, in this connection, can be seen to *define* the generalized gradient system. This connection has many philosophical and practical implications, which are discussed in the references above.

We show how in such systems, described by a rate functional, ‘passing to a limit’ is facilitated by the duality structure that a rate function inherits from the large-deviation context, in a way that meshes particularly well with coarse-graining.

1.1 Variational approach—an outline

The systems that we consider in this paper are evolution equations in a space of measures. Typical examples are the forward Kolmogorov equations associated with stochastic processes, but also various nonlinear equations, as in one of the examples below.

Consider the family of evolution equations

$$\begin{aligned}\partial_t \rho^\varepsilon &= \mathcal{N}^\varepsilon \rho^\varepsilon, \\ \rho^\varepsilon|_{t=0} &= \rho_0^\varepsilon,\end{aligned}\tag{1}$$

where \mathcal{N}^ε is a linear or nonlinear operator. The unknown ρ^ε is a time-dependent Borel measure on a state space \mathcal{X} , i.e. $\rho^\varepsilon : [0, T] \rightarrow \mathcal{M}(\mathcal{X})$. In the systems of this paper, (1) has a variational formulation characterized by a functional I^ε such that

$$I^\varepsilon \geq 0 \quad \text{and} \quad \rho^\varepsilon \text{ solves (1)} \iff I^\varepsilon(\rho^\varepsilon) = 0.\tag{2}$$

This variational formulation is closely related to the Brezis-Ekeland-Nayroles variational principle [BE76, Nay76, Ste08, Gho09] and the integrated energy-dissipation identity for gradient flows [AGS08]; see Section 5.

Our interest in this paper is the limit $\varepsilon \rightarrow 0$, and we wish to study the behaviour of the system in this limit. If we postpone the aspect of coarse-graining for the moment, this corresponds to studying the limit of ρ^ε as $\varepsilon \rightarrow 0$. Since ρ^ε is characterized by I^ε , establishing the limiting behaviour consists of answering two questions:

1. *Compactness*: Do solutions of $I^\varepsilon(\rho^\varepsilon) = 0$ have useful compactness properties, allowing one to extract a subsequence that converges in a suitable topology, say ζ ?

2. *Liminf inequality*: Is there a limit functional $I \geq 0$ such that

$$\rho^\varepsilon \xrightarrow{c} \rho \implies \liminf_{\varepsilon \rightarrow 0} I^\varepsilon(\rho^\varepsilon) \geq I(\rho)? \quad (3)$$

And if so, does one have

$$I(\rho) = 0 \iff \rho \text{ solves } \partial_t \rho = \mathcal{N}\rho,$$

for some operator \mathcal{N} ?

A special aspect of the method of the present paper is that it also applies to *approximate* solutions. By this we mean that we are interested in sequences of time-dependent Borel measures ρ^ε such that $\sup_{\varepsilon > 0} I^\varepsilon(\rho^\varepsilon) \leq C$ for some $C \geq 0$. The exact solutions are special cases when $C = 0$. The main message of our approach is that all the results then follow from this uniform bound and assumptions on well-prepared initial data.

The compactness question will be answered by the first crucial property of the functionals I^ε , which is that they provide an *a priori* bound of the type

$$S^\varepsilon(\rho_t^\varepsilon) + \int_0^t R^\varepsilon(\rho_s^\varepsilon) ds \leq S^\varepsilon(\rho_0^\varepsilon) + I^\varepsilon(\rho^\varepsilon), \quad (4)$$

where ρ_t^ε denotes time slice at time t and S^ε and R^ε are functionals. In the examples of this paper S^ε is a free energy and R^ε a relative Fisher Information, but the structure is more general. This inequality is reminiscent of the energy-dissipation inequality in the gradient-flow setting. The uniform bound, by assumption, of the right-hand side of (4) implies that each term in the left-hand side of (4), i.e., the free energy at any time $t > 0$ and the integral of the Fisher information, is also bounded. This will be used to apply the Arzelà-Ascoli theorem to obtain certain compactness and ‘local-equilibrium’ properties. All this discussion will be made clear in each example in this paper.

The second crucial property of the functionals I^ε is that they satisfy a duality relation of the type

$$I^\varepsilon(\rho) = \sup_f \mathcal{J}^\varepsilon(\rho, f), \quad (5)$$

where the supremum is taken over a class of smooth functions f . It is well known how such duality structures give rise to good convergence properties such as (3), but the focus in this paper is on how this duality structure combines well with coarse-graining.

In this paper we define *coarse-graining* to be a shift to a reduced, lower dimensional description via a coarse-graining map $\xi : \mathcal{X} \rightarrow \mathcal{Y}$ which identifies relevant information and is typically highly non-injective. Note that ξ may depend on ε . A typical example of such a coarse-graining map is a ‘reaction coordinate’ in molecular dynamics. The coarse-grained equivalent of $\rho^\varepsilon : [0, T] \rightarrow \mathcal{M}(\mathcal{X})$ is the push-forward $\hat{\rho}^\varepsilon := \xi_{\#}\rho^\varepsilon : [0, T] \rightarrow \mathcal{M}(\mathcal{Y})$. If ρ^ε is the law of a stochastic process X^ε , then $\xi_{\#}\rho^\varepsilon$ is the law of the process $\xi(X^\varepsilon)$.

There might be several reasons to be interested in $\xi_{\#}\rho^\varepsilon$ rather than ρ^ε itself. The push-forward $\xi_{\#}\rho^\varepsilon$ obeys a dynamics with fewer degrees of freedom, since ξ is non-injective; this might allow for more efficient computation. Our first example (see Section 1.3), the overdamped limit in the Vlasov-Fokker-Planck equation, is an example of this. As a second reason, by removing certain degrees of freedom, some specific behaviour of ρ^ε might become clearer; this is the case with our second and third examples (Section 1.3), where the effect of ξ is to remove a rapid oscillation, leaving behind a slower diffusive movement. Whatever the reason, in this paper we assume that some ξ is given, and that we wish to study the limit of $\xi_{\#}\rho^\varepsilon$ as $\varepsilon \rightarrow 0$.

The core of the arguments of this paper, that leads to the characterization of the equation satisfied by

the limit of $\xi_{\#}\rho^\varepsilon$, is captured by the following formal calculation:

$$\begin{aligned}
I^\varepsilon(\rho^\varepsilon) &= \sup_f \mathcal{J}^\varepsilon(\rho^\varepsilon, f) \\
&\stackrel{f=g\circ\xi}{\geq} \sup_g \mathcal{J}^\varepsilon(\rho^\varepsilon, g\circ\xi) \\
&\quad \downarrow \varepsilon \rightarrow 0 \\
&\sup_g \mathcal{J}(\rho, g\circ\xi) \\
&\stackrel{(*)}{=} \sup_g \hat{\mathcal{J}}(\hat{\rho}, g) \quad \stackrel{(**)}{=} \hat{I}(\hat{\rho})
\end{aligned}$$

Let us go through the lines one by one.

First, the inequality in the calculation above is due to reduction to a subset of special functions f , namely those of the form $f = g \circ \xi$. This is in fact an implementation of coarse-graining: in the supremum we decide to limit ourselves to observables of the form $g \circ \xi$ which only have access to the information provided by ξ . After this reduction we pass to the limit and show that $\mathcal{J}^\varepsilon(\rho^\varepsilon, g \circ \xi)$ converges to some $\mathcal{J}(\rho, g \circ \xi)$ —at least for appropriately chosen coarse-graining maps.

In the step $(*)$ one requires that the loss-of-information in passing from ρ to $\hat{\rho}$ is consistent with the loss-of-resolution in considering only functions $f = g \circ \xi$. This step requires a proof of *local equilibrium*, which describes how the behaviour of ρ that is *not* represented explicitly by the push-forward $\hat{\rho}$, can nonetheless be deduced from $\hat{\rho}$. This local-equilibrium property is at the core of various coarse-graining methods and is typically determined case by case.

We finally define \hat{I} by duality in terms of $\hat{\mathcal{J}}$ as in $(**)$. In a *successful* application of this method, the resulting functional \hat{I} at the end has ‘good’ properties *despite* the loss-of-accuracy introduced by the restriction to functions of the form $g \circ \xi$, and this fact acts as a test of success. Such good properties should include, for instance, the property that $\hat{I} = 0$ has a unique solution in an appropriate sense.

Now let us explain the origin of the functionals I^ε .

1.2 Origin of the functional I^ε : large deviations of a stochastic particle system

The abstract methodology that we described above arises naturally in the context of *large deviations*, and we next describe this in the context of the three examples that we discuss in the next section. All three originate from (slight modifications of) one stochastic process, that models a collection of interacting particles with inertia in the physical space \mathbb{R}^d :

$$dQ_i^n(t) = \frac{P_i^n(t)}{m} dt, \tag{6a}$$

$$dP_i^n(t) = -\nabla V(Q_i^n(t))dt - \frac{1}{n} \sum_{j=1}^n \nabla \psi(Q_j^n(t) - Q_i^n(t))dt - \frac{\gamma}{m} P_i^n(t)dt + \sqrt{2\gamma\theta} dW_i(t). \tag{6b}$$

Here $Q_i^n \in \mathbb{R}^d$ and $P_i^n \in \mathbb{R}^d$ are the position and momentum of particles $i = 1, \dots, n$ with mass m . Equation (6a) is the usual relation between Q_i^n and P_i^n , and (6b) is a force balance which describes the forces acting on the particle. For this system, corresponding to the first example below, these forces are (a) a force arising from a fixed potential V , (b) an interaction force deriving from a potential ψ , (c) a friction force, and (d) a stochastic force characterized by independent d -dimensional Wiener measures W_i . Throughout this paper we collect Q_i^n and P_i^n into a single variable $X_i^n = (Q_i^n, P_i^n)$.

The parameter γ characterizes the intensity of collisions of the particle with the solvent; it is present in both the friction term and the noise term, since they both arise from these collisions (and in accordance with the Einstein relation). The parameter $\theta = kT_a$, where k is the Boltzmann constant and T_a is the absolute temperature, measures the mean kinetic energy of the solvent molecules, and therefore characterizes the

magnitude of collision noise. Typical applications of this system are for instance as a simplified model for chemical reactions, or as a model for particles interacting through Coulomb, gravitational, or volume-exclusion forces. However, our focus in this paper is on methodology, not on technicality, so we will assume that ψ is sufficiently smooth later on.

We now consider the many-particle limit $n \rightarrow \infty$ in (6). It is a well-known fact that the empirical measure

$$\rho_n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{X_i^n(t)} \quad (7)$$

converges almost surely to the unique solution of the *Vlasov-Fokker-Planck (VFP) equation* [Oel84]

$$\partial_t \rho = (\mathcal{L}_\rho)^* \rho, \quad (\mathcal{L}_\mu)^* \rho := -\operatorname{div}_q \left(\rho \frac{p}{m} \right) + \operatorname{div}_p \rho \left(\nabla_q V + \nabla_q \psi * \mu + \gamma \frac{p}{m} \right) + \gamma \theta \Delta_p \rho, \quad (8)$$

$$= -\operatorname{div} \rho J \nabla (H + \psi * \mu) + \gamma \operatorname{div}_p \rho \frac{p}{m} + \gamma \theta \Delta_p \rho, \quad (9)$$

with an initial datum that derives from the initial distribution of X_i^n . The spatial domain here is \mathbb{R}^{2d} with coordinates $(q, p) \in \mathbb{R}^d \times \mathbb{R}^d$, and subscripts such as in ∇_q and Δ_p indicate that differential operators act only on corresponding variables. The convolution is defined by $(\psi * \rho)(q) = \int_{\mathbb{R}^{2d}} \psi(q - q') \rho(q', p') dq' dp'$. In the second line above we use a slightly shorter way of writing \mathcal{L}_μ^* , by introducing the Hamiltonian $H(q, p) = p^2/2m + V(q)$ and the canonical symplectic matrix $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. This way of writing also highlights that the system is a combination of conservative effects, described by J , H , and ψ , and dissipative effects, which are parametrized by γ . For future reference we also give the primal form \mathcal{L}_μ explicitly:

$$\mathcal{L}_\mu f = J \nabla (H + \psi * \mu) \cdot \nabla f - \gamma \frac{p}{m} \cdot \nabla_p f + \gamma \theta \Delta_p f.$$

The almost-sure convergence of ρ_n to the solution ρ of the (deterministic) VFP equation is the starting point for a *large-deviation* result. In particular it has been shown that the sequence (ρ_n) has a *large-deviation property* [DG87, BDF12, DPZ13] which characterizes the probability of finding the empirical measure far from the limit ρ , written informally as

$$\operatorname{Prob}(\rho_n \approx \rho) \sim \exp\left(-\frac{n}{2} I(\rho)\right),$$

in terms of a *rate functional* $I : C([0, T]; \mathcal{P}(\mathbb{R}^{2d})) \rightarrow \mathbb{R}$. Assuming that the initial data X_i^n are chosen to be deterministic, and such that the initial empirical measure $\rho_n(0)$ converges narrowly to some ρ_0 ; then I has the form, see [DPZ13],

$$I(\rho) := \sup_{f \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^{2d})} \int_{\mathbb{R}^{2d}} f_T d\rho_T - \int_{\mathbb{R}^{2d}} f_0 d\rho_0 - \int_0^T \int_{\mathbb{R}^{2d}} (\partial_t f + \mathcal{L}_{\rho_t} f) d\rho_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2d}} \Lambda(f, f) d\rho_t dt, \quad (10)$$

provided $\rho_t|_{t=0} = \rho_0$, where Λ is the carré-du-champ operator (e.g. [BGL⁺14, Section 1.4.2])

$$\Lambda(f, g) := \frac{1}{2} (\mathcal{L}_\mu(fg) - f \mathcal{L}_\mu g - g \mathcal{L}_\mu f) = \gamma \theta \nabla_p f \nabla_p g.$$

If the initial measure $\rho_t|_{t=0}$ is not equal to the limit ρ_0 of the stochastic initial empirical measures, then $I(\rho) = \infty$.

Note that the functional I in (10) is non-negative, since $f \equiv 0$ is admissible. If $I(\rho) = 0$, then by replacing f by λf and letting λ tend to zero we find that ρ is the weak solution of (8) (which is unique, given initial data ρ_0 [Fun84]). Therefore I is of the form that we discussed in Section 1.1: $I \geq 0$, and $I(\rho) = 0$ iff ρ solves (8), which is a realization of (1).

1.3 Concrete Problems

We now apply the coarse-graining method of Section 1.1 to three limits: the *overdamped* limit $\gamma \rightarrow \infty$, and two *small-noise* limits $\theta \rightarrow 0$. In each of these three limits, the VFP equation (8) is the starting point, and we prove convergence to a limiting system using appropriate coarse-graining maps. Note that the convergence is therefore from one deterministic equation to another one; but the method makes use of the large-deviation structure that the VFP equation has inherited from its stochastic origin.

1.3.1 Overdamped limit of the Vlasov-Fokker-Planck equation

The first limit that we consider is the limit of large friction, $\gamma \rightarrow \infty$, in the Vlasov-Fokker-Planck equation (8), setting $\theta = 1$ for convenience. To motivate what follows, we divide (8) throughout by γ and formally let $\gamma \rightarrow \infty$ to find

$$\operatorname{div}_p \rho \left(\frac{p}{m} \right) + \Delta_p \rho = 0,$$

which suggests that in the limit $\gamma \rightarrow \infty$, ρ should be Maxwellian in p , i.e.

$$\rho_t(dq, dp) = Z^{-1} \exp\left(-\frac{p^2}{2m}\right) dp \sigma_t(dq), \quad (11)$$

where Z is the normalization constant for the Maxwellian distribution. The main result in Section 2 shows that after an appropriate time rescaling, in the limit $\gamma \rightarrow \infty$, the remaining unknown $\sigma \in C([0, T]; \mathcal{P}(\mathbb{R}^d))$ solves the Vlasov-Fokker-Planck equation

$$\partial_t \sigma = \operatorname{div}(\sigma \nabla V(q)) + \operatorname{div}(\sigma (\nabla \psi * \sigma)) + \Delta \sigma. \quad (12)$$

In his seminal work [Kra40], Kramers formally discussed these results for the ‘Kramers equation’, which corresponds to (8) with $\psi = 0$, and this limit has become known as the *Smoluchowski-Kramers approximation*. Nelson made these ideas rigorous [Nel67] by studying the corresponding stochastic differential equations (SDEs); he showed that under suitable rescaling the solution to the Langevin equation converges almost surely to the solution of (12) with $\psi = 0$. Since then various generalizations and related results have been proved [Fre04, CF06, Nar94, HVW12], mostly using stochastic and asymptotic techniques.

In this article we recover some of the results mentioned above for the VFP equation using the variational technique described in Section 1.1. Our proof is made up of the following three steps. Theorem 2.4 provides the necessary compactness properties to pass to the limit, Lemma 2.5 gives characterization (11) of the limit, and in Theorem 2.6 we prove the convergence of the solution of the VFP equation to the solution of (12).

1.3.2 Small-noise limit of a randomly perturbed Hamiltonian system with one degree of freedom

In our second example we consider the following equation

$$\partial_t \rho = -\operatorname{div}_q \left(\rho \frac{p}{m} \right) + \operatorname{div}_p (\rho \nabla_q V) + \varepsilon \Delta_p \rho \quad \text{on } \mathbb{R} \times \mathbb{R}^2, \quad (13)$$

where $(q, p) \in \mathbb{R}^2$, $t \in \mathbb{R}$ and div_q , div_p , Δ_p are one-dimensional derivatives. This equation can also be written as

$$\partial_t \rho = -\operatorname{div}(\rho J \nabla H) + \varepsilon \Delta_p \rho, \quad \text{on } \mathbb{R} \times \mathbb{R}^2. \quad (14)$$

This corresponds to the VFP equation (8) with $\psi = 0$, without friction and with small noise $\varepsilon = \gamma \theta$.

In addition to the interpretation as the many-particle limit of (6), Equation (14) also is the forward Kolmogorov equation of a randomly perturbed Hamiltonian system in \mathbb{R}^2 with Hamiltonian H :

$$X = \begin{pmatrix} Q \\ P \end{pmatrix}, \quad dX_t = J \nabla H(X_t) + \sqrt{2\varepsilon} \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t, \quad (15)$$

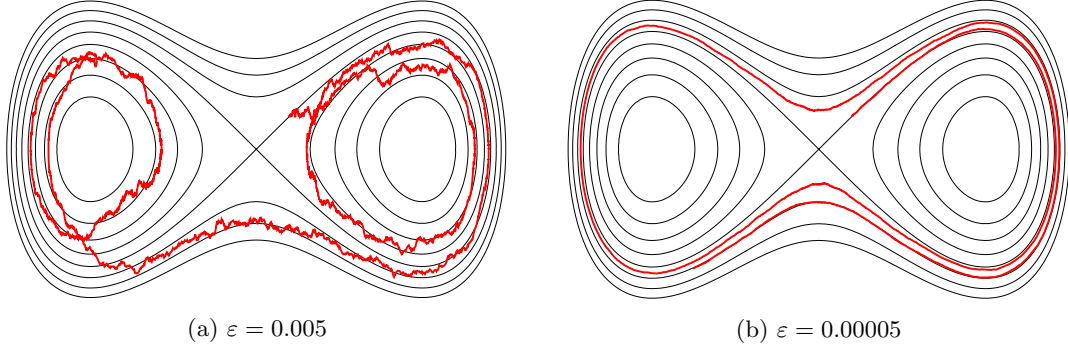


Figure 1: Simulation of (15) for varying ε . Shown are the level curves of the Hamiltonian H and for each case a single trajectory.

where W_t is a 1-dimensional Wiener process. This system is a prototype for a large class of Hamiltonian systems perturbed by random noise. When the amplitude ε of the noise is small, the dynamics (14) splits into fast and slow components. The fast component approximately follows an unperturbed trajectory of the Hamiltonian system, which is a level set of H . The slow component is visible as a slow modification of the value of H , corresponding to a motion transversal to the level sets of H . Figure 1 illustrates this.

Following [FW94] and others, in order to focus on the slow, Hamiltonian-changing motion, we rescale time such that the Hamiltonian, level-set-following motion is fast, of rate $O(1/\varepsilon)$, and the level-set-changing motion is of rate $O(1)$. In other words, the process (15) ‘whizzes round’ level sets of H , while shifting from one level set to another at rate $O(1)$.

This behaviour suggests choosing a coarse-graining map $\xi : \mathbb{R}^2 \rightarrow \Gamma$, which maps a whole *level set* to a single point in a new space Γ ; because of the structure of level sets of H , the set Γ has a structure that is called a *graph*, a union of one-dimensional intervals locally parametrized by the value of the Hamiltonian. Figure 2 illustrates this, and in Section 3 we discuss it in full detail.

After projecting onto the graph Γ , the process turns out to behave like a diffusion process on Γ . This property was first made rigorous in [FW94] for a system with one degree of freedom, as here, and non-degenerate noise, using probabilistic techniques. In [FW98] the authors consider the case of degenerate noise by using probabilistic and analytic techniques based on hypoelliptic operators. More recently this problem has been handled using PDE techniques [IS12] (the elliptic case) and Dirichlet forms [BvR14]. In Section 3 we give a new proof, using the structure outlined in Section 1.1.

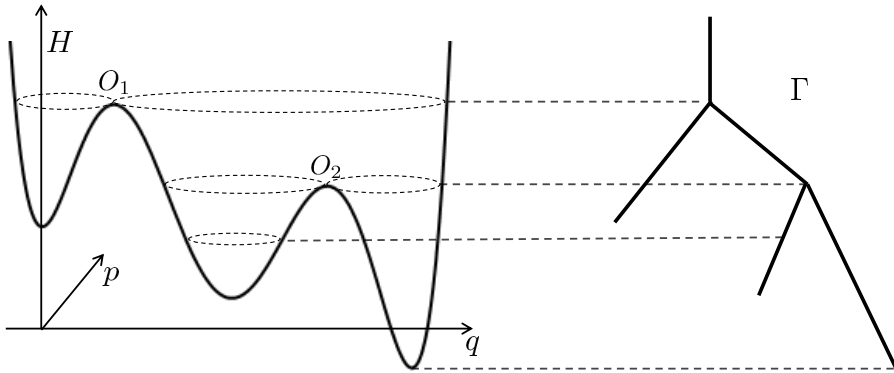


Figure 2: Left: Hamiltonian $\mathbb{R}^2 \ni (q, p) \mapsto H(q, p)$, Right: Graph Γ

1.3.3 Small-noise limit of a randomly perturbed Hamiltonian system with d degrees of freedom

The convergence of solutions of (14) as $\varepsilon \rightarrow 0$ to a diffusion process on a graph requires that the non-perturbed system has a unique invariant measure on each connected component of a level set. While this is true for a Hamiltonian system with one degree of freedom, in the higher-dimensional case one might have additional first integrals of motion. In such a system the slow component will not be a one-dimensional process but a more complicated object—see [FW04]. However, by introducing an additional stochastic perturbation that destroys all first integrals except the Hamiltonian, one can regain the necessary ergodicity, such that the slow dynamics again lives on a graph.

In Section 4 we discuss this case. Equation (14) gains an additional noise term, and reads

$$\partial_t \rho = -\operatorname{div}(\rho J \nabla H) + \kappa \operatorname{div}(a \nabla \rho) + \varepsilon \Delta_p \rho, \quad (16)$$

where $a : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d \times 2d}$ with $a \nabla H = 0$, $\dim(\operatorname{Kernel}(a)) = 1$ and $\kappa, \varepsilon > 0$ with $\kappa \gg \varepsilon$. The spatial domain is \mathbb{R}^{2d} , $d > 1$ with coordinates $(q, p) \in \mathbb{R}^d \times \mathbb{R}^d$ and the unknown is a trajectory in the space of probability measures $\rho : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^{2d})$. As before the aim is to derive the dynamics as $\varepsilon \rightarrow 0$. This problem was studied in [FW01] and the results closely mirror the previous case. The main difference lies in the proof of the local equilibrium statement, which we discuss in Section 4.

1.4 Comparison with other work

The novelty of the present paper lies in the following.

1. *In comparison with existing literature on the three concrete examples treated in this paper:* The results of the three examples are known in the literature (see for instance [Nel67, FW94, FW98, FW01]), but they are proved by different techniques and in a different setting. The variational approach of this paper, which has a clear microscopic interpretation from the large-deviation principle, to these problems is new. We provide alternative proofs, recovering known results, in a unified framework. In addition, we obtain all the results on compactness, local-equilibrium properties and liminf inequalities solely from the variational structures. The approach also is applicable to approximate solutions, which obey the original fine-grained dynamics only to some error. This allows us to work with larger class of measures and to relax many regularity conditions required by the exact solutions. Furthermore, our abstract setting has potential applications to many other systems.
2. *In comparison with recently developed variational-evolutionary methods:* Many recently developed variational techniques for ‘passing to a limit’ such as the Sandier-Safarty method based on the Ψ - Ψ^* structure [SS04, AMP⁺12, Mie14] only apply to gradient flows, i.e. dissipative systems. The approach of this paper also applies to certain variational-evolutionary systems that include non-dissipative effects, such as GENERIC systems [Ött05, DPZ13], as in the examples. Since our approach only uses the duality structure of the rate functionals, which holds true for more general systems, we expect that our method works for other limits in non-gradient-flow systems such as the Langevin limit of the Nosé-Hoover-Langevin thermostat [FG11, OP11].
3. *Quantification of the coarse-graining error.* The use of the rate functional as a central ingredient in ‘passing to a limit’ and coarse-graining also allows us to obtain quantitative estimates of the coarse-graining error. One intermediate result of our analysis is a functional inequality similar to the energy-dissipation inequality in the gradient-flow setting (see (4)). This inequality provides an upper bound on the free energy and the integral of the Fisher information by the rate functional and initial free energy. This offers an alternative to the Talagrand and log-Sobolev inequalities used in the literature [LL10, GOVW09] to obtain quantification of the coarse-graining error. To keep the paper to a reasonable length, we address this issue in details separately in a companion article [DLP⁺15].

We provide further comments in Section 5.

1.5 Outline of the article

The rest of the paper is devoted to the study of three concrete problems: the overdamped limit of the VFP equation in Section 2, diffusion on a graph with one degree of freedom in Section 3 and diffusion on a graph with many degrees of freedom in Section 4. In each Section, the main steps in the abstract framework are performed in detail. Section 5 provides further discussion. Finally, detailed proofs of some theorems are given in Appendices A and B.

1.6 Summary of notation

\pm_{kj}	± 1 , depending on which end vertex O_j lies of edge I_k	Sec. 3.1
\mathcal{F}	Free energy	(22), (45)
Γ, γ	The graph Γ and its elements γ	Sec. 3.1
$\mathcal{H}(\cdot \cdot)$	relative entropy	(21)
$H(q, p)$	$H(q, p) = p^2/2m + V(q)$, the Hamiltonian	
\mathcal{H}^n	n -dimensional Hausdorff measure	
$\mathcal{I}(\cdot \cdot)$	relative Fisher Information	(24)
Int	The interior of a set	
I^ε	Large-deviation rate functional for the diffusion-on-graph problem	(46)
I^γ	Large-deviation rate functional for the VFP equation	(19)
J	$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, the canonical symplectic matrix	
\mathcal{L}	Lebesgue measure	
$\mathcal{M}(\mathcal{X})$	space of finite, non-negative Borel measures on \mathcal{X}	
$\mathcal{P}(\mathcal{X})$	space of probability measures on \mathcal{X}	
$\hat{\rho}$	push-forward under ξ of ρ	(44)
$T(\gamma)$	period of the periodic orbit at $\gamma \in \Gamma$	(48)
$V(q)$	potential on position	
x	$x = (q, p)$ joint variable	
ξ^γ, ξ	coarse-graining maps	(30), (43)

Throughout we use measure notation and terminology. For a given topological space \mathcal{X} , the space $\mathcal{M}(\mathcal{X})$ is the space of non-negative, finite Borel measures on \mathcal{X} ; $\mathcal{P}(\mathcal{X})$ is the space of probability measures on \mathcal{X} . For a measure $\rho \in \mathcal{M}([0, T] \times \mathbb{R}^{2d})$, for instance, we often write $\rho_t \in \mathcal{M}(\mathbb{R}^{2d})$ for the time slice at time t ; we also often use both the notation $\rho(x)dx$ and $\rho(dx)$ when ρ is Lebesgue-absolutely-continuous. We equip $\mathcal{M}(\mathcal{X})$ and $\mathcal{P}(\mathcal{X})$ with the *narrow* topology, in which convergence is characterized by duality with continuous and bounded functions on \mathcal{X} .

2 Overdamped Limit of the VFP equation

2.1 Setup of the system

In this section we prove the large-friction limit $\gamma \rightarrow \infty$ of the VFP equation (8). Setting $\theta = 1$ for convenience, and speeding time up by a factor γ , the VFP equation reads

$$\partial_t \rho = \mathcal{L}_\rho^* \rho, \quad \mathcal{L}_\nu^* \rho := -\gamma \operatorname{div} \rho J \nabla (H + \psi * \nu) + \gamma^2 \left[\operatorname{div}_p \left(\rho \frac{p}{m} \right) + \Delta_p \rho \right], \quad (17)$$

where, as before, $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ and $H(q, p) = p^2/2m + V(q)$. The spatial domain is \mathbb{R}^{2d} with coordinates $(q, p) \in \mathbb{R}^d \times \mathbb{R}^d$ with $d \geq 1$, and $\rho \in C([0, T]; \mathcal{P}(\mathbb{R}^{2d}))$. For later reference we also mention the primal form of the operator \mathcal{L}_ν^* :

$$\mathcal{L}_\nu f = \gamma J \nabla (H + \psi * \nu) \cdot \nabla f - \gamma^2 \frac{p}{m} \cdot \nabla_p f + \gamma^2 \Delta_p f. \quad (18)$$

We assume

- (V1) The potential $V \in C^2(\mathbb{R}^d)$ has globally bounded second derivative. Furthermore $V \geq 0$, $|\nabla V|^2 \leq C(1+V)$ for some $C > 0$, and $e^{-V} \in L^1(\mathbb{R}^d)$.
- (V2) The interaction potential $\psi \in C^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ is symmetric, has globally bounded first and second derivatives, and the mapping $\nu \mapsto \int \nu * \psi d\nu$ is convex (and therefore non-negative).

As we described in Section 1.1, the study of the limit $\gamma \rightarrow \infty$ contains the following steps:

1. Prove compactness;
2. Prove a local-equilibrium property;
3. Prove a liminf inequality.

Each of these results is based on the large-deviation structure, which for Equation (17) is

$$I^\gamma(\rho) = \sup_{f \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^{2d})} \left[\int_{\mathbb{R}^{2d}} f_T d\rho_T - \int_{\mathbb{R}^{2d}} f_0 d\rho_0 - \int_0^T \int_{\mathbb{R}^{2d}} (\partial_t f_t + \mathcal{L}_{\rho_t} f_t) d\rho_t dt - \frac{\gamma^2}{2} \int_0^T \int_{\mathbb{R}^{2d}} |\nabla_p f_t|^2 d\rho_t dt \right]. \quad (19)$$

Alternatively the rate functional can be written as [DPZ13, Theorem 2.5]

$$I^\gamma(\rho) = \begin{cases} \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2d}} |h_t|^2 d\rho_t dt & \text{if } \partial_t \rho_t = \mathcal{L}_{\rho_t}^* \rho_t - \gamma \operatorname{div}_p(\rho_t h_t), \text{ for } h \in L^2(0, T; L^2_{\nabla}(\rho_t)), \\ +\infty & \text{otherwise,} \end{cases} \quad (20)$$

where \mathcal{L}_ν is given in (18), and $L^2_{\nabla}(\rho_t)$ is the completion of $\{\nabla_p \varphi : \varphi \in C_c^\infty(\mathbb{R}^{2d})\}$ in the ρ_t -weighted L^2 norm. This second form shows clearly how $I^\gamma(\rho) = 0$ is equivalent to the property that ρ solves the VFP equation (17). It also shows that if $I^\gamma(\rho) > 0$ then ρ is an approximative solution in the sense that it satisfies the VFP equation up to some error $-\gamma \operatorname{div}_p(\rho_t h_t)$ whose norm is controlled by the rate functional.

2.2 A priori bounds

We give ourselves a sequence, indexed by γ , of solutions ρ^γ to the VFP equation (17) with initial datum $\rho_t^\gamma|_{t=0} = \rho_0$. We will deduce the compactness of the sequence ρ^γ from a priori estimates, that are themselves derived from the rate function I^γ .

For nonnegative measures ν, ζ on \mathbb{R}^{2d} we first introduce:

- Relative entropy:

$$\mathcal{H}(\nu || \zeta) = \begin{cases} \int_{\mathbb{R}^{2d}} [f \log f] d\zeta & \text{if } \nu = f\zeta, \\ \infty & \text{otherwise.} \end{cases} \quad (21)$$

- The free energy for this system:

$$\mathcal{F}(\nu) := \mathcal{H}(\nu | Z_H^{-1} e^{-H} dx) + \frac{1}{2} \int_{\mathbb{R}^{2d}} \psi * \nu d\nu = \int_{\mathbb{R}^{2d}} \left[\log g + H + \frac{1}{2} \psi * g \right] g dx + \log Z_H, \quad (22)$$

where $Z_H = \int e^{-H}$ and the second expression makes sense whenever $\nu = g dx$.

The convexity of the term involving ψ (condition (V2)) implies that the free energy \mathcal{F} is strictly convex and has a unique minimizer $\mu \in \mathcal{P}(\mathbb{R}^{2d})$. This minimizer is a stationary point of the evolution (17), and has the implicit characterization

$$\mu \in \mathcal{P}(\mathbb{R}^{2d}) : \mu(dqdp) = Z^{-1} \exp\left(-[H(q, p) + (\psi * \mu)(q)]\right) dqdp, \quad (23)$$

where Z is the normalization constant for μ . Note that $\nabla_p \mu = -\mu \nabla_p H = -p\mu/m$.

We also define the *relative Fisher Information* with respect to μ (in the p -variable only):

$$\mathcal{I}(\nu|\mu) = \sup_{\varphi \in C_c^\infty(\mathbb{R}^{2d})} 2 \int_{\mathbb{R}^{2d}} \left[\Delta_p \varphi - \frac{p}{m} \nabla_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right] d\nu. \quad (24)$$

In the more common case in which the derivatives Δ_p and ∇_p are replaced by the full derivatives Δ and ∇ , the relative Fisher Information has an equivalent formulation in terms of the Lebesgue density of ν . In our case such equivalence only holds when ν is absolutely-continuous with respect to the Lebesgue measure in both q and p :

Lemma 2.1 (Equivalence of relative-Fisher-Information expressions for a.c. measures). *If $\nu \in \mathcal{P}(\mathbb{R}^{2d})$, $\nu(dx) = f(x)dx$ with $f \in L^1(\mathbb{R}^{2d})$, then*

$$\mathcal{I}(\nu|\mu) = \begin{cases} \int_{\mathbb{R}^{2d}} \left| \frac{\nabla_p f}{f} \mathbb{1}_{\{f>0\}} + \frac{p}{m} \right|^2 f dqdp, & \text{if } \nabla_p f \in L^1_{\text{loc}}(dqdp), \\ \infty, & \text{otherwise,} \end{cases} \quad (25)$$

where $\mathbb{1}_{\{f>0\}}$ denotes the indicator function of the set $\{x \in \mathbb{R}^{2d} \mid f(x) > 0\}$.

For a measure of the form $\zeta(dq)f(p)dp$, with $\zeta \llneqq dq$, \mathcal{I} in (24) may be finite while the integral in (25) is not defined. Because of the central role of duality in this paper, definition (24) is a natural one, as we shall see below. The proof of Lemma 2.1 is given in Appendix A.

In the introduction we mentioned that we expect ρ^γ to become Maxwellian in the limit $\gamma \rightarrow \infty$. This will be driven by a vanishing relative Fisher Information, as we shall see below. For a.c. measures, the characterization (25) already provides the property

$$\mathcal{I}(f dx|\mu) = 0 \quad \implies \quad f(q, p) = \tilde{f}(q) \exp\left(-\frac{p^2}{2m}\right).$$

This property holds more generally:

Lemma 2.2 (Zero relative Fisher Information implies Maxwellian). *If $\nu \in \mathcal{P}(\mathbb{R}^{2d})$ with $\mathcal{I}(\nu|\mu) = 0$, then there exists $\sigma \in \mathcal{P}(\mathbb{R}^d)$ such that*

$$\nu(dqdp) = Z^{-1} \exp\left(-\frac{p^2}{2m}\right) \sigma(dq)dp,$$

where $Z = \int_{\mathbb{R}^d} e^{-p^2/2m} dp$ is the normalization constant for the Maxwellian distribution.

Proof. From

$$\mathcal{I}(\nu|\mu) = \sup_{\varphi \in C_c^\infty(\mathbb{R}^{2d})} 2 \int_{\mathbb{R}^{2d}} \left(\Delta_p \varphi - \frac{p}{m} \cdot \nabla_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right) d\nu = 0 \quad (26)$$

we conclude upon disintegrating ν as $\nu(dqdp) = \sigma(dq)\nu_q(dp)$,

$$\text{for } \sigma\text{-a.e. } q: \quad \sup_{\phi \in C_c^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} \left(\Delta_p \phi - \frac{p}{m} \cdot \nabla_p \phi - \frac{1}{2} |\nabla_p \phi|^2 \right) \nu_q(dp) = 0.$$

By replacing ϕ by $\lambda\phi$, $\lambda > 0$, and taking $\lambda \rightarrow 0$ we find

$$\forall \phi \in C_c^\infty(\mathbb{R}^d) : \int_{\mathbb{R}^d} \left(\Delta_p \phi - \frac{p}{m} \cdot \nabla_p \phi \right) \nu_q(dp) = 0,$$

which is the weak form of an elliptic equation on \mathbb{R}^d with unique solution

$$\nu_q(dp) = \frac{1}{Z} \exp\left(-\frac{p^2}{2m}\right) dp.$$

This proves the lemma. \square

In the following theorem we give the central *a priori* estimate, in which free energy and relative Fisher Information are bounded from above by the rate functional and the relative entropy at initial time.

Theorem 2.3 (A priori bounds). *Fix $\gamma > 0$ and let $\rho \in C([0, T]; \mathcal{P}(\mathbb{R}^{2d}))$ with $\rho_t|_{t=0} =: \rho_0$ satisfy*

$$I^\gamma(\rho) < \infty, \quad \mathcal{F}(\rho_0) < \infty. \quad (27)$$

Then for any $t \in [0, T]$ we have

$$\mathcal{F}(\rho_t) + \frac{\gamma^2}{2} \int_0^t \mathcal{I}(\rho_s | \mu) ds \leq I^\gamma(\rho) + \mathcal{F}(\rho_0). \quad (28)$$

From (28) we obtain the separate inequality

$$\int_{\mathbb{R}^{2d}} H d\rho_t \leq \mathcal{F}(\rho_0) + I^\gamma(\rho) - \log \int_{\mathbb{R}^{2d}} e^{-H}. \quad (29)$$

This estimate will lead to a priori bounds in two ways. First, the bound on the free energy gives tightness estimates, and therefore compactness in space (Theorem 2.4); secondly, the relative Fisher Information is bounded by C/γ^2 and therefore vanishes in the limit $\gamma \rightarrow \infty$. This fact is used to prove that the limiting measure is Maxwellian (Lemma 2.5).

Proof. We give a heuristic motivation here; Appendix B contains a full proof. Given a trajectory ρ as in the theorem, note that by (20) ρ satisfies

$$\partial_t \rho_t = -\gamma \operatorname{div} \rho_t J \nabla (H + \psi * \rho_t) + \gamma^2 \left(\operatorname{div}_p \rho_t \frac{p}{m} + \Delta_p \rho_t \right) - \gamma \operatorname{div}_p \rho_t h_t.$$

We then formally calculate

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(\rho_t) &= \int_{\mathbb{R}^{2d}} [\log \rho_t + 1 + H + \psi * \rho_t] \left(-\gamma \operatorname{div} \rho_t J \nabla (H + \psi * \rho_t) + \gamma^2 \left(\operatorname{div}_p \rho_t \frac{p}{m} + \Delta_p \rho_t \right) \right. \\ &\quad \left. - \gamma \operatorname{div}_p \rho_t h_t \right) \\ &= -\gamma^2 \int_{\mathbb{R}^{2d}} \frac{1}{\rho_t} \left| \nabla_p \rho_t + \rho_t \frac{p}{m} \right|^2 + \gamma \int_{\mathbb{R}^{2d}} h_t \left(\nabla_p \rho_t + \rho_t \frac{p}{m} \right) \\ &\leq -\frac{\gamma^2}{2} \int_{\mathbb{R}^{2d}} \frac{1}{\rho_t} \left| \nabla_p \rho_t + \rho_t \frac{p}{m} \right|^2 + \frac{1}{2} \int_{\mathbb{R}^{2d}} \rho_t h_t^2, \end{aligned}$$

where the first $O(\gamma)$ term cancels because of the antisymmetry of J . After integration in time this latter expression yields (28).

For exact solutions of the VFP equation, i.e. when $I^\gamma(\rho) = 0$, this argument can be made rigorous following e.g. [BCS97]. However, the fairly low regularity of the right-hand side in (20) prevents these techniques from working. ‘Mild’ solutions, defined using the variation-of-constants formula and the Green function for the hypoelliptic operator, are not well-defined either, for the same reason: the term $\iint \nabla_p G \cdot h d\rho$ that appears in such an expression is generally not integrable. In the appendix we give a different proof, using the method of dual equations. \square

2.3 Coarse-graining and compactness

As we described in the introduction, in the overdamped limit $\gamma \rightarrow \infty$ we expect that ρ will resemble a Maxwellian distribution $Z^{-1} \exp(-p^2/2m) \sigma_t(dq)$, and that the q -dependent part σ will solve the Vlasov-Fokker-Planck equation (12). We will prove this statement using the method described in Section 1.1.

It would be natural to define ‘coarse-graining’ in this context as the projection $\xi(q, p) := q$, since that should eliminate the fast dynamics of p and focus on the slower dynamics of q . However, this choice fails: it completely decouples the dynamics of q from that of p , thereby preventing the noise in p from transferring to q . Following the lead of Kramers [Kra40], therefore, we define a slightly different coarse-graining map

$$\xi^\gamma : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d, \quad \xi^\gamma(q, p) := q + \frac{p}{\gamma}. \quad (30)$$

In the limit $\gamma \rightarrow \infty$, $\xi^\gamma \rightarrow \xi$ locally uniformly, recovering the projection onto the q -coordinate.

The theorem below gives the compactness properties of the solutions ρ^γ of the rescaled VFP equation that allow us to pass to the limit. There are two levels of compactness, a weaker one in the original space \mathbb{R}^{2d} , and a stronger one in the coarse-grained space $\mathbb{R}^d = \xi^\gamma(\mathbb{R}^{2d})$. This is similar to other multilevel compactness results as in e.g. [GOVW09].

Theorem 2.4 (Compactness). *Let a sequence $\rho^\gamma \in C([0, T]; \mathcal{P}(\mathbb{R}^{2d}))$ satisfy for a suitable constant $C > 0$ and every γ the estimate*

$$I^\gamma(\rho^\gamma) + \mathcal{F}(\rho_t^\gamma|_{t=0}) \leq C. \quad (31)$$

Then there exist a subsequence (not relabelled) such that

1. $\rho^\gamma \rightarrow \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ with respect to the narrow topology.
2. $\xi_{\#}^\gamma \rho^\gamma \rightarrow \xi_{\#} \rho$ in $C([0, T]; \mathcal{P}(\mathbb{R}^d))$ with respect to the uniform topology in time and narrow topology on $\mathcal{P}(\mathbb{R}^d)$.

For a.e. $t \in [0, T]$ the limit ρ_t satisfies

$$\mathcal{I}(\rho_t|\mu) = 0 \quad (32)$$

Proof. To prove part 1, note that the positivity of the convolution integral involving ψ and the free-energy-dissipation inequality (28) imply that $\mathcal{H}(\rho_t^\gamma | Z_H^{-1} e^{-H} dx)$ is bounded uniformly in t and γ . By an argument as in [ASZ09, Prop. 4.2] this implies that $\{\rho_t^\gamma : t \in [0, T], \gamma > 1\}$ is tight, upon which compactness in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ follows.

To prove (32) we remark that

$$0 \leq \sup_{\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}^{2d})} 2 \int_0^T \int_{\mathbb{R}^{2d}} \left[\Delta_p \varphi - \frac{p}{m} \nabla_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right] d\rho_t^\gamma dt \leq \int_0^T \mathcal{I}(\rho_t^\gamma|\mu) dt \leq \frac{C}{\gamma^2} \xrightarrow{\gamma \rightarrow \infty} 0,$$

and by passing to the limit on the left-hand side we find

$$\sup_{\varphi \in C_c^\infty(\mathbb{R} \times \mathbb{R}^{2d})} 2 \int_0^T \int_{\mathbb{R}^{2d}} \left[\Delta_p \varphi - \frac{p}{m} \nabla_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right] d\rho_t dt = 0.$$

By disintegrating ρ in time as $\rho(dt dq dp) = \rho_t(dq dp) dt$, we find that $\mathcal{I}(\rho_t|\mu) = 0$ for (Lebesgue-) almost all t .

We prove part 2 with the Arzelà-Ascoli theorem. For any $t \in [0, T]$ the sequence $\xi_{\#}^\gamma \rho_t^\gamma$ is tight, which follows from the tightness of ρ_t^γ proved above and the local uniform convergence $\xi^\gamma \rightarrow \xi$ (see e.g. [AGS08, Lemma 5.2.1]).

To prove equicontinuity we will show

$$\sup_{\gamma > 1} \sup_{t \in [0, T-h]} \sup_{\substack{\varphi \in C_c^2(\mathbb{R}^d) \\ \|\varphi\|_{C^2(\mathbb{R}^d)} \leq 1}} \int_{\mathbb{R}^d} \varphi(\xi_{\#}^\gamma \rho_{t+h}^\gamma - \xi_{\#}^\gamma \rho_t^\gamma) \xrightarrow{h \rightarrow 0} 0. \quad (33)$$

Note that the boundedness of the rate functional, definition (20), and tightness of ρ^γ imply that there exists some $h^\gamma \in L^2(0, T; L^2_{\#}(\rho_t^\gamma))$ with

$$\partial_t \rho_t^\gamma = (\mathcal{L}_{\rho_t^\gamma})^* \rho_t^\gamma - \gamma \operatorname{div}_p(\rho_t^\gamma h_t^\gamma). \quad (34)$$

in duality with $C_b^2(\mathbb{R}^{2d})$. Therefore for any $f \in C_b^2(\mathbb{R}^{2d})$ we have in the sense of distributions on $[0, T]$,

$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} f \rho_t^\gamma = \int_{\mathbb{R}^{2d}} \left(\gamma \frac{p}{m} \cdot \nabla_q f - \gamma \nabla_q V \cdot \nabla_p f - \gamma \nabla_p f \cdot (\nabla_q \psi * \rho^\gamma) - \gamma^2 \frac{p}{m} \cdot \nabla_p f + \gamma^2 \Delta_p f + \gamma \nabla_p f \cdot h_t^\gamma \right) d\rho_t^\gamma.$$

To prove (33), make the choice $f = \varphi \circ \xi^\gamma$ for $\varphi \in C_c^2(\mathbb{R}^d)$ and integrate over $[t, t+h]$ to arrive at

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(\xi_{\#}^\gamma \rho_{t+h}^\gamma - \xi_{\#}^\gamma \rho_t^\gamma) &= \int_t^{t+h} \int_{\mathbb{R}^{2d}} \left(-\nabla V(q) \cdot \nabla \varphi \left(q + \frac{p}{\gamma} \right) - (\nabla_q \psi * \rho_s^\gamma)(q) \cdot \nabla \varphi \left(q + \frac{p}{\gamma} \right) \right. \\ &\quad \left. + \Delta \varphi \left(q + \frac{p}{\gamma} \right) + \nabla \varphi \left(q + \frac{p}{\gamma} \right) \cdot h_s^\gamma(q, p) \right) d\rho_s^\gamma ds. \end{aligned}$$

We estimate the first term on the right hand side by using Hölder's inequality and growth condition (V1),

$$\begin{aligned} \left| \int_t^{t+h} \int_{\mathbb{R}^{2d}} \nabla V(q) \cdot \nabla \varphi \left(q + \frac{p}{\gamma} \right) d\rho_s^\gamma ds \right| &\leq \|\nabla \varphi\|_\infty \sqrt{h} \left(\int_t^{t+h} \int_{\mathbb{R}^{2d}} |\nabla V(q)|^2 d\rho_s^\gamma ds \right)^{1/2} \\ &\leq \|\nabla \varphi\|_\infty \sqrt{h} \left(\int_t^{t+h} \int_{\mathbb{R}^{2d}} C(1+V(q)) \rho_s^\gamma ds \right)^{1/2} \leq \tilde{C} \|\nabla \varphi\|_\infty h \end{aligned}$$

where the last inequality follows from the free-energy-dissipation inequality (28). For the second term we use $|\nabla_q \psi * \rho_s^\gamma| \leq \|\nabla_q \psi\|_\infty$ and the last term is estimated by Hölder's inequality,

$$\begin{aligned} \left| \int_t^{t+h} \int_{\mathbb{R}^{2d}} \nabla \varphi \left(q + \frac{p}{\gamma} \right) h_s^\gamma(q, p) d\rho_s^\gamma ds \right| &\leq \|\nabla \varphi\|_\infty \sqrt{h} \left(\int_t^{t+h} \int_{\mathbb{R}^{2d}} |h_s^\gamma|^2 d\rho_s^\gamma ds \right)^{1/2} \\ &\leq \|\nabla \varphi\|_\infty \sqrt{h} (2I^\gamma(\rho^\gamma))^{\frac{1}{2}} \leq C \|\nabla \varphi\|_\infty \sqrt{h}. \end{aligned}$$

To sum up we have

$$\left| \int_{\mathbb{R}^d} \varphi(\xi_{\#}^\gamma \rho_{t+h}^\gamma - \xi_{\#}^\gamma \rho_t^\gamma) \right| \leq C \sqrt{h} \xrightarrow{h \rightarrow 0} 0,$$

where C is independent of t and γ .

Thus by the Arzelà-Ascoli theorem there exists a $\nu \in C([0, T]; \mathcal{P}(\mathbb{R}^d))$ such that $\xi_{\#}^\gamma \rho^\gamma \rightarrow \nu$ with respect to uniform topology in time and narrow topology on $\mathcal{P}(\mathbb{R}^d)$. Since $\rho^\gamma \rightarrow \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ and $\xi^\gamma \rightarrow \xi$ locally uniformly, we have $\xi_{\#}^\gamma \rho^\gamma \rightarrow \xi_{\#} \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^d)$ (again using [AGS08, Lemma 5.2.1]), implying that $\nu = \xi_{\#} \rho$. This concludes the proof of Theorem 2.4. \square

2.4 Local equilibrium

A central step in any coarse-graining method is the treatment of the information that is 'lost' upon coarse-graining. The lemma below uses the a priori estimate (28) to reconstruct this information, which for this system means showing that ρ^γ becomes Maxwellian in p as $\gamma \rightarrow \infty$.

Lemma 2.5 (Local equilibrium). *Under the same conditions as in Theorem 2.4 let us assume that $\rho^\gamma \rightarrow \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ with respect to the narrow topology. Then there exists $\sigma \in \mathcal{M}([0, T] \times \mathbb{R}^d)$, $\sigma(dt dq) = \sigma_t(dq) dt$, such that for almost all $t \in [0, T]$,*

$$\rho_t(dq dp) = Z^{-1} \exp\left(-\frac{p^2}{2m}\right) \sigma_t(dq) dp, \quad (35)$$

where $Z = \int_{\mathbb{R}^d} e^{-p^2/2m} dp$ is the normalization constant for the Maxwellian distribution. Furthermore $\xi_{\#}^\gamma \rho_t^\gamma \rightarrow \sigma_t$ narrowly for every $t \in [0, T]$.

Proof. Since $\rho^\gamma \rightarrow \rho$ narrowly in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$, the limit ρ also has the disintegration structure $\rho(dtdpdq) = \rho_t(dpdq)dt$, with $\rho_t \in \mathcal{P}(\mathbb{R}^{2d})$. From the *a priori* estimate (28) and the duality definition of \mathcal{I} we have $\mathcal{I}(\rho_t|\mu) = 0$ for almost all t , and the characterization (35) then follows from Lemma 2.2. The compactness results in Theorem 2.4 imply that $\xi_{\#}^\gamma \rho_t^\gamma \rightarrow \xi_{\#} \rho_t = \sigma_t$ for all $t \in [0, T]$. \square

2.5 Liminf inequality

The final step in the variational technique is proving an appropriate liminf inequality which also provides the structure of the limiting coarse-grained evolution. The following theorem makes this step rigorous.

Define the (limiting) functional $I : C([0, T]; \mathcal{P}(\mathbb{R}^d)) \rightarrow \mathbb{R}$ by

$$I(\sigma) := \sup_{g \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^d)} \int_{\mathbb{R}^d} g_T d\sigma_T - \int_{\mathbb{R}^d} g_0 d\sigma_0 - \int_0^T \int_{\mathbb{R}^d} \left(\partial_t g - \nabla V \cdot \nabla g - (\nabla \psi * \sigma) \cdot \nabla g + \Delta g \right) d\sigma_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla g|^2 d\sigma_t dt. \quad (36)$$

Note that $I \geq 0$ (since $g = 0$ is admissible); we have the equivalence

$$I(\sigma) = 0 \iff \partial_t \sigma = \operatorname{div} \sigma \nabla V(q) + \operatorname{div} \sigma (\nabla \psi * \sigma) + \Delta \sigma \quad \text{in } [0, T] \times \mathbb{R}^d.$$

Theorem 2.6 (Liminf inequality). *Under the same conditions as in Theorem 2.4 we assume that $\rho^\gamma \rightarrow \rho$ narrowly in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ and $\xi_{\#}^\gamma \rho^\gamma \rightarrow \xi_{\#} \rho \equiv \sigma$ in $C([0, T]; \mathcal{P}(\mathbb{R}^d))$. Then*

$$\liminf_{\gamma \rightarrow \infty} I^\gamma(\rho^\gamma) \geq I(\sigma).$$

Proof. Write the large deviation rate functional $I^\gamma : C([0, T]; \mathcal{P}(\mathbb{R}^{2d})) \rightarrow \mathbb{R}$ in (19) as

$$I^\gamma(\rho) = \sup_{f \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^{2d})} \mathcal{J}^\gamma(\rho, f), \quad (37)$$

where

$$\begin{aligned} \mathcal{J}^\gamma(\rho, f) = & \int_{\mathbb{R}^{2d}} f_T d\rho_T - \int_{\mathbb{R}^{2d}} f_0 d\rho_0 - \int_0^T \int_{\mathbb{R}^{2d}} \left(\partial_t f + \gamma \frac{p}{m} \cdot \nabla_q f - \gamma \nabla_q V \cdot \nabla_p f - \gamma \nabla_p f \cdot (\nabla_q \psi * \rho_t) \right. \\ & \left. - \gamma^2 \frac{p}{m} \cdot \nabla_p f + \gamma^2 \Delta_p f \right) d\rho_t dt - \frac{\gamma^2}{2} \int_0^T \int_{\mathbb{R}^{2d}} |\nabla_p f|^2 d\rho_t dt. \end{aligned}$$

Define $\mathcal{A} := \{f = g \circ \xi^\gamma \text{ with } g \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^d)\}$. Then we have

$$I^\gamma(\rho^\gamma) \geq \sup_{f \in \mathcal{A}} \mathcal{J}^\gamma(\rho^\gamma, f),$$

and

$$\begin{aligned} \mathcal{J}^\gamma(\rho^\gamma, g \circ \xi^\gamma) = & \int_{\mathbb{R}^{2d}} g_T \circ \xi^\gamma d\rho_T^\gamma - \int_{\mathbb{R}^{2d}} g_0 \circ \xi^\gamma d\rho_0^\gamma - \int_0^T \int_{\mathbb{R}^{2d}} \left[\partial_t (g \circ \xi^\gamma) - \nabla_q V(q) \cdot \nabla g \left(q + \frac{p}{\gamma} \right) \right. \\ & \left. + \Delta g \left(q + \frac{p}{\gamma} \right) - \nabla g \left(q + \frac{p}{\gamma} \right) \cdot (\nabla_q \psi * \rho_t^\gamma)(q) \right] d\rho_t^\gamma dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2d}} |\nabla (g \circ \xi^\gamma)|^2 d\rho_t^\gamma dt. \quad (38) \end{aligned}$$

Note how the specific dependence of $\xi^\gamma(q, p) = q + p/\gamma$ on γ has caused the coefficients γ and γ^2 in the expression above to vanish. Adding and subtracting $\nabla V(q+p/\gamma) \cdot \nabla g(q+p/\gamma)$ in (38) and defining $\hat{\rho}^\gamma := \xi_{\#}^\gamma \rho^\gamma$, \mathcal{J}^γ can be rewritten as

$$\begin{aligned} \mathcal{J}^\gamma(\rho, g \circ \xi^\gamma) &= \int_{\mathbb{R}^d} g_T d\hat{\rho}_T^\gamma - \int_{\mathbb{R}^d} g_0 d\hat{\rho}_0^\gamma - \int_0^T \int_{\mathbb{R}^d} (\partial_t g - \nabla V \cdot \nabla g + \Delta g)(\zeta) \hat{\rho}_t^\gamma(d\zeta) dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^d} |\nabla g|^2 d\hat{\rho}_t^\gamma dt \\ &\quad - \int_0^T \int_{\mathbb{R}^{2d}} \left(\nabla V \left(q + \frac{p}{\gamma} \right) - \nabla V(q) \right) \cdot \nabla g \left(q + \frac{p}{\gamma} \right) d\rho_t^\gamma dt + \int_0^T \int_{\mathbb{R}^{2d}} \nabla g \left(q + \frac{p}{\gamma} \right) \cdot (\nabla_q \psi * \rho_t^\gamma)(q) d\rho_t^\gamma dt. \end{aligned} \quad (39)$$

We now show that (39) converges to the right-hand side of (36), term by term. Since $\xi_{\#}^\gamma \rho^\gamma \rightarrow \xi_{\#} \rho = \sigma$ narrowly in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ and $g \in C_b^2(\mathbb{R} \times \mathbb{R}^d)$ we have

$$\int_0^T \int_{\mathbb{R}^d} \left(\partial_t g - \nabla V \cdot \nabla g + \Delta g + \frac{1}{2} |\nabla g|^2 \right) d\hat{\rho}_t^\gamma dt \xrightarrow{\gamma \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} \left(\partial_t g - \nabla V \cdot \nabla g + \Delta g + \frac{1}{2} |\nabla g|^2 \right) d\sigma_t dt.$$

Taylor expansion of ∇V around q and estimate (29) give

$$\begin{aligned} \left| \int_0^T \int_{\mathbb{R}^{2d}} \left(\nabla V \left(q + \frac{p}{\gamma} \right) - \nabla V(q) \right) \cdot \nabla g \left(q + \frac{p}{\gamma} \right) d\rho_t^\gamma dt \right| &\leq \\ &\leq \|D^2 V\|_\infty \|\nabla g\|_\infty \sqrt{T} \left(\int_0^T \int_{\mathbb{R}^{2d}} \frac{p^2}{\gamma^2} d\rho_t^\gamma dt \right)^{1/2} \leq \frac{C}{\gamma} \xrightarrow{\gamma \rightarrow \infty} 0. \end{aligned}$$

Adding and subtracting $\nabla g(q) \cdot (\nabla_q \psi * \rho_t^\gamma)(q)$ in (39) we find

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^{2d}} \nabla g \left(q + \frac{p}{\gamma} \right) \cdot (\nabla_q \psi * \rho_t^\gamma)(q) d\rho_t^\gamma dt &= \int_0^T \int_{\mathbb{R}^{2d}} \nabla g(q) \cdot (\nabla_q \psi * \rho_t^\gamma)(q) d\rho_t^\gamma dt \\ &\quad + \int_0^T \int_{\mathbb{R}^{2d}} \left[\nabla g \left(q + \frac{p}{\gamma} \right) - \nabla g(q) \right] \cdot (\nabla_q \psi * \rho_t^\gamma)(q) d\rho_t^\gamma dt. \end{aligned}$$

Since $\rho^\gamma \rightarrow \rho$ we have $\rho^\gamma \otimes \rho^\gamma \rightarrow \rho \otimes \rho$ and therefore passing to the limit in the first term and using the local-equilibrium characterization of Lemma 2.5, we obtain

$$\int_0^T \int_{\mathbb{R}^{2d}} \nabla g(q) \cdot (\nabla_q \psi * \rho^\gamma)(q) d\rho_t^\gamma dt \xrightarrow{\gamma \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} \nabla g \cdot (\nabla \psi * \sigma) d\sigma_t dt.$$

For the second term we calculate

$$\begin{aligned} &\left| \int_0^T \int_{\mathbb{R}^{2d}} \left[\nabla g \left(q + \frac{p}{\gamma} \right) - \nabla g(q) \right] \cdot (\nabla_q \psi * \rho_t^\gamma)(q) d\rho_t^\gamma dt \right| \\ &\leq \|D^2 g\|_\infty \|\nabla_q \psi\|_\infty \sqrt{T} \left(\int_0^T \int_{\mathbb{R}^{2d}} \frac{p^2}{\gamma^2} d\rho_t^\gamma dt \right)^{1/2} \leq \frac{C}{\gamma} \xrightarrow{\gamma \rightarrow \infty} 0. \end{aligned}$$

Therefore

$$\int_0^T \int_{\mathbb{R}^{2d}} \nabla g \left(q + \frac{p}{\gamma} \right) \cdot (\nabla_q \psi * \rho^\gamma)(q) d\rho_t^\gamma dt \xrightarrow{\gamma \rightarrow \infty} \int_0^T \int_{\mathbb{R}^d} \nabla g \cdot (\nabla \psi * \sigma) d\sigma_t dt.$$

□

2.6 Discussion

The ingredients of the convergence proof above are, as mentioned before, (a) a compactness result, (b) a local-equilibrium result, and (c) a liminf inequality. All three follow from the large-deviation structure, through the rate functional I^γ . We now comment on these.

Compactness. Compactness in the sense of measures is, both for ρ^γ and for $\xi_{\#}^\gamma \rho^\gamma$, a simple consequence of the confinement provided by the growth of H . In Theorem 2.4 we provide a stronger statement for $\xi_{\#}^\gamma \rho^\gamma$, by showing continuity in time, in order for the limiting functional $I(\sigma)$ in (36) to be well defined. This continuity depends on the boundedness of I^γ .

Local equilibrium. The local-equilibrium statement depends crucially on the structure of I^γ , and more specifically on the large coefficient γ^2 multiplying the derivatives in p . This coefficient also ends up as a prefactor of the relative Fisher Information in the *a priori* estimate (28), and through this estimate it drives the local-equilibrium result.

Liminf inequality. As remarked in the introduction, the duality structure of I^γ is the key to the liminf inequality, as it allows for relatively weak convergence of ρ^γ and $\xi_{\#}^\gamma \rho^\gamma$. The role of the local equilibrium is to allow us to replace the p -dependence in some of the integrals by the Maxwellian dependence, and therefore to reduce all terms to dependence on the macroscopic information $\xi_{\#}^\gamma \rho^\gamma$ only.

As we have shown, the choice of the coarse-graining map has the advantage that it has caused the (large) coefficients γ and γ^2 in the expression of the rate functionals to vanish. In other words, it cancels out the inertial effects and transforms a Laplacian in p variable to a Laplacian in the coarse-grained variable while rescaling it to be of order 1. The choice $\xi(q, p) = q$, on the other hand, would lose too much information by completely discarding the diffusion.

3 Diffusion on a Graph, $d = 1$

In this section we derive the small-noise limit of a randomly perturbed Hamiltonian system, which corresponds to passing to the limit $\varepsilon \rightarrow 0$ in (14). In terms of a rescaled time, in order to focus on the time scale of the noise, equation (14) becomes

$$\partial_t \rho^\varepsilon = -\frac{1}{\varepsilon} \operatorname{div}(\rho^\varepsilon J \nabla H) + \Delta_p \rho^\varepsilon. \quad (40)$$

Here $\rho^\varepsilon \in C([0, T], \mathcal{P}(\mathbb{R}^2))$, $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is again the canonical symplectic matrix, Δ_p is the Laplacian in the p -direction, and the equation holds in the sense of distributions. The Hamiltonian $H \in C^2(\mathbb{R}^{2d}; \mathbb{R})$ is again defined by $H(q, p) = p^2/2m + V(q)$ for some potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$. We make the following assumptions (that we formulate on H for convenience):

- (A1) $H \geq 0$, and H is coercive, i.e. $H(x) \xrightarrow{|x| \rightarrow \infty} \infty$;
- (A2) $|\nabla H|, |\Delta H|, |\nabla_p H|^2 \leq C(1 + H)$;
- (A3) H has a finite number of non-degenerate (i.e. non-singular Hessian) saddle points O_1, \dots, O_n with $H(O_i) \neq H(O_j)$ for every $i, j \in \{1, \dots, n\}$, $i \neq j$.

As explained in the introduction, and in contrast to the VFP equation of the previous section, equation (40) has two equally valid interpretations: as a PDE in its own right, or as the Fokker-Planck (forward Kolmogorov) equation of the stochastic process

$$X^\varepsilon = \begin{pmatrix} Q^\varepsilon \\ P^\varepsilon \end{pmatrix}, \quad dX_t^\varepsilon = \frac{1}{\varepsilon} J \nabla H(X_t^\varepsilon) dt + \sqrt{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t. \quad (41)$$

For the sequel we will think of ρ^ε as the law of the process X_t^ε ; although this is not strictly necessary, it helps in illustrating the ideas.

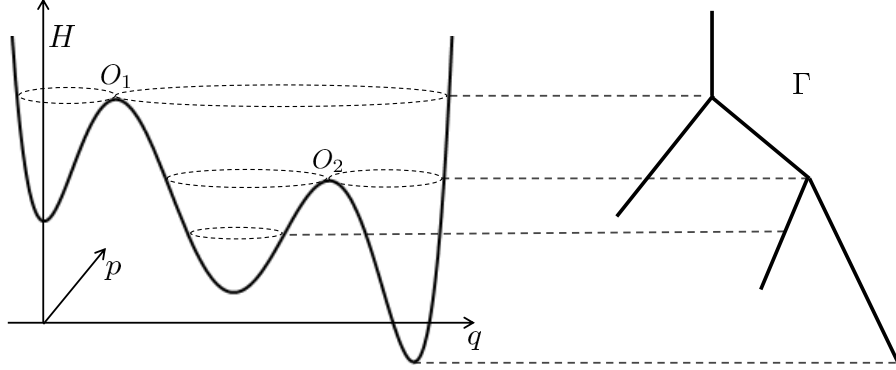


Figure 3: Left: Hamiltonian $\mathbb{R}^2 \ni (q, p) \mapsto H(q, p)$, Right: Graph Γ

3.1 Construction of the graph Γ

As mentioned in the introduction, the dynamics of (40) has two time scales when $0 < \varepsilon \ll 1$, a fast and a slow one. The fast time scale, of scale ε , is described by the (deterministic) equation

$$\dot{x} = \frac{1}{\varepsilon} J \nabla H(x) \quad \text{in } \mathbb{R}^2, \quad (42)$$

whereas the slow time scale, of order 1, is generated by the noise term.

The solutions of (42) follow level sets of H . There exist three types of such solutions: stationary ones, periodic orbits, and homoclinic orbits. Stationary solutions of (42) correspond to stationary points of H (where $\nabla H = 0$); periodic orbits to connected components of level sets along which $\nabla H \neq 0$; and homoclinic orbits to components of level sets of H that are terminated on each end by a stationary point. Since we have assumed in (A3) that there is at most one stationary point in each level sets, heteroclinic orbits do not exist, and the orbits necessarily connect a stationary point with itself.

Looking ahead towards coarse-graining, we define Γ to be the set of all connected components of level sets of H , and we identify Γ with a union of one-dimensional line segments, as shown in Figure 3. Each periodic orbit corresponds to an interior point of one of the edges of Γ ; the vertices of Γ correspond to connected components of level sets containing a stationary point of H . Each saddle point O corresponds to a vertex connected by three edges.

For practical purposes we also introduce a coordinate system on Γ . We represent the edges by closed intervals $I_k \subset \mathbb{R}$, and number them with numbers $k = 1, 2, \dots, n$; the pair (h, k) is then a coordinate for a point $\gamma \in \Gamma$, if k is the index of the edge containing γ , and h the value of H on the level set represented by γ . For a vertex $O \in \Gamma$, we write $O \sim I_k$ if O is at one end of edge I_k ; we use the shorthand notation \pm_{k_j} to mean 1 if O_j is at the upper end of I_k , and -1 in the other case. Note that if $O \sim I_{k_1}$, $O \sim I_{k_2}$ and $O \sim I_{k_3}$ and h_0 is the value of H at the point corresponding to O , then the coordinates (h_0, k_1) , (h_0, k_2) and (h_0, k_3) correspond to the same point O . With a slight abuse of notation, we also define the function $k : \mathbb{R}^2 \rightarrow \{1, \dots, n\}$ as the index of the edge $I_k \subset \Gamma$ corresponding to the component containing (q, p) .

The rigorous construction of the graph Γ and the topology on it has been done several times [FW93, FW94, BvR14]; for our purposes it suffices to note that (a) inside each edge, the usual topology and geometry of \mathbb{R}^1 apply, and (b) across the whole graph there is a natural concept of distance, and therefore of continuity. It will be practical to think of functions $f : \Gamma \rightarrow \mathbb{R}$ as defined on the disjoint union $\sqcup_k I_k$. A function $f : \Gamma \rightarrow \mathbb{R}$ is then called well-defined if it is a single-valued function on Γ (i.e., it takes the same value on those vertices that are multiply represented). A well-defined function $f : \Gamma \rightarrow \mathbb{R}$ is *continuous* if $f|_{I_k} \in C(I_k)$ for every k .

We also define a concept of *differentiability* of a function $f : \Gamma \rightarrow \mathbb{R}$. A *subgraph* of Γ is defined as any union of edges such that each interior vertex connects exactly two edges, one from above and one from below—i.e., a subtree without bifurcations. A continuous function on Γ is called differentiable on Γ if it is differentiable on each of its subgraphs.

Finally, in order to integrate over Γ , we write $d\gamma$ for the measure on Γ which is defined on each I_k as the local Lebesgue measure dh . Whenever we write \int_Γ , this should be interpreted as $\sum_k \int_{I_k}$.

3.2 Adding noise: diffusion on the graph

In the noisy evolution (41), for small but finite $\varepsilon > 0$, the evolution follows fast trajectories that nearly coincide with the level sets of H ; the noise breaks the conservation of H , and causes a slower drift of X_t across the levels of H . In order to remove the fast deterministic dynamics, we now define the coarse-graining map as

$$\xi : \mathbb{R}^2 \rightarrow \Gamma, \quad \xi(q, p) := (H(q, p), k(q, p)), \quad (43)$$

where the mapping $k : \mathbb{R}^2 \rightarrow \{1, \dots, n\}$ indexes the edges of the graph, as above.

We now consider the process $\xi(X_t^\varepsilon)$, which contains no fast dynamics. For each finite $\varepsilon > 0$, $\xi(X_t^\varepsilon)$ is not a Markov process; but as $\varepsilon \rightarrow 0$, the fast movement should result in a form of averaging, such that the influence of the missing information vanishes; then the limit process is a diffusion on the graph Γ .

The results of this section are stated and proved in terms of the corresponding objects ρ^ε and $\hat{\rho}^\varepsilon$, where $\hat{\rho}^\varepsilon$ is the push-forward

$$\hat{\rho}^\varepsilon := \xi_\# \rho^\varepsilon, \quad (44)$$

as explained in Section 1.1, and similar to Section 2. The corresponding statement about ρ^ε and $\hat{\rho}^\varepsilon$ is that $\hat{\rho}^\varepsilon$ should converge to some $\hat{\rho}$, which in the limit satisfies a (convection-) diffusion equation on Γ . Theorems 3.1 and 3.5 make this statement precise.

3.3 Compactness

As in the case of the VFP equation, equation (40) has a free energy, which in this case is simply the relative entropy with respect to the Lebesgue measure \mathcal{L} :

$$\mathcal{F}(\rho) = \mathcal{H}(\rho | \mathcal{L}). \quad (45)$$

The corresponding ‘relative’ Fisher Information is the same as the usual Fisher Information,

$$\mathcal{I}(\rho | \mathcal{L}) = \sup_{\varphi \in C_c^\infty(\mathbb{R}^2)} 2 \int_{\mathbb{R}^2} \left[\Delta_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right] d\nu,$$

and satisfies for $\rho = f\mathcal{L}$,

$$\mathcal{I}(f\mathcal{L} | \mathcal{L}) = \int_{\mathbb{R}^2} |\nabla_p \log f|^2 f dqdp,$$

whenever this is finite.

The large deviation functional $I^\varepsilon : C([0, T]; \mathcal{P}(\mathbb{R}^2)) \rightarrow \mathbb{R}$ is given by

$$I^\varepsilon(\rho) = \sup_{f \in C_c^{1,2}(\mathbb{R} \times \mathbb{R}^2)} \left[\int_{\mathbb{R}^2} f_T d\rho_T - \int_{\mathbb{R}^2} f_0 d\rho_0 - \int_0^T \int_{\mathbb{R}^2} \left(\partial_t f + \frac{1}{\varepsilon} J \nabla H \cdot \nabla f + \Delta_p f \right) d\rho_t dt - \frac{1}{2} \int_0^T \int_{\mathbb{R}^2} |\nabla_p f|^2 d\rho_t dt \right]. \quad (46)$$

For fixed $\varepsilon > 0$, ρ^ε solves (40) iff $I^\varepsilon(\rho^\varepsilon) = 0$.

The following theorem states the relevant *a priori* estimates and the ensuing compactness properties for any sequence ρ^ε with $\sup_\varepsilon I^\varepsilon(\rho^\varepsilon) < \infty$.

Theorem 3.1 (A priori estimates and compactness). *Let $\rho^\varepsilon \in C([0, T]; \mathcal{P}(\mathbb{R}^2))$ with $\rho_t^\varepsilon|_{t=0} =: \rho_0^\varepsilon$ satisfy for a suitable constant $C > 0$ and all $\varepsilon > 0$*

$$I^\varepsilon(\rho^\varepsilon) + \mathcal{F}(\rho_0^\varepsilon) \leq C.$$

Then for any $t \in [0, T]$ we have

$$\mathcal{F}(\rho_t^\varepsilon) + \frac{1}{2} \int_0^t \mathcal{I}(\rho_s^\varepsilon | \mathcal{L}) ds \leq I^\varepsilon(\rho^\varepsilon) + \mathcal{F}(\rho_0^\varepsilon) \leq C. \quad (47)$$

Moreover, there exist subsequences (not relabelled) such that

1. $\rho^\varepsilon \rightarrow \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^2)$ in the narrow topology;
2. $\hat{\rho}^\varepsilon \rightarrow \hat{\rho} = \xi_{\#}\rho$ in $C([0, T]; \mathcal{P}(\Gamma))$ with respect to the uniform topology in time and narrow topology on $\mathcal{P}(\Gamma)$.

Finally, we have the estimate

$$\mathcal{F}(\rho_t) + \frac{1}{2} \int_0^t \mathcal{I}(\rho_s | \mathcal{L}) ds \leq C \quad \text{for all } t \in [0, T].$$

The proof of this theorem follows along the lines of the proofs of Theorems 2.3 and 2.4, and we omit it here.

Note that the estimate (47) implies that $\mathcal{H}(\rho_t^\varepsilon | \mathcal{L})$ is finite for all t , and therefore ρ_t^ε is Lebesgue absolutely continuous. We will often therefore write $\rho_t^\varepsilon(x)$ for the Lebesgue density of ρ_t^ε .

3.4 Local equilibrium

Theorem 3.1 states that ρ^ε converges narrowly on $[0, T] \times \mathbb{R}^2$ to some ρ . In fact we need a stronger statement, in which the behaviour of ρ on each connected component of H is fully determined by the limit $\hat{\rho}$.

Lemma 3.2 below makes this statement precise. Before proceeding we define $T : \Gamma \rightarrow \mathbb{R}$ as

$$T(\gamma) := \int_{\xi^{-1}(\gamma)} \frac{\mathcal{H}^1(dx)}{|\nabla H(x)|}, \quad (48)$$

where \mathcal{H}^1 is the one-dimensional Hausdorff measure. T has a natural interpretation as the period of the periodic orbit of the deterministic equation (42) corresponding to γ . When γ is an interior vertex, such that the orbit is homoclinic, not periodic, $T(\gamma) = +\infty$. T also has a second natural interpretation: the measure $T(\gamma)d\gamma = T(h, k)dh$ on Γ is the push-forward under ξ of the Lebesgue measure on \mathbb{R}^2 , and the measure $T(\gamma)d\gamma$ therefore appears in various places.

Lemma 3.2 (Local Equilibrium). *Under the assumptions in Theorem 3.1, assume that $\rho^\varepsilon \rightarrow \rho$ in $\mathcal{M}([0, T] \times \mathbb{R}^{2d})$ with respect to the narrow topology. Let $\hat{\rho}$ be the push-forward $\xi_{\#}\rho$ of the limit ρ , as above.*

Then for a.e. t , the limit ρ_t is absolutely continuous with respect to the Lebesgue measure, $\hat{\rho}_t$ is absolutely continuous with respect to the measure $T(\gamma)d\gamma$, where $T(\gamma)$ is defined in (48). Writing

$$\rho_t(dx) = \rho_t(x)dx \quad \text{and} \quad \hat{\rho}_t(d\gamma) = \alpha_t(\gamma)T(\gamma)d\gamma,$$

we have

$$\rho_t(x) = \alpha_t(\xi(x)) \quad \text{for almost all } x \in \mathbb{R}^2 \text{ and } t \in [0, T]. \quad (49)$$

Proof. From the boundedness of $I^\varepsilon(\rho^\varepsilon)$ and the narrow convergence $\rho^\varepsilon \rightarrow \rho$ we find, passing to the limit in the rate functional (46), for any $f \in C_c^{1,2}(\mathbb{R} \times \mathbb{R}^2)$

$$\int_0^T \int_{\mathbb{R}^2} J \nabla H \cdot \nabla f d\rho_t dt = 0. \quad (50)$$

Now choose any $\varphi \in C_c^2([0, T] \times \mathbb{R}^2)$ and any $\zeta \in C_b^2(\Gamma)$ such that ζ is constant in a neighbourhood of each vertex; then the function $f(t, x) = \zeta(\xi(x))\varphi(t, x)$ is well-defined and in $C_c^2([0, T] \times \mathbb{R}^2)$. We substitute this special function in (50); since $J\nabla H\nabla(\zeta \circ \xi) = 0$, we have $J\nabla H\nabla f = (\zeta \circ \xi)J\nabla H\nabla\varphi$. Applying the disintegration theorem to ρ , writing $\rho_t(dx) = \hat{\rho}_t(d\gamma)\tilde{\rho}_t(dx|\gamma)$ with $\text{supp } \tilde{\rho}_t(\cdot|\gamma) \subset \xi^{-1}(\gamma)$, we obtain

$$0 = \int_0^T \int_{\Gamma} \zeta(\gamma)\hat{\rho}_t(d\gamma) \int_{\xi^{-1}(\gamma)} \nabla\varphi \cdot \frac{J\nabla H}{|\nabla H|} |\nabla H| \tilde{\rho}_t(\cdot|\gamma) d\mathcal{H}^1 = \int_0^T \int_{\Gamma} \zeta(\gamma)\hat{\rho}_t(d\gamma) \int_{\xi^{-1}(\gamma)} \partial_{\tau}\varphi |\nabla H| \tilde{\rho}_t(\cdot|\gamma) d\mathcal{H}^1 dt,$$

where ∂_{τ} is the tangential derivative. By varying ζ and φ we conclude that for $\hat{\rho}$ -almost every (γ, t) , $|\nabla H|\tilde{\rho}_t(\cdot|\gamma) = C_{\gamma,t}$ for some γ, t -dependent constant $C_{\gamma,t} > 0$, and since $\tilde{\rho}$ is normalized, we find that

$$\text{for } \hat{\rho}\text{-a.e. } (\gamma, t) : \tilde{\rho}_t(dx|\gamma) = \frac{1}{T(\gamma)|\nabla H(x)|} \mathcal{H}^1|_{\xi^{-1}(\gamma)}(dx). \quad (51)$$

This also implies that $\tilde{\rho}_t(\cdot|\gamma)$ is in fact t -independent.

For measurable f we now compare the two relations

$$\begin{aligned} \int_{\mathbb{R}^2} f d\rho_t &= \int_{\mathbb{R}^2} f(y)\rho_t(y) dy &= \int_{\Gamma} d\gamma \int_{\xi^{-1}(\gamma)} \frac{f(y)}{|\nabla H(y)|} \rho_t(y) \mathcal{H}^1(dy) \\ \int_{\mathbb{R}^2} f d\rho_t &= \int_{\Gamma} \hat{\rho}_t(d\gamma) \int_{\xi^{-1}(\gamma)} f(y)\tilde{\rho}(dy|\gamma) &= \int_{\Gamma} \frac{\hat{\rho}_t(d\gamma)}{T(\gamma)} \int_{\xi^{-1}(\gamma)} \frac{f(y)}{|\nabla H(y)|} \mathcal{H}^1(dy) \end{aligned}$$

where we have used the co-area formula in the first line and (51) in the second one. Since f was arbitrary, (49) follows for almost all t . \square

3.5 Continuity of ρ and $\hat{\rho}$

As a consequence of the local-equilibrium property (49) and the boundedness of the Fisher Information, we will show in the following that ρ and its push-forward $\hat{\rho}$ satisfy an important continuity property. We first motivate this property heuristically.

The local-equilibrium result Lemma 3.2 states that the limit measure ρ depends on x only through $\xi(x)$. Take any measure $\rho \in \mathcal{P}(\mathbb{R}^{2d})$ of that form, i.e. $\rho(dx) = f(\xi(x))dx$, with finite free energy and finite relative Fisher Information. Setting $f = f \circ \xi$, by Lemma 2.1, $\nabla_p f$ is well-defined and locally integrable.

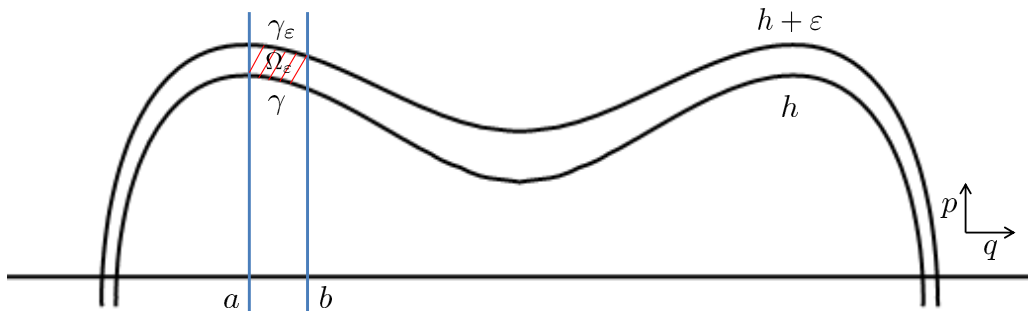


Figure 4: Section Ω in which $H^{-1}(h)$ is transverse to p .

Consider a section Ω_{ε} of the (q, p) -plane as shown in Figure 4, bounded by $q = a$ and $q = b$ and level sets $H = h$ and $H = h + \varepsilon$. The top and bottom boundaries γ and γ_{ε} correspond to elements of Γ that we also call γ and γ_{ε} ; they might be part of the same edge k of the graph, or they might belong to different edges. As $\varepsilon \rightarrow 0$, γ_{ε} converges to γ .

By simple integration we find that

$$\int_{\Omega_\varepsilon} \nabla_p \tilde{f} = \int_{\gamma_\varepsilon \cup \gamma} \tilde{f} n_p dr = (f(\gamma_\varepsilon) - f(\gamma))(b - a),$$

where dr is the scalar line element and n_p the p -component of the normal n . Applying Hölder's inequality we find

$$|b - a| |f(\gamma_\varepsilon) - f(\gamma)| = \left| \int_{\Omega_\varepsilon} \nabla_p \rho \right| \leq \left(\int_{\Omega_\varepsilon} \frac{1}{\rho} |\nabla_p \rho|^2 \right)^{\frac{1}{2}} \left(\int_{\Omega_\varepsilon} \rho \right)^{\frac{1}{2}} \xrightarrow{\varepsilon \rightarrow 0} 0.$$

This argument shows that f is continuous from the right at the point $\gamma \in \Gamma$.

The following lemma generalizes this argument to the case at hand, in which ρ also depends on time. Note that $\text{Int } \Gamma$ is the interior of the graph Γ , which is Γ without the lower exterior vertices.

Lemma 3.3 (Continuity of ρ). *Let $\rho \in \mathcal{P}([0, T] \times \mathbb{R}^2)$, $\rho(dt dx) = f(t, \xi(x)) dt dx$ for a Borel measurable $f : [0, T] \times \Gamma \rightarrow \mathbb{R}$, and assume that*

$$\int_0^T \mathcal{I}(\rho_t | \mathcal{L}) dt + \sup_{t \in [0, T]} \mathcal{F}(\rho_t) < \infty.$$

Then for almost all $t \in [0, T]$, $\gamma \mapsto f(t, \gamma)$ is continuous on $\text{Int } \Gamma$.

Proof. The argument is essentially the same as the one above. For almost all t , ρ_t is Lebesgue-absolutely-continuous and $\mathcal{I}(\rho_t | \mathcal{L})$ is finite, and the argument above can be applied to the neighbourhood of any point x with $\nabla H(x) \neq 0$, and to both right and left limits. The only elements of Γ that have no representative $x \in \mathbb{R}^2$ with $\nabla H(x) \neq 0$ are the lower ends of the graph, corresponding to the bottoms of the wells of H . At all other points of Γ we obtain continuity. \square

Corollary 3.4 (Continuity of $\hat{\rho}$). *Let ρ be the limit given by Theorem 3.1, and $\hat{\rho} := \xi_{\#} \rho$ its push-forward. For almost all t , $\hat{\rho}_t \ll T(\gamma) d\gamma$, and $d\hat{\rho}_t / T(\gamma) d\gamma$ is continuous on $\text{Int } \Gamma$.*

This corollary follows by combining Lemma 3.3 with Lemma 3.2.

3.6 Liminf inequality

We now derive the final ingredient of the proof, the liminf inequality. Define

$$\hat{I}(\hat{\rho}) := \begin{cases} \sup_{g \in C_c^{1,2}(\mathbb{R} \times \Gamma)} \hat{\mathcal{J}}(\hat{\rho}, g) & \text{if } \hat{\rho}_t \ll T(\gamma) d\gamma, \hat{\rho}_t(d\gamma) = f_t(\gamma) T(\gamma) d\gamma \text{ with } f \text{ continuous on } \text{Int } \Gamma, \\ +\infty & \text{for almost all } t \in [0, T], \\ +\infty & \text{otherwise,} \end{cases} \quad (52)$$

where

$$\begin{aligned} \hat{\mathcal{J}}(\hat{\rho}, g) := & \int_{\Gamma} g_T d\hat{\rho}_T - \int_{\Gamma} g_0 d\hat{\rho}_0 - \int_0^T \int_{\Gamma} (\partial_t g_t(\gamma) + A(\gamma) g_t''(\gamma) + B(\gamma) g_t'(\gamma)) \hat{\rho}_t(d\gamma) dt \\ & - \frac{1}{2} \int_0^T \int_{\Gamma} A(\gamma) (g_t'(\gamma))^2 \hat{\rho}_t(d\gamma) dt, \end{aligned} \quad (53)$$

and we use g' and g'' to indicate derivatives with respect to h . For $\gamma \in \Gamma$, the coefficients are defined by

$$A(\gamma) := \frac{1}{T(\gamma)} \int_{\xi^{-1}(\gamma)} \frac{(\nabla_p H)^2}{|\nabla H|} d\mathcal{H}^1, \quad B(\gamma) := \frac{1}{T(\gamma)} \int_{\xi^{-1}(\gamma)} \frac{\Delta_p H}{|\nabla H|} d\mathcal{H}^1, \quad T(\gamma) := \int_{\xi^{-1}(\gamma)} \frac{1}{|\nabla H|} d\mathcal{H}^1. \quad (54)$$

Note that for our particular choice of $H(q, p) = p^2/2m + V(q)$, we have $B(\gamma) = 1/m$.

The class of test functions in (52) is $C_c^{1,2}(\mathbb{R} \times \Gamma)$; recall that differentiability of a function $f : \Gamma \rightarrow \mathbb{R}$ is defined by restriction to one-dimensional subgraphs, and $C_c^{1,2}(\mathbb{R} \times \Gamma)$ therefore consists of functions $g : \Gamma \rightarrow \mathbb{R}$ that are twice continuously differentiable in h in this sense. The subscript c indicates that we restrict to functions that vanish for sufficiently large h (i.e. somewhere along the top edge of Γ).

Note that again $\hat{I} \geq 0$; formally, $\hat{I}(\hat{\rho}) = 0$ iff $\hat{\rho}$ satisfies the diffusion equation

$$\partial_t \hat{\rho} = (A\hat{\rho})'' - (B\hat{\rho})',$$

and we will investigate this equation in more detail in the next section.

Theorem 3.5 (Liminf inequality). *Under the same conditions as Theorem 3.1, let us assume that $\rho^\varepsilon \rightarrow \rho$ in $\mathcal{M}([0, T]; \mathbb{R}^2)$ and $\hat{\rho}^\varepsilon := \xi_{\#}\rho^\varepsilon \rightarrow \xi_{\#}\rho =: \hat{\rho}$ in $C([0, T]; \mathcal{P}(\Gamma))$. Then*

$$\liminf_{\varepsilon \rightarrow 0} I^\varepsilon(\rho^\varepsilon) \geq I(\hat{\rho}).$$

Proof. Recall the rate functional from (46)

$$I^\varepsilon(\rho^\varepsilon) = \sup_{f \in C_c^{1,2}(\mathbb{R} \times \mathbb{R}^2)} \mathcal{J}^\varepsilon(\rho^\varepsilon, f), \quad \text{where} \quad (55)$$

$$\mathcal{J}^\varepsilon(\rho^\varepsilon, f) := \int_{\mathbb{R}^2} f_T d\rho_T^\varepsilon - \int_{\mathbb{R}^2} f_0 d\rho_0^\varepsilon - \int_0^T \int_{\mathbb{R}^2} \left(\partial_t f + \frac{1}{\varepsilon} J \nabla H \cdot \nabla f + \Delta_p f \right) d\rho_t^\varepsilon - \frac{1}{2} \int_0^T \int_{\mathbb{R}^2} |\nabla_p f|^2 d\rho_t^\varepsilon.$$

Define $\hat{A} := \{f = g \circ \xi : g \in C_c^{1,2}(\mathbb{R} \times \Gamma)\}$. Then we have

$$I^\varepsilon(\rho^\varepsilon) \geq \sup_{f \in \hat{A}} \mathcal{J}^\varepsilon(\rho^\varepsilon, f).$$

Since $J \nabla H \nabla(g \circ \xi) = 0$, upon substituting $f = g \circ \xi$ into \mathcal{J}^ε the $O(1/\varepsilon)$ term vanishes. Using the notation g' for the partial derivative with respect to h , $\partial_t g$ for the time derivative, and suppressing the dependence of g on time, we find

$$\begin{aligned} \mathcal{J}^\varepsilon(\rho^\varepsilon, g \circ \xi) &:= \int_{\Gamma} g_T d\hat{\rho}_T^\varepsilon - \int_{\Gamma} g_0 d\hat{\rho}_0^\varepsilon - \int_0^T \int_{\mathbb{R}^2} \left(\partial_t g(\xi(x)) + g''(\xi(x)) (\nabla_p H(x))^2 + g'(\xi(x)) \Delta_p H(x) \right) \rho_t^\varepsilon(dx) dt \\ &\quad - \frac{1}{2} \int_0^T \int_{\mathbb{R}^2} |g'(\xi(x)) \nabla_p H(x)|^2 \rho_t^\varepsilon(dx) dt. \end{aligned} \quad (56)$$

The limit of (56) is determined term by term. Taking the fourth term as an example, using the co-area formula and the local-equilibrium result of Lemma 3.2, the fourth term on the right-hand side of (56) gives

$$\begin{aligned} &\int_0^T \int_{\mathbb{R}^2} g''(\xi(x)) (\nabla_p H(x))^2 \rho_t^\varepsilon(dx) dt \xrightarrow{\varepsilon \rightarrow 0} \int_0^T \int_{\mathbb{R}^2} g''(\xi(x)) (\nabla_p H(x))^2 \rho_t(dx) dt \\ &= \int_0^T dt \int_{\Gamma} \frac{g''(\gamma) \hat{\rho}_t(d\gamma)}{T(\gamma)} \left(\int_{\xi^{-1}(\gamma)} \frac{(\nabla_p H(y))^2}{|\nabla H(y)|} \mathcal{H}^1(dy) \right) = \int_0^T \int_{\Gamma} A(\gamma) g''(\gamma) \hat{\rho}_t(d\gamma) dt, \end{aligned}$$

where $A : \Gamma \rightarrow \mathbb{R}$ is defined in (54). Proceeding similarly with the other terms we find

$$\liminf_{\varepsilon \rightarrow 0} I^\varepsilon(\rho^\varepsilon) \geq \sup_{g \in C_c^{1,2}(\mathbb{R} \times \Gamma)} \hat{\mathcal{J}}(\hat{\rho}, g). \quad (57)$$

This concludes the proof of Theorem 3.5. □

3.7 Study of the limit problem

We now investigate the limiting functional \hat{I} from (52) a little further. The two main results of this section are that $\hat{\mathcal{J}}$ can be written as

$$\hat{\mathcal{J}}(\hat{\rho}, g) = \int_{\Gamma} g_T d\hat{\rho}_T - \int_{\Gamma} g_0 d\hat{\rho}_0 - \int_0^T \int_{\Gamma} \left[\partial_t g_t d\hat{\rho}_t + \left((TA g_t') + \frac{1}{2} TA g_t'^2 \right) \frac{d\hat{\rho}_t}{T} \right] dt, \quad (58)$$

and that \hat{I} satisfies

$$\hat{I}(\hat{\rho}) \geq \sup_{g \in \mathcal{A}} \hat{\mathcal{J}}(\hat{\rho}, g) \quad \text{for all } \hat{\rho} \in C([0, T]; \mathcal{P}(\Gamma)), \quad (59)$$

where \mathcal{A} is the larger class

$$\mathcal{A} := \left\{ g : C^{1,0}(\mathbb{R} \times \Gamma) : g|_{I_k} \in C_b^{1,2}(\mathbb{R} \times I_k), \quad \forall \text{ interior vertex } O_j \forall t : \sum_{k: I_k \sim O_j} \pm_{kj} g_t'(O_j, k) TA(O_j, k) = 0 \right\}. \quad (60)$$

The admissible set \mathcal{A} relaxes the conditions on g at interior vertices: instead of requiring g to have identical derivatives coming from each edge, only a single scalar combination of the derivatives has to vanish. (In fact it can be shown that equality holds in (59), but that requires a further study of the limiting equation that takes us too far here.)

Both results use some special properties of T , A , and B , which are given by the following lemma. In this lemma and below we use TA and TB for the functions obtained by multiplying T with A and B ; these combinations play a special role, and we treat them as separate functions.

Lemma 3.6 (Properties of TA and TB). *The functions TA and TB have the following properties.*

1. $TA \in C^1(I_k)$ for each k , and $(TA)' = TB$;
2. TA is bounded on compact subsets of Γ ;
3. At each interior vertex O_j , for each k such that $I_k \sim O_j$, $TA(O_j, k) := \lim_{\substack{h \in I_k \\ h \rightarrow O_j}} TA(h, k)$ exists, and

$$\sum_{k: I_k \sim O_j} \pm_{kj} TA(O_j, k) = 0. \quad (61)$$

From this lemma the expression (58) follows by simple manipulation.

With these two results, we can obtain a differential-equation characterization of those $\hat{\rho}$ with $\hat{I}(\hat{\rho}) = 0$. Assume that a $\hat{\rho}$ with $\hat{I}(\hat{\rho}) = 0$ is given. By rescaling we find that for all $g \in \mathcal{A}$,

$$\int_{\Gamma} g_T d\hat{\rho}_T - \int_{\Gamma} g_0 d\hat{\rho}_0 = \int_0^T \int_{\Gamma} \left[\partial_t g_t d\hat{\rho} + (TA g_t')' \frac{d\hat{\rho}_t}{T} \right] dt. \quad (62)$$

As already remarked we find a parabolic equation inside each edge of Γ ,

$$\partial_t \hat{\rho}_t = \left(TA \left(\frac{\hat{\rho}_t}{T} \right)' \right)' = (A \hat{\rho}_t)'' - (B \hat{\rho}_t)'. \quad (63)$$

We next determine the boundary and connection conditions at the vertices.

Consider a single interior vertex O_j , and choose a function $g \in \mathcal{A}$ such that $\text{supp } g$ contains no other vertices. Writing $\hat{\rho}_t(d\gamma) = f_t(\gamma)T(\gamma)d\gamma$ we find first that f_t is continuous at O_j , by the definition (52) of \hat{I} .

Then, assuming that $\hat{\rho}$ is smooth enough for the following expressions to make sense¹, we perform two partial integrations in γ and one in time on (62) and substitute (63) to find

$$0 = \int_0^T f_t(O_j) \sum_{k:I_k \sim O_j} \pm_{kj} TA(O_j, k) g'_t(O_j, k) dt - \int_0^T g_t(O_j) \sum_{k:I_k \sim O_j} \pm_{kj} TA(O_j, k) f'_t(O_j, k) dt.$$

The first term vanishes since $g \in \mathcal{A}$, while the second term leads to the connection condition

$$\text{at each interior vertex } O_j : \sum_{k:I_k \sim O_j} \pm_{kj} TA(O_j, k) f'_t(O_j, k) = 0.$$

The lower exterior vertices and the top vertex are *inaccessible*, in the language of [Fel52, Man68], and therefore require no boundary condition. Summarizing, we find that if $\hat{I}(\hat{\rho}) = 0$, then $\hat{\rho} =: fT d\gamma$ satisfies a weak version of equation (63) with connection conditions

$$\text{at each interior vertex } O_j : f \text{ is continuous and } \sum_{k:I_k \sim O_j} \pm_{kj} TA(O_j, k) f'_t(O_j, k) = 0.$$

This combination of equation and boundary conditions can be proved to characterize a well-defined semigroup using e.g. the Hille-Yosida theorem and the characterization of one-dimensional diffusion processes by Feller (e.g. [Fel52]).

We now prove the inequality (59).

Lemma 3.7 (Comparison of \hat{I} and \tilde{I}). *We have*

$$\hat{I}(\hat{\rho}) \geq \tilde{I}(\hat{\rho}) := \sup_{g \in \mathcal{A}} \hat{\mathcal{J}}(\hat{\rho}, g).$$

Proof. Take $\hat{\rho}$ such that $\hat{I}(\hat{\rho}) < \infty$, implying that $\hat{\rho}_t(d\gamma) = f_t(\gamma)T(\gamma)d\gamma$ with f_t continuous on $\text{Int } \Gamma$ for almost all t . Choose $g \in \mathcal{A}$; we will show that $\hat{I}(\hat{\rho}) \geq \hat{\mathcal{J}}(\hat{\rho}, g)$, thus proving the lemma. For simplicity we only treat the case of a single interior vertex, called O ; the case of multiple vertices is a simple generalization. For convenience we also assume that O corresponds to $h = 0$.

Define

$$g_{\delta,t}(h, k) = g_t(h, k)\zeta_\delta(h) + (1 - \zeta_\delta(h))g_t(0), \quad (64)$$

where ζ_δ is a sequence of smooth functions such that

- ζ_δ is identically zero in a δ -neighbourhood of O , and identically 1 away from a 2δ -neighbourhood of O ;
- ζ_δ satisfies the growth conditions $|\zeta'_\delta| \leq 2/\delta$ and $|\zeta''_\delta| \leq 4/\delta^2$.

We calculate $\hat{\mathcal{J}}(\hat{\rho}, g_\delta)$. The limit of the first three terms is straightforward: by dominated convergence we obtain

$$\int_\Gamma g_{\delta,T} d\hat{\rho}_T - \int_\Gamma g_{\delta,0} d\hat{\rho}_0 - \int_0^T \int_\Gamma \partial_t g_{\delta,t} d\hat{\rho}_t \xrightarrow{\delta \rightarrow 0} \int_\Gamma g_T d\hat{\rho}_T - \int_\Gamma g_0 d\hat{\rho}_0 - \int_0^T \int_\Gamma \partial_t g_t d\hat{\rho}_t.$$

Next consider the term

$$\int_0^T \int_\Gamma A(\gamma) g''_\delta(\gamma) \hat{\rho}_t(d\gamma) dt = \int_0^T \int_\Gamma \left[g''(h, k)\zeta_\delta(h) + 2\zeta'_\delta(h)g'(h, k) + \zeta''_\delta(h)[hg'(0, k) + O(h^2)] \right] A(\gamma) \hat{\rho}_t(d\gamma) dt. \quad (65)$$

¹This can actually be proved using the properties of A and B near the vertices and applying standard parabolic regularity theory on each of the edges.

Since the function $(\gamma, t) \mapsto A(\gamma)g_t''(\gamma) \in L^\infty(\hat{\rho}_t)$ the first term in (65) again converges by dominated convergence :

$$\int_0^T \int_\Gamma g_t''(h, k) \zeta_\delta(h) A(h, k) \hat{\rho}_t(d\gamma) dt \xrightarrow{\delta \rightarrow 0} \int_0^T \int_\Gamma g_t''(h, k) A(h, k) \hat{\rho}_t(d\gamma) dt.$$

Abbreviate $f_t(\gamma)TA(\gamma)$ as $a(\gamma)$; note that a is continuous and bounded in a neighbourhood of O . Write the second term on the right-hand side in (65) as (supressing the time integral for the moment)

$$\begin{aligned} 2 \int_\Gamma \zeta_\delta'(h) g'(h, k) a(h, k) dh &= 2 \int_\Gamma \zeta_\delta'(h) g'(h, k) (a(h, k) - a(0, k)) d\gamma + 2 \sum_k a(0, k) \int_{I_k} \zeta_\delta'(h) (g'(h, k) - g'(0, k)) dh \\ &\quad + 2 \sum_k a(0, k) g'(0, k) \int_{\Gamma_k} \zeta_\delta'(h) dh \\ &\xrightarrow{\delta \rightarrow 0} 0 + 0 - 2 \sum_{k: I_k \sim O} \pm_{kO} g'(0, k) a(0, k) = 2 \sum_{k: I_k \sim O} \pm_{kO} g'(0, k) f(0, k) TA(0, k). \end{aligned}$$

The limit above holds since $-\zeta_\delta'(\cdot, k)$ converges weakly to a signed Dirac, $\pm_{kO}\delta_0$, as $\delta \rightarrow 0$. Proceeding similarly with the remaining terms we have

$$\begin{aligned} \hat{I}(\hat{\rho}) \geq \hat{\mathcal{J}}(\hat{\rho}, g_\delta) &\xrightarrow{\delta \rightarrow 0} \int_\Gamma g_T d\hat{\rho}_T - \int_\Gamma g_0 d\hat{\rho}_0 - \int_0^T \int_\Gamma (\partial_t g_t + A(\gamma)g_t''(\gamma) + B(\gamma)g_t'(\gamma)) \hat{\rho}_t(d\gamma) dt \\ &\quad - \frac{1}{2} \int_0^T \int_\Gamma A(\gamma)g_t'(\gamma)^2 \hat{\rho}_t(d\gamma) dt - \int_0^T f_t(0, k) \left[\sum_{k: I_k \sim O} \pm_{kO} TA(0, k) g_t'(0, k) \right] dt. \end{aligned}$$

Note that the final term vanishes by the requirement that $g \in \mathcal{A}$, and therefore the right-hand side above equals $\hat{\mathcal{J}}(\hat{\rho}, g)$. This concludes the proof of the lemma. \square

We still owe the reader the proof of Lemma 3.6.

Proof of Lemma 3.6. We first prove part 1. For simplicity, assume first that H has a single well, and therefore Γ has only one edge, $k = 1$. Since

$$\operatorname{div} \begin{pmatrix} 0 \\ \nabla_p H \end{pmatrix} = \Delta_p H,$$

and remarking that the exterior normal n to the set $H \leq h$ equals $(0, \nabla_p H / |\nabla H|)^T$, we calculate that

$$\int_{\{H \leq h\}} \Delta_p H = \int_{\{H=h\}} \frac{(\nabla_p H)^2}{|\nabla H|} d\mathcal{H}^1 = TA(h). \quad (66)$$

By the smoothness of H , the derivative of the left-hand integral is well-defined for all h such that $\nabla H \neq 0$ at that level. At such h we then have

$$TB(h) = \int_{\{H=h\}} \frac{\Delta_p H}{|\nabla H|} d\mathcal{H}^1 = \partial_h \int_{\{H \leq h\}} \Delta_p H = \partial_h TA(h).$$

For the multi-well case, this argument can simply be applied to each branch of Γ .

For part 2, since H is coercive, $\{H \leq h\}$ is bounded for each h ; since H is smooth, therefore $\Delta_p H$ is bounded on bounded sets. From (66) it follows that TA also is bounded on bounded sets of Γ .

Finally, for part 3, note first that TB is bounded near each interior vertex. This follows by an explicit calculation and our assumption that each interior vertex corresponds to exactly one, non-degenerate, saddle point. Since $(TA)' = TB$, TA has a well-defined and finite limit at each interior saddle. The summation property (61) follows from comparing (66) for values of h just above and below the critical value. For

instance, in the case of a single saddle at value $h = 0$, with two lower edges $k = 1, 2$ and upper edge $k = 0$, we have

$$\begin{aligned} \lim_{h \uparrow 0} TA(h, 1) + TA(h, 2) &= \lim_{h \uparrow 0} \int_{\xi^{-1}((-\infty, h] \times \{1\})} \Delta_p H + \int_{\xi^{-1}((-\infty, h] \times \{2\})} \Delta_p H \\ &= \lim_{h \uparrow 0} \int_{\{H \leq h\}} \Delta_p H = \lim_{h \downarrow 0} \int_{\{H \leq h\}} \Delta_p H = \lim_{h \downarrow 0} TA(h, 0). \end{aligned}$$

This concludes the proof of Lemma 3.6. □

3.8 Conclusion and Discussion

The combination of Theorems 3.1 and 3.5 give us that along subsequences $\hat{\rho}^\varepsilon := \xi_{\#} \rho^\varepsilon$ converges in an appropriate manner to some $\hat{\rho}$, and that

$$\hat{I}(\hat{\rho}) \leq \liminf_{\varepsilon \rightarrow 0} I^\varepsilon(\rho^\varepsilon).$$

In addition, any $\hat{\rho}$ satisfying $I(\hat{\rho}) = 0$ is a weak solution of the PDE

$$\partial_t \hat{\rho} = (A\hat{\rho})'' - (B\hat{\rho})'$$

on the graph Γ . This is the central coarse-graining statement of this section. We also obtain the boundary conditions, similarly as in the conventional weak-formulation method, by expanding the admissible set of test functions.

In switching from the VFP equation (9) to equation (40) we removed two terms, representing the friction with the environment and the interaction between particles. Mathematically, it is straightforward to treat the case with friction, which leads to an additional drift term in the limit equation in the direction of decreasing h . We left this out simply for the convenience of shorter expressions.

As for the interaction, represented by the interaction potential ψ , again there is no mathematical necessity for setting $\psi = 0$ in this section; the analysis continues rather similarly. However, the limiting equation will now be non-local, since the particles at some $\gamma \in \Gamma$, which can be thought of as ‘living’ on a full connected level set of H , will feel a force exerted by particles at a different $\gamma' \in \Gamma$, i.e. at a different level set component. This makes the interpretation of the limiting equation somewhat convoluted.

The results of the current and the next sections were proved by Freidlin and co-authors in a series of papers [FW93, FW94, FW98, FW01, FW04], using probabilistic techniques. Recently, Barret and Von Renesse [BvR14] provided an alternative proof using Dirichlet forms and their convergence. The latter approach is closer to ours in the sense that it is mainly PDE-based method and of variational type. However, in [BvR14] the authors consider a perturbation of the Hamiltonian by a friction term and a non-degenerate noise, i.e. the noise is present in both space and momentum variables; this non-degeneracy appears to be essential in their method. Moreover, their approach invokes a reference measure which is required to satisfy certain non-trivial conditions. In contrast, the approach of this paper is applicable to degenerate noise and does not require such a reference measure. In addition, certain non-linear evolutions can be treated, such as the example of the VFP equation.

4 Diffusion on a Graph, $d > 1$

We now switch to our final example. As described in the introduction, the higher-dimensional analogue of the diffusion-on-graph system has an additional twist: in order to obtain unique stationary measures on level

sets of ξ , we need to add an additional noise in the SDE, or equivalently, an additional diffusion term in the PDE. This leads to the equation

$$\partial_t \rho = -\frac{1}{\varepsilon} \operatorname{div}(\rho J \nabla H) + \frac{\kappa}{\varepsilon} \operatorname{div}(a \nabla \rho) + \Delta_\rho \rho, \quad (67)$$

where $a : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d \times 2d}$ with $a \nabla H = 0$, $\dim(\operatorname{Ker}(a)) = 1$ and $\kappa, \varepsilon > 0$ with $\kappa \gg \varepsilon$. The spatial domain is \mathbb{R}^{2d} , $d > 1$, with coordinates $(q, p) \in \mathbb{R}^d \times \mathbb{R}^d$. Here the unknown is trajectory in the space of probability measures $\rho : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^{2d})$; the Hamiltonian is the same as in the previous section, $H : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ given by $H(q, p) = p^2/2m + V(q)$.

The results for the limit $\varepsilon \rightarrow 0$ in (67) closely mirror the one-degree-of-freedom diffusion-on-graph problem of the previous section; the only real difference lies in the proof of local equilibrium (Lemma 3.2). For a rigorous proof of this lemma in this case, based on probabilistic techniques, we refer to [FW01, Lemma 3.2]; here we only outline a possible analytic proof.

Along the lines of Theorem 3.1, and using boundedness of the rate functional $I^\varepsilon(\rho^\varepsilon)$, one can show that

$$\frac{1}{2} \int_0^T \int_{\mathbb{R}^2} \frac{|\nabla_p \rho^\varepsilon|^2}{\rho^\varepsilon} + \frac{\kappa}{\varepsilon} \int_0^T \int_{\mathbb{R}^2} \frac{a \nabla \rho^\varepsilon \cdot \nabla \rho^\varepsilon}{\rho^\varepsilon} \leq C.$$

Multiplying this inequality by ε/κ and using the weak convergence $\rho^\varepsilon \rightharpoonup \rho$ along with the lower-semicontinuity of the Fisher information [FK06, Theorem D.45] we find

$$\int_0^T \int_{\mathbb{R}^2} \frac{a \nabla \rho \cdot \nabla \rho}{\rho} = 0,$$

or in variational form, for almost all $t \in [0, T]$,

$$\begin{aligned} 0 &= \sup_{\varphi \in C_c^\infty(\mathbb{R}^{2d})} \int_{\mathbb{R}^{2d}} \operatorname{div}(a \nabla \varphi) \rho_t - \frac{1}{2} \int_{\mathbb{R}^{2d}} a \nabla \varphi \cdot \nabla \varphi \rho_t \\ &\iff 0 = \int_{\mathbb{R}^{2d}} \operatorname{div}(a \nabla \varphi) \rho_t, \quad \forall \varphi \in C_c^\infty(\mathbb{R}^{2d}). \end{aligned}$$

Applying the co-area formula we find

$$\int_{\xi^{-1}(\gamma)} \frac{\rho(x)}{|\nabla H(x)|} \operatorname{div}(a(x) \nabla \varphi(x)) \mathcal{H}^{2d-1}(dx) = 0, \quad (68)$$

where \mathcal{H}^{2d-1} is the $(2d-1)$ dimensional Hausdorff measure. Let \mathcal{M}_γ be the $(2d-1)$ dimensional manifold $\xi^{-1}(\gamma)$ with volume element $|\nabla H|^{-1} \mathcal{H}^{2d-1}$. Then (68) becomes

$$\int_{\mathcal{M}_\gamma} \rho(x) \operatorname{div}_{\mathcal{M}}(a(x) \nabla_{\mathcal{M}} \varphi(x)) \operatorname{vol}_{\mathcal{M}}(dx) = 0,$$

where $\operatorname{div}_{\mathcal{M}}$ and $\nabla_{\mathcal{M}}$ are the corresponding differential operators on \mathcal{M}_γ , and $\operatorname{vol}_{\mathcal{M}}$ is the induced volume measure. Since $a \nabla H = 0$, $\dim(\operatorname{Ker}(a)) = 1$, a is non-degenerate on the tangent space of \mathcal{M}_γ . Therefore, given $\psi \in C^\infty(\mathcal{M}_\gamma)$ with $\int_{\mathcal{M}_\gamma} \psi d \operatorname{vol}_{\mathcal{M}} = 0$, we can solve the corresponding Laplace-Beltrami-Poisson equation for φ ,

$$\operatorname{div}_{\mathcal{M}}(a \nabla_{\mathcal{M}} \varphi) = \psi,$$

and therefore

$$\int_{\mathcal{M}_\gamma} \rho \psi d \operatorname{vol}_{\mathcal{M}} = 0, \quad \forall \psi \in C^\infty(\mathcal{M}_\gamma) \text{ with } \int_{\mathcal{M}_\gamma} \psi d \operatorname{vol}_{\mathcal{M}} = 0.$$

Since \mathcal{M}_γ is connected by definition, it follows that ρ constant on \mathcal{M}_γ ; this is the statement of Lemma 3.2.

5 Conclusion and discussion

In this paper we have presented a structure in which coarse-graining and ‘passing to a limit’ combine in a natural way, and which extends also naturally to a class of approximate solutions. The central object is the rate function I , which is minimal and vanishes at solutions; in the dual formulation of this rate function, coarse-graining has a natural interpretation, and the inequalities of the dual formulation and of the coarse-graining combine in a convenient way.

We now comment on a number of issues related with this method.

Why does this method work? One can wonder why the different pieces of the arguments of this paper fit together. Why do the relative entropy and the relative Fisher information appear? To some extent this can be recognized in the similarity between the duality definition of the rate function I and the duality characterization of relative entropy and relative Fisher Information. The details of Appendix B show this most clearly, but the similarity between the duality definition of the relative Fisher information and the duality structure of I can readily be recognized: in (19) combined with (18) we collect the $O(\gamma^2)$ terms

$$\int_0^T \int_{\mathbb{R}^{2d}} \left[\Delta_p f_t - \frac{p}{m} \nabla_p f_t - \frac{1}{2} |\nabla_p f_t|^2 \right] d\rho_t dt,$$

and these match one-to-one to the definition (24). This shows how the structure of the relative Fisher Information is to some extent ‘built-in’ in this system.

Relation with other variational formulations. Our variational formulation (2) to ‘passing to a limit’ is closely related to other variational formulations in the literature, notably the Ψ - Ψ^* formulation and the method in [PSCF05, ASZ09]. In the Ψ - Ψ^* formulation, a gradient flow of the energy $\mathcal{E}_\varepsilon : \mathcal{Z} \rightarrow \mathbb{R}$ with respect to the dissipation Ψ_ε^* is defined to be a curve $\rho^\varepsilon \in C([0, T], \mathcal{Z})$ such that

$$\mathcal{A}^\varepsilon(\rho) := \mathcal{E}_\varepsilon(\rho_T) - \mathcal{E}_\varepsilon(\rho_0) + \int_0^T [\Psi_\varepsilon(\dot{\rho}_t, \rho_t) + \Psi_\varepsilon^*(-D\mathcal{E}_\varepsilon(\rho_t), \rho_t)] dt = 0. \quad (69)$$

‘Passing to a limit’ in a Ψ - Ψ^* structure is then accomplished by studying (Gamma-) limits of the functionals \mathcal{A}^ε . The method introduced in [PSCF05, ASZ09] is slightly different. Therein ‘passing to a limit’ in the evolution equation is executed by studying (Gamma-)limits of the functionals that appear in the approximating discrete minimizing-movement schemes.

The similarities between these two approaches and ours is that all the methods hinge on duality structure of the relevant functionals, allow one to obtain both compactness and limiting results, and can work with approximate solutions, see e.g. [AMP⁺12] and the papers above for details. In addition, all methods assume some sort of well-prepared initial data, such as bounded initial free energy and boundedness of the functionals. Our assumptions on the boundedness of the rate functionals arise naturally in the context of large-deviation principle since this assumption describes events of a certain degree of ‘improbability’.

The main difference is that the method of this paper makes no use of the gradient-flow structure, and therefore also applies to non-gradient-flow systems as in this paper. The first example, of the overdamped limit of the VFP equation, also is interesting in the sense that it derives a dissipative system from a non-dissipative one. Since the GENERIC framework unifies both dissipative and non-dissipative systems, we expect that the method of this paper could be used to derive evolutionary convergence for GENERIC systems (see the next point). Finally, we emphasize that using the duality of the rate functional is mathematically convenient because we do not need to treat the three terms in the right-hand side of (69) separately. Note that although the entropy and energy functionals as well as the dissipation mechanism are not explicitly present in this formulation, we are still able to derive an energy-dissipation inequality in (4).

Relation with GENERIC. As mentioned in the introduction, the Vlasov-Fokker-Planck system (8) combines both conservative and dissipative effects. In fact it can be cast into the GENERIC form by introducing an excess-energy variable e , depending only on time, that captures the fluctuation of energy due to dissipative effects (but does not change the evolution of the system). The building blocks of the GENERIC for the

augmented system for (ρ, e) can be easily deduced from the conservative and dissipative effects of the original Vlasov-Fokker-Planck equation. Moreover, this GENERIC structure can be derived from the large-deviation rate functional of the empirical process (7). We refer to [DPZ13] for more information. This suggests that our method could be applied to other GENERIC systems.

Gradient flows and large-deviation principles. As mentioned in the introduction, this approach using the duality formulation of the rate functionals is motivated by our recent results on the connection between generalised gradient flows and large-deviation principles [ADPZ11, ADPZ13, DPZ14, DPZ13, DLZ12, MPR14]. We want to discuss here how the two overlap but are not the same. In [MPR14], the authors show that if \mathcal{N}^ε is the adjoint operator of a generator of a Markov process that satisfies a *detailed balance condition*, then the evolution (1) is the same as the generalised gradient flow induced from a large-deviation rate functional, which is of the form $\int_0^T \mathcal{L}^\varepsilon(\rho_t, \dot{\rho}_t) dt$, of the underlying empirical process. The generalised gradient flow is described via the Ψ - Ψ^* structure as in (69) with $\mathcal{L}^\varepsilon(z, \dot{z}) = \Psi_\varepsilon(z, \dot{z}) + \Psi_\varepsilon^*(z, -D\mathcal{E}_\varepsilon(z)) + \langle D\mathcal{E}_\varepsilon(z), \dot{z} \rangle$. Moreover, \mathcal{E}_ε and Ψ_ε can be determined from \mathcal{L}^ε [MPR14, Theorem 3.3]. However, it is not clear if such characterisation holds true for systems that do not satisfy detailed balance. In addition, there exist (generalised) gradient flows for which we currently do not know of any corresponding microscopic particle systems, such as the Allen-Cahn and Cahn-Hilliard equations.

Quantification of coarse-graining error. The use of the rate functional in a central role allows us not only to derive the limiting coarse-grained system but also to obtain quantitative estimates of the coarse-graining error. Existing quantitative methods such as [LL10] and [GOVW09] only work for gradient flows systems since they use crucially the gradient flow structures. The essential estimate that they need is the energy-dissipation inequality, which is similar to (4). Since we are able to obtain this inequality from the duality formulation of the rate functionals, our method would offer an alternative technique for obtaining quantitative estimate of the coarse-graining error for both dissipative and non-dissipative systems. We address this issue in detail in a companion article [DLP⁺15].

Other stochastic processes. The key ingredient of the method is the duality structure of the rate functional (5) and (10). This duality formulation holds true for many other stochastic processes; indeed, the ‘Feng-Kurtz’ algorithm (see chapter 1 of [FK06]) suggests that the large-deviation rate functional for a very wide class of Markov processes can be written as

$$I(\rho) = \sup_f \left\{ \langle f_T, \rho_T \rangle - \langle f_0, \rho_0 \rangle - \int_0^T \langle \dot{f}_t, \rho_t \rangle dt - \int_0^T \mathcal{H}(\rho_t, f_t) dt \right\},$$

where \mathcal{H} is an appropriate limit of ‘non-linear’ generators. The formula (10) is a special case. As a result, we expect that the method can be extended to this same wide class of Markov processes.

A Proof of Lemma 2.1

Define $\tilde{\mathcal{I}}(f)$ to be the right-hand side in (25),

$$\tilde{\mathcal{I}}(f) := \begin{cases} \int_{\mathbb{R}^{2d}} \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right|^2 f dqdp, & \text{if } \nabla_p f \in L^1_{\text{loc}}(dqdp), \\ \infty & \text{otherwise.} \end{cases}$$

for $f \in L^1(\mathbb{R}^{2d})$. We need to show that $\tilde{\mathcal{I}}(f) = \mathcal{I}(f dqdp | \mu)$.

First assume that $\tilde{\mathcal{I}}$ is finite. Then $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \in L^2(f dqdp)$, which implies the following stronger statement.

Lemma A.1. *One has*

$$\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \in L^2_{\nabla}(f dqdp),$$

where the space $L^2_{\nabla}(fdqdp)$ is defined as the closure of $\{\nabla_p \varphi : \varphi \in C_c^\infty(\mathbb{R}^{2d})\}$ with respect to the norm $\|\cdot\|_{fdqdp}^2 := \int_{\mathbb{R}^{2d}} |\cdot|^2 fdqdp$.

Assuming Lemma A.1 for the moment we rewrite $\tilde{\mathcal{I}}(f)$ as

$$\begin{aligned}\tilde{\mathcal{I}}(f) &= \int_{\mathbb{R}^{2d}} \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right|^2 fdqdp = \left\| -\nabla_p \cdot \left(f \left(\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right) \right) \right\|_{-1, (fdqdp)}^2 \\ &= \left\| -\nabla_p \cdot \left(\mathbf{1}_{\{f>0\}} \nabla_p f + f \frac{p}{m} \right) \right\|_{-1, (fdqdp)}^2 \\ &= \left\| -\Delta_p(\mathbf{1}_{\{f>0\}} f) - \nabla_p \cdot \left(f \frac{p}{m} \right) \right\|_{-1, (fdqdp)}^2,\end{aligned}$$

where $\|\cdot\|_{-1, fdqdp}$ is the dual norm (in duality with $L^2_{\nabla}(fdqdp)$) from [DPZ13] and $\mathbf{1}_{\{f>0\}} \nabla_p f = \nabla_p(\mathbf{1}_{\{f>0\}} f)$ holds due to Stampacchia's Lemma [KS00, Theorem A.1]. Following the variational characterization of $\|\cdot\|_{-1, (fdqdp)}$ from [DPZ13, (11)] we finally obtain

$$\begin{aligned}\tilde{\mathcal{I}}(f) &= \sup_{\varphi \in C_c^\infty(\mathbb{R}^{2d})} 2 \int_{\mathbb{R}^{2d}} \left(\nabla_p \varphi \cdot \frac{p}{m} - \mathbf{1}_{\{f>0\}} \Delta_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right) fdqdp \\ &= \sup_{\varphi \in C_c^\infty(\mathbb{R}^{2d})} 2 \int_{\mathbb{R}^{2d}} \left(\nabla_p \varphi \cdot \frac{p}{m} - \Delta_p \varphi - \frac{1}{2} |\nabla_p \varphi|^2 \right) fdqdp,\end{aligned}$$

which is the claimed result. The same reference also provides that $\tilde{\mathcal{I}} = \infty$ iff $\mathcal{I}(fdqdp|\mu) = \infty$.

Proof of Lemma A.1. We assume that $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \in L^2(fdqdp)$ and show that the two individual terms $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}}$ and $\frac{p}{m}$ are in $L^2_{\nabla}(fdqdp)$. Choose a smooth cut-off function $\eta_R = \eta(x/R)$ with $\eta : \mathbb{R}^{2d} \rightarrow \mathbb{R}$, $\eta = 1$ on $B_1(0)$ and $\eta = 0$ in $\mathbb{R}^{2d} \setminus B_2(0)$. Then

$$\begin{aligned}- \int_{\mathbb{R}^{2d}} \eta_R \frac{p}{m} \cdot \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} f &= - \int_{\mathbb{R}^{2d}} \eta_R \frac{p}{m} \cdot \nabla_p f \mathbf{1}_{\{f>0\}} = - \int_{\mathbb{R}^{2d}} \eta_R \frac{p}{m} \cdot \nabla_p(\mathbf{1}_{\{f>0\}} f) \\ &= + \frac{1}{m} \int_{\mathbb{R}^{2d}} \left[\eta_R d + p \cdot \nabla_p \eta_R \right] \mathbf{1}_{\{f>0\}} f \leq \frac{d}{m} + \int_{\mathbb{R}^{2d}} p \cdot \nabla_p \eta_R f =: b(R).\end{aligned}$$

As $R \rightarrow \infty$, the bound $b(R)$ converges to d/m .

Therefore we have

$$\begin{aligned}\int_{\mathbb{R}^{2d}} \eta_R \left[\left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} \right|^2 + \left| \frac{p}{m} \right|^2 \right] f &= \int_{\mathbb{R}^{2d}} \eta_R \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right|^2 f - 2 \int_{\mathbb{R}^{2d}} \eta_R \nabla_p f \cdot \frac{p}{m} \mathbf{1}_{\{f>0\}} \\ &\leq 2b(R) + \int_{\mathbb{R}^{2d}} \eta_R \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right|^2 f.\end{aligned}$$

By passing to the limit $R \rightarrow \infty$ we obtain

$$\lim_{R \rightarrow \infty} \int_{\mathbb{R}^{2d}} \eta_R \left[\left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} \right|^2 + \left| \frac{p}{m} \right|^2 \right] f \leq \int_{\mathbb{R}^{2d}} \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} + \frac{p}{m} \right|^2 f + \frac{2d}{m} < \infty$$

and thus $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}}, \frac{p}{m} \in L^2(fdqdp)$. To conclude the proof of Lemma A.1 it remains to show that $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}}, \frac{p}{m}$ can be approximated by gradients of C_c^∞ -functions. For $\varepsilon > 0$ we therefore define

$$\varphi_\varepsilon := \log(f \vee \varepsilon) - \log \varepsilon.$$

Then φ_ε has compact support in \mathbb{R}^{2d} and $\nabla_p \varphi_\varepsilon = \frac{\nabla_p f}{f}$ for $f \geq \varepsilon$ and 0 otherwise. Note that φ_ε is not necessarily smooth, but by convolution with a mollifier we can also achieve smoothness. Our aim is to show that $\|\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} - \nabla_p \varphi_\varepsilon\|_{fdqdp} \rightarrow 0$ as $\varepsilon \rightarrow 0$. Indeed,

$$\int_{\mathbb{R}^{2d}} \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} - \nabla_p \varphi_\varepsilon \right|^2 f = \int_{\{f < \varepsilon\}} \left| \frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} \right|^2 f \xrightarrow{\varepsilon \rightarrow 0} 0.$$

This proves that $\frac{\nabla_p f}{f} \mathbf{1}_{\{f>0\}} \in L^2_{\nabla}(fdqdp)$. The calculation for $\frac{p}{m} = \nabla_p \left(\frac{|p|^2}{2m} \right)$ is similar. \square

B Proof of Theorem 2.3

We prove Theorem 2.3 using the method of the duality equation; see e.g. [ACP82, RK82, BKP85, Eid90] for examples. Throughout this appendix γ is fixed.

In addition to the duality definition of relative Fisher Information (24) we also use the Donsker-Varadhan duality characterization of the relative entropy (21) for two probability measures (see e.g. [DE97, Lemma. 1.4.3]):

$$\mathcal{H}(\nu|\mu) = \sup_{\phi \in C_c^\infty(\mathbb{R}^{2d})} \int_{\mathbb{R}^{2d}} \phi d\nu - \log \int_{\mathbb{R}^{2d}} e^\phi d\mu,$$

which implies the corresponding characterization of the free energy (22),

$$\mathcal{F}(\nu) = \sup_{\phi \in C_c^\infty(\mathbb{R}^{2d})} \int_{\mathbb{R}^{2d}} \left[\phi + H + \frac{1}{2} \psi * \nu \right] d\nu - \log \int_{\mathbb{R}^{2d}} e^\phi dx + \log Z_H. \quad (70)$$

Lemma B.1. *For all $g \in C_b^{1,2}(\mathbb{R} \times \mathbb{R}^{2d})$, all $\rho \in C([0, T]; \mathcal{P}(\mathbb{R}^{2d}))$, and $\tau \in [0, T]$,*

$$\begin{aligned} I^\gamma(\rho) \geq \int_{\mathbb{R}^{2d}} \rho \left(\frac{1}{2} \psi * \rho + H + g \right) \Big|_0^\tau - \int_0^\tau \int_{\mathbb{R}^{2d}} \left\{ \gamma^2 \left(\frac{d}{m} - \frac{1}{2} \frac{p^2}{m^2} \right) \right. \\ \left. + \partial_t g + \gamma J \nabla (H + \psi * \rho_t) \cdot \nabla g + \gamma^2 \Delta_p g + \frac{\gamma^2}{2} |\nabla_p g|^2 \right\} d\rho_t dt. \end{aligned} \quad (71)$$

Lemma B.2. *Given $\phi \in C_c^\infty(\mathbb{R}^d)$ and $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d)$, there exists a function $g \in L^\infty(0, T; (W^{1,\infty} \cap L^1)(\mathbb{R}^{2d})) \cap L^2([0, T] \times \mathbb{R}_q^d; H^2(\mathbb{R}_p^d))$ with $\partial_t g \in (L^\infty + L^2)([0, T] \times \mathbb{R}^{2d})$, which satisfies the following equation pointwise a.e.:*

$$\partial_t g + \gamma J \nabla (H + \psi * \rho_t) \cdot \nabla g + \gamma^2 \Delta_p g + \frac{\gamma^2}{2} |\nabla_p g|^2 = \gamma^2 \left(\Delta_p \varphi + \frac{1}{2} |\nabla_p \varphi|^2 \right) \quad \text{in } [0, T] \times \mathbb{R}^d \quad (72a)$$

$$g = \phi \quad \text{at } t = T. \quad (72b)$$

This solution satisfies

$$t \mapsto \int_{\mathbb{R}^{2d}} e^{g_t} dx \quad \text{is non-decreasing.} \quad (73)$$

The proof of Theorem 2.3 follows from these two lemmas. Using a smoothed version of the function g given by Lemma B.2 (on the time interval $[0, \tau]$) in Lemma B.1, we find that

$$\sup_{\varphi, \phi} \int_{\mathbb{R}^d} \rho \left(\frac{1}{2} \psi * \rho + H + \phi \right) \Big|_{t=\tau} - \gamma^2 \int_0^\tau \int_{\mathbb{R}^{2d}} \rho \left(\frac{d}{m} - \frac{1}{2} \frac{p^2}{m^2} + \Delta_p \varphi + \frac{1}{2} |\nabla_p \varphi|^2 \right) \leq I^\gamma(\rho) + \int_{\mathbb{R}^{2d}} \rho \left(\frac{1}{2} \psi * \rho + H + g \right) \Big|_{t=0}. \quad (74)$$

Using (70) we estimate the integral on the right-hand side by

$$\int_{\mathbb{R}^{2d}} \rho \left(\frac{1}{2} \psi * \rho + H + g \right) \Big|_{t=0} \leq \mathcal{F}(\rho_0) + \log \int_{\mathbb{R}^{2d}} e^{g_0} dx - \log Z_H.$$

From inequality (74) we deduce that, using $g_\tau = \phi$,

$$\begin{aligned} \sup_{\varphi, \phi} \int_{\mathbb{R}^d} \rho \left(\frac{1}{2} \psi * \rho + H + \phi \right) \Big|_{t=\tau} - \log \int_{\mathbb{R}^{2d}} e^\phi dx - \gamma^2 \int_0^\tau \int_{\mathbb{R}^{2d}} \rho \left(\frac{d}{m} - \frac{1}{2} \frac{p^2}{m^2} + \Delta_p \varphi + \frac{1}{2} |\nabla_p \varphi|^2 \right) \leq \\ \leq I^\gamma(\rho) + \mathcal{F}(\rho_0) + \log \frac{\int_{\mathbb{R}^{2d}} e^{g_0} dx}{\int_{\mathbb{R}^{2d}} e^{g_\tau} dx} - \log Z_H. \end{aligned} \quad (75)$$

By setting $\varphi(q, p) = -\tilde{\varphi}(q, p) - p^2/2m$ (which is allowable since $\int H\rho < \infty$, as shown in Lemma B.3 below), the third integral in (75) becomes

$$-\gamma^2 \int_0^\tau \int_{\mathbb{R}^{2d}} \rho \left(\frac{p}{m} \nabla_p \tilde{\varphi} - \Delta_p \tilde{\varphi} + \frac{1}{2} |\nabla_p \tilde{\varphi}|^2 \right).$$

We now combine (75), the duality characterizations of free energy (70) and of the relative Fisher information (24), and estimate (73), to find

$$\mathcal{F}(\rho_\tau) + \gamma^2 \sup_{\tilde{\varphi} \in C_c^\infty(\mathbb{R} \times \mathbb{R}^{2d})} \int_0^\tau \int_{\mathbb{R}^{2d}} \rho \left(-\frac{p}{m} \nabla_p \tilde{\varphi} + \Delta_p \tilde{\varphi} - \frac{1}{2} |\nabla_p \tilde{\varphi}|^2 \right) \leq \mathcal{F}(\rho_0) + I^\gamma(\rho).$$

By a standard argument, based on the C^2 -separability of C_c^∞ , we can move the supremum inside of the time integral to find

$$\mathcal{F}(\rho_\tau) + \frac{\gamma^2}{2} \int_0^\tau \mathcal{I}(\rho_t | \mu) dt \leq \mathcal{F}(\rho_0) + I^\gamma(\rho).$$

This proves Theorem 2.3.

Proof of Lemma B.1. Formally, (71) follows from substituting in (19) $f = \psi * \rho + H + g$. To prove the statement rigorously, we use

Lemma B.3. *Let $\rho \in C([0, T]; \mathcal{P}(\mathbb{R}^{2d}))$.*

1. *The map $t \mapsto \psi * \rho_t$ is continuous from $[0, T]$ to $C_b(\mathbb{R}^d)$;*
2. *If $I^\gamma(\rho), \mathcal{H}(\rho | Z_H^{-1} e^{-H}) < \infty$, then $\int H \rho_t < \infty$ for all $t \in [0, T]$.*

Accepting this lemma for the moment, the result follows from choosing in (19) the function

$$f_n = \delta_n *_t (\chi \delta_n *_t \psi *_x \rho) + H \xi_n + g,$$

for some $g \in C^{2,1}([0, T] \times \mathbb{R}^{2d})$ and $\chi \in C_c^\infty((0, T))$. Here $\delta_n(t) := n\delta(nt)$ is an approximation of the identity, $\xi_n(x) = \xi(xn^{-2d})$ is a truncation sequence, and we write $*_t$ and $*_x$ for convolution in time and in space ($x = (q, p)$), respectively. (The convolution $\psi *_x \rho$ is the same as the notation $\psi * \rho$ used in the rest of this paper.) Upon rearranging the time convolutions, letting $n \rightarrow \infty$, using Lemma B.3, and letting χ converge to the function 1, we recover (74). \square

Proof of Lemma B.3. The first part follows from the bound $\psi \in L^1(\mathbb{R}^d) \cap C_b^1(\mathbb{R}^d)$. Fix $\varepsilon > 0$, $t \in [0, T]$, and take a sequence $t_n \rightarrow t$. For each n , choose $x_n \in \mathbb{R}^{2d}$ such that $|\psi * (\rho_t - \rho_{t_n})(x_n)| \geq \|\psi * (\rho_t - \rho_{t_n})\|_\infty - \varepsilon/2$. Since $\rho_{t_n} \rightarrow \rho_t$ narrowly, $\{\rho_{t_n}\}_n$ is tight, implying that x_n can be chosen bounded; let $x_{n_k} \rightarrow x$ as $k \rightarrow \infty$. Then

$$\begin{aligned} |(\psi * \rho_t)(x_n) - (\psi * \rho_{t_n})(x_n)| &\leq |(\psi * \rho_t)(x_n) - (\psi * \rho_t)(x)| + |(\psi * \rho_t)(x) - (\psi * \rho_{t_n})(x)| \\ &\quad + |(\psi * \rho_{t_n})(x) - (\psi * \rho_{t_n})(x_n)|. \end{aligned}$$

The last term on the right-hand side satisfies

$$\int |\psi(x - y) - \psi(x_n - y)| \rho_{t_n}(dy) \rightarrow 0 \quad \text{since } \psi(x_n - \cdot) \rightarrow \psi(x - \cdot) \text{ uniformly,}$$

and a similar argument applies to the first term. The middle term converges to zero by the narrow convergence of ρ_{t_n} to ρ_t . This proves the first part.

For the second part, we take in (19) the function $f(q, p, t) = \zeta(H(q, p))$, where $\zeta \in C^\infty([0, \infty))$ is a smooth, bounded, increasing truncation of the function $f(s) = s$, satisfying $0 \leq \zeta' \leq 1$ and $\zeta'' \leq 0$. Then we find

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \zeta(H) \rho_\tau - \int_{\mathbb{R}^{2d}} \zeta(H) \rho_0 - I^\gamma(\rho) &\leq \int_0^\tau \int_{\mathbb{R}^{2d}} \left(-\gamma \zeta' \frac{p}{m} \cdot \nabla_q \psi * \rho_t + \gamma^2 \left(\zeta'' + \frac{1}{2} \zeta'^2 - \zeta' \right) \frac{p^2}{m^2} + \gamma^2 \zeta' \frac{d}{m} \right) d\rho_t dt \\ &\leq \int_0^\tau \int_{\mathbb{R}^{2d}} \left(\frac{1}{2} \zeta' |\nabla_q \psi * \rho_t|^2 + \gamma^2 \zeta' \frac{d}{m} \right) d\rho_t dt \leq \frac{\tau}{2} \|\nabla_q \psi\|_\infty^2 + \gamma^2 \frac{d}{m} \tau. \end{aligned}$$

The result follows upon letting ζ converge to the identity.

Note that this inequality gives a bound on $\int H \rho_t$ for fixed γ , but this bound breaks down when $\gamma \rightarrow \infty$. The bound (29), which is directly derived from (28), gives a γ -independent estimate. \square

Proof of Lemma B.2. The Hopf-Cole transformation $g = 2 \log(G + 1)$ transforms equation (72a) into

$$\partial_t G + \gamma J \nabla(H + \psi * \rho_t) \cdot \nabla G + \gamma^2 \Delta_p G = \frac{\gamma^2}{2} (G + 1) \left(\Delta_p \varphi + \frac{1}{2} |\nabla_p \varphi|^2 \right), \quad (76)$$

with final datum $G_T = e^{\phi/2} - 1$.

The existence and uniqueness of a solution of this equation is standard; a proof follows along these lines:

- First approximate V by a function V_δ with the property that ∇V_δ is bounded for each δ , and $d^2 V_\delta$ is bounded uniformly in δ (this is possible by assumption (V1));
- Solve the equation (76), with V replaced by V_δ , by e.g. [Deg86, Appendix 1];
- Applying the maximum principle to weighted derivatives of G_δ , as in [Deg86, Section III], obtain δ -independent bounds on G_δ in $L^\infty([0, T] \times \mathbb{R}^{2d})$ and $L^\infty(0, T; L^1(\mathbb{R}^{2d}))$, on $(1 + |q|^2 + |p|^2)^\alpha \nabla G_\delta$ in $L^\infty([0, T] \times \mathbb{R}^{2d})$ for all $\alpha > 0$, on $\Delta_p G_\delta$ in $L^2([0, T] \times \mathbb{R}^{2d})$, and on $\partial_t G_\delta$ in $(L^\infty + L^2)([0, T] \times \mathbb{R}^{2d})$;
- Pass to the limit $\delta \rightarrow 0$, with $G_\delta \xrightarrow{*} G$ in L^∞ ; then G solves equation (76) in the pointwise sense, with the same bounds on G and its derivatives.

The estimate (73) follows from the calculation

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} \int (G + 1)^2 &= \int (G + 1) \partial_t G = \gamma^2 \int (G + 1) \left(-\Delta_p G + \frac{1}{2} (G + 1) \left(\frac{1}{2} |\nabla_p \varphi|^2 + \Delta_p \varphi \right) \right) \\ &= \gamma^2 \int |\nabla_p G - \frac{1}{2} (G + 1) \nabla_p \varphi|^2 \geq 0. \end{aligned}$$

Since $\int (G + 1)^2 = \int e^g$, (73) follows. \square

Acknowledgement

US thanks Giovanni Bonaschi, Xiulei Cao, Joep Evers and in particular Patrick van Meurs for insightful discussions regarding Theorem 2.4 and Theorem 3.2. MAP and US kindly acknowledge support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) VICI grant 639.033.008. Part of this work has appeared in Oberwolfach proceedings [PDS13].

References

- [ACP82] D. Aronson, M. G. Crandall, and L. A. Peletier. Stabilization of solutions of a degenerate nonlinear diffusion problem. *Nonl. Anal.*, 6(10):1001–1022, 1982.
- [ADPZ11] 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. *Communications in Mathematical Physics*, 307:791–815, 2011.
- [ADPZ13] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. Large deviations and gradient flows. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2005):20120341, 2013.
- [AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in mathematics ETH Zürich. Birkhäuser, 2008.
- [All92] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23:1482, 1992.
- [AMP⁺12] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction. *Calculus of Variations and Partial Differential Equations*, 44:419–454, 2012.
- [ASZ09] L. Ambrosio, G. Savaré, and L. Zambotti. Existence and stability for Fokker–Planck equations with log-concave reference measure. *Probability theory and related fields*, 145(3):517–564, 2009.
- [BCS97] L. L. Bonilla, J. A. Carrillo, and J. Soler. Asymptotic behavior of an initial-boundary value problem for the Vlasov–Poisson–Fokker–Planck system. *SIAM Journal on Applied Mathematics*, 57(5):1343–1372, 1997.
- [BDF12] A. Budhiraja, P. Dupuis, and M. Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *The Annals of Probability*, 40(1):74–102, 2012.
- [BE76] H. Brezis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *Comptes Rendus de l’Académie des Sciences de Paris, Série A*, 282:971–974, 1976.
- [BGL⁺14] D. Bakry, I. Gentil, M. Ledoux, et al. *Analysis and geometry of Markov diffusion operators*. Springer, 2014.
- [BKP85] M. Bertsch, R. Kersner, and L. A. Peletier. Positivity versus localization in degenerate diffusion equations. *Nonl. Anal.*, 9(9):987–1008, 1985.
- [BvR14] F. Barret and M. von Renesse. Averaging principle for diffusion processes via Dirichlet forms. *Potential Anal.*, 41(4):1033–1063, 2014.
- [CDG02] D. Cioranescu, A. Damlamian, and G. Griso. Periodic unfolding and homogenization. *Comptes Rendus Mathématique*, 335(1):99–104, 2002.
- [CDG08] D. Cioranescu, A. Damlamian, and G. Griso. The periodic unfolding method in homogenization. *SIAM Journal on Mathematical Analysis*, 40:1585, 2008.
- [CF06] S. Cerrai and M. Freidlin. On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom. *Probab. Theory Related Fields*, 135(3):363–394, 2006.
- [CIL92] M. Crandall, H. Ishii, and P. Lions. User’s guide to viscosity solutions of second order partial differential equations. In *Amer. Math. Soc.*, volume 27, pages 1–67, 1992.

- [DE97] P. Dupuis and R. S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*, volume 902. John Wiley & Sons, 1997.
- [Deg86] P. Degond. Global existence of smooth solutions for the Vlasov-Fokker-Planck equation in 1 and 2 space dimensions. *Annales scientifiques de l'École Normale Supérieure*, 19(4):519–542, 1986.
- [DG87] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.
- [DLP⁺15] M. H. Duong, A. Lamacz, M. A. Peletier, A. Schlichting, and U. Sharma. Quantification of coarse-graining error in Langevin and overdamped Langevin dynamics, 2015.
- [DLZ12] N. Dirr, V. Laschos, and J. Zimmer. Upscaling from particle models to entropic gradient flows. *Journal of Mathematical Physics*, 53(6), 2012.
- [dPC07] J. J. de Pablo and W. A. Curtin. Multiscale modeling in advanced materials research: challenges, novel methods, and emerging applications. *Mrs Bulletin*, 32(11):905–911, 2007.
- [DPZ13] M. H. Duong, M. A. Peletier, and J. Zimmer. GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(2951-2971), 2013.
- [DPZ14] M. H. Duong, M. A. Peletier, and J. Zimmer. Conservative-dissipative approximation schemes for a generalized kramers equation. *Mathematical Methods in the Applied Sciences*, 37(16):2517–2540, 2014.
- [DS10] S. Daneri and G. Savaré. Lecture notes on gradient flows and optimal transport. *arXiv preprint arXiv:1009.3737*, 2010.
- [Eid90] D. Eidus. The Cauchy problem for the non-linear filtration equation in an inhomogeneous medium. *J. Diff. Eqns*, 84:309–318, 1990.
- [Fel52] W. Feller. The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics*, pages 468–519, 1952.
- [FG11] J. Frank and G. A. Gottwald. The Langevin limit of the Nosé-Hoover-Langevin thermostat. *J. Stat. Phys.*, 143(4):715–724, 2011.
- [FK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2006.
- [FR07] G. Fleming and M. Ratner, editors. *Directing Matter and Energy: Five Challenges for Science and the Imagination*. Basic Energy Sciences Advisory Committee, 2007.
- [Fre04] M. Freidlin. Some remarks on the Smoluchowski-Kramers approximation. *Journal of Statistical Physics*, 117(3-4):617–634, 2004.
- [Fun84] T. Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3):331–348, 1984.
- [FW93] M. I. Freidlin and A. D. Wentzell. Diffusion processes on graphs and the averaging principle. *The Annals of Probability*, pages 2215–2245, 1993.
- [FW94] M. I. Freidlin and A. D. Wentzell. Random perturbations of Hamiltonian systems. *Mem. Amer. Math. Soc.*, 109 (523), 1994.
- [FW98] M. Freidlin and M. Weber. Random perturbations of nonlinear oscillators. *The Annals of Probability*, 26(3):925–967, 1998.

- [FW01] M. Freidlin and M. Weber. On random perturbations of Hamiltonian systems with many degrees of freedom. *Stochastic processes and their applications*, 94(2):199–239, 2001.
- [FW04] M. I. Freidlin and A. D. Wentzell. Diffusion processes on an open book and the averaging principle. *Stochastic processes and their applications*, 113(1):101–126, 2004.
- [Gho09] N. Ghoussoub. *Self-Dual Partial Differential Systems and Their Variational Principles*. Springer, 2009.
- [GOVW09] 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(2):302–351, 2009.
- [HVW12] S. Hottovy, G. Volpe, and J. Wehr. Noise-induced drift in stochastic differential equations with arbitrary friction and diffusion in the Smoluchowski-Kramers limit. *J. Stat. Phys.*, 146(4):762–773, 2012.
- [IS12] H. Ishii and P. E. Souganidis. A pde approach to small stochastic perturbations of Hamiltonian flows. *Journal of Differential Equations*, 252(2):1748–1775, 2012.
- [Kra40] H. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284 – 304, 1940.
- [Kru70] S. N. Kružkov. First order quasilinear equations in several independent variables. *Mat. USSR Sbornik*, 10(2):217–243, 1970.
- [KS00] D. Kinderlehrer and G. Stampacchia. *An Introduction to Variational Inequalities and Their Applications*. Classics in Applied Mathematics. SIAM, 2000.
- [LL10] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23(9):2131, 2010.
- [Man68] P. Mandl. *Analytical treatment of one-dimensional Markov processes*. Academia, Publishing House of the Czechoslovak Academy of Sciences, 1968.
- [Mie14] A. Mielke. On evolutionary Gamma-convergence for gradient systems. Technical Report 1915, WIAS Berlin, 2014.
- [MPR14] A. Mielke, M. A. Peletier, and D. R. M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.
- [MRS08] A. Mielke, T. Roubíček, and U. Stefanelli. Γ -limits and relaxations for rate-independent evolutionary problems. *Calculus of Variations and Partial Differential Equations*, 31(3):387–416, 2008.
- [MRS12] A. Mielke, R. Rossi, and G. Savaré. Variational convergence of gradient flows and rate-independent evolutions in metric spaces. *Milan Journal of Mathematics*, 80(2):381–410, 2012.
- [Mur87] F. Murat. A survey on compensated compactness. *Contributions to modern calculus of variations*, pages 145–183, 1987.
- [Nar94] K. Narita. Asymptotic behavior of fluctuation and deviation from limit system in the Smoluchowski-Kramers approximation for SDE. *Yokohama Math. J.*, 42(1):41–76, 1994.
- [Nay76] B. Nayroles. Deux théoremes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris, Ser. A-B*, 282:A1035–A1038, 1976.

- [Nel67] E. Nelson. *Dynamical theories of Brownian motion*, volume 17. Princeton University Press Princeton, 1967.
- [NN12] G. Nicolis and C. Nicolis. *Foundations of Complex Systems: Emergence, Information and Prediction*. World Scientific, 2012.
- [Oel84] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *The Annals of Probability*, pages 458–479, 1984.
- [OP11] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629–1653, 2011.
- [Ött05] H. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley, 2005.
- [PDS13] M. A. Peletier, M. H. Duong, and U. Sharma. Coarse-graining and fluctuations: Two birds with one stone. In *Oberwolfach Reports*, volume 10 (4), 2013.
- [PSCF05] M. Pennacchio, G. Savaré, and P. Colli Franzone. Multiscale modeling for the bioelectric activity of the heart. *SIAM J. Math. Anal.*, 37(4):1333–1370 (electronic), 2005.
- [RK82] P. Rosenau and S. Kamin. Non-linear diffusion in a finite mass medium. *Comm. Pure Appl. Math.*, 35:113–127, 1982.
- [SATS07] D. A. Stainforth, M. R. Allen, E. R. Tredger, and L. A. Smith. Confidence, uncertainty and decision-support relevance in climate predictions. *Philosophical Transactions A*, 365(1857):2145, 2007.
- [Ser11] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete and Continuous Dynamical Systems A*, 31(4):1427–1451, 2011.
- [Smo94] J. Smoller. *Shock waves and reaction-diffusion equations*. Springer, 1994.
- [SS04] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Communications on Pure and Applied Mathematics*, 57(12):1627–1672, 2004.
- [Ste08] U. Stefanelli. The Brezis–Ekeland principle for doubly nonlinear equations. *SIAM Journal on Control and Optimization*, 47:1615, 2008.
- [Tar79] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt symposium*, volume 4, pages 136–212, 1979.

Preprints ab 2012/17

- 2015-07 **Manh Hong Duong, Agnes Lamacz, Mark A. Peletier and Upanshu Sharma**
Variational approach to coarse-graining of generalized gradient flows
- 2015-06 **Agnes Lamacz and Ben Schweizer**
A negative index meta-material for Maxwell's equations
- 2015-05 **Michael Voit**
Dispersion and limit theorems for random walks associated with hypergeometric functions of type BC
- 2015-04 **Andreas Rätz**
Diffuse-interface approximations of osmosis free boundary problems
- 2015-03 **Margit Rösler and Michael Voit**
A multivariate version of the disk convolution
- 2015-02 **Christina Dörlemann, Martin Heida, Ben Schweizer**
Transmission conditions for the Helmholtz-equation in perforated domains
- 2015-01 **Frank Klinker**
Program of the International Conference
Geometric and Algebraic Methods in Mathematical Physics
March 16-19, 2015, Dortmund
- 2014-10 **Frank Klinker**
An explicit description of $SL(2, \mathbb{C})$ in terms of $SO^+(3, 1)$ and vice versa
- 2014-09 **Margit Rösler and Michael Voit**
Integral representation and sharp asymptotic results for some Heckman-Opdam hypergeometric functions of type BC
- 2014-08 **Martin Heida and Ben Schweizer**
Stochastic homogenization of plasticity equations
- 2014-07 **Margit Rösler and Michael Voit**
A central limit theorem for random walks on the dual of a compact Grassmannian
- 2014-06 **Frank Klinker**
Eleven-dimensional symmetric supergravity backgrounds, their geometric superalgebras, and a common reduction
- 2014-05 **Tomáš Dohnal and Hannes Uecker**
Bifurcation of nonlinear Bloch waves from the spectrum in the Gross-Pitaevskii equation
- 2014-04 **Frank Klinker**
A family of non-restricted $D = 11$ geometric supersymmetries
- 2014-03 **Martin Heida and Ben Schweizer**
Non-periodic homogenization of infinitesimal strain plasticity equations
- 2014-02 **Ben Schweizer**
The low frequency spectrum of small Helmholtz resonators

- 2014-01 **Tomáš Dohnal, Agnes Lamacz, Ben Schweizer**
Dispersive homogenized models and coefficient formulas for waves in general periodic media
- 2013-16 **Karl Friedrich Siburg**
Almost opposite regression dependence in bivariate distributions
- 2013-15 **Christian Palmes and Jeannette H. C. Woerner**
The Gumbel test and jumps in the volatility process
- 2013-14 **Karl Friedrich Siburg, Katharina Stehling, Pavel A. Stoimenov, Jeannette H. C. Wörner**
An order for asymmetry in copulas, and implications for risk management
- 2013-13 **Michael Voit**
Product formulas for a two-parameter family of Heckman-Opdam hypergeometric functions of type BC
- 2013-12 **Ben Schweizer and Marco Veneroni**
Homogenization of plasticity equations with two-scale convergence methods
- 2013-11 **Sven Glaser**
A law of large numbers for the power variation of fractional Lévy processes
- 2013-10 **Christian Palmes and Jeannette H. C. Woerner**
The Gumbel test for jumps in stochastic volatility models
- 2013-09 **Agnes Lamacz, Stefan Neukamm and Felix Otto**
Moment bounds for the corrector in stochastic homogenization of a percolation model
- 2013-08 **Frank Klinker**
Connections on Cahen-Wallach spaces
- 2013-07 **Andreas Rätz and Matthias Röger**
Symmetry breaking in a bulk-surface reaction-diffusion model for signaling networks
- 2013-06 **Gilles Francfort and Ben Schweizer**
A doubly non-linear system in small-strain visco-plasticity
- 2013-05 **Tomáš Dohnal**
Traveling solitary waves in the periodic nonlinear Schrödinger equation with finite band potentials
- 2013-04 **Karl Friedrich Siburg, Pavel Stoimenov and Gregor N. F. Weiß**
Forecasting portfolio-value-at-risk with nonparametric lower tail dependence estimates
- 2013-03 **Martin Heida**
On thermodynamics of fluid interfaces
- 2013-02 **Martin Heida**
Existence of solutions for two types of generalized versions of the Cahn-Hilliard equation
- 2013-01 **Tomáš Dohnal, Agnes Lamacz, Ben Schweizer**
Dispersive effective equations for waves in heterogeneous media on large time scales
- 2012-19 **Martin Heida**
On gradient flows of nonconvex functional in Hilbert spaces with Riemannian metric and application to Cahn-Hilliard equations
- 2012-18 **Robert V. Kohn, Jianfeng Lu, Ben Schweizer and Michael I. Weinstein**
A variational perspective on cloaking by anomalous localized resonance
- 2012-17 **Margit Rösler and Michael Voit**
Olshanski spherical functions for infinite dimensional motion groups of fixed rank