Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2009, Vol. 45, No. 2, 302–351
DOI: 10.1214/07-AIHP200

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit

Natalie Grunewald[a,1], Felix Otto[b,2], Cédric Villani[c,3] and Maria G. Westdickenberg[d,4]

[a] *Institute für Angewandte Mathematik, Universität Bonn. E-mail: grunewald@iam.uni-bonn.de*
[b] *Institute für Angewandte Mathematik, Universität Bonn. E-mail: otto@iam.uni-bonn.de*
[c] *UMPA, École Normale Supérieure de Lyon, and Institut Universitaire de France. E-mail: cvillani@umpa.ens-lyon.fr*
[d] *School of Mathematics, Georgia Institute of Technology. E-mail: maria@math.gatech.edu*

Received 30 May 2007; accepted 7 November 2007

Abstract. We consider the coarse-graining of a lattice system with continuous spin variable. In the first part, two abstract results are established: sufficient conditions for a logarithmic Sobolev inequality with constants independent of the dimension (Theorem 3) and sufficient conditions for convergence to the hydrodynamic limit (Theorem 8). In the second part, we use the abstract results to treat a specific example, namely the Kawasaki dynamics with Ginzburg–Landau-type potential.

Résumé. Nous étudions un système sur réseau à variable de spin continue. Dans la première partie, nous établissons deux résultats abstraits : des conditions suffisantes pour une inégalité de Sobolev logarithmique avec constante indépendante de la dimension (Théorème 3), et des conditions suffisantes pour la convergence vers la limite hydrodynamique (Theorème 8). Dans la seconde partie, nous utilisons ces résultats abstraits pour traiter un exemple spécifique, à savoir la dynamique de Kawasaki avec un potentiel de type Ginzburg–Landau.

MSC: Primary 60K35; secondary 60J25; 82B21

Keywords: Logarithmic Sobolev inequality; Hydrodynamic limit; Spin system; Kawasaki dynamics; Canonical ensemble; Coarse-graining

Contents

[1] Supported by the BMBF Förderbereich Mathematik.
[2] Supported in part by the DFG through SFB 611.
[3] Supported in part by the Institut Universitaire de France.
[4] Supported in part by the National Science Foundation under Grants No. DMS-04-02762 and DMS-07-06026.

Introduction

We consider the coarse-graining of a lattice system with continuous spin variable. It has been well known since the end of the eighties [10] that such problems can be attacked by entropy methods; the reference book [12] summarizes a decade of works in this direction. It was later discovered that logarithmic Sobolev inequalities are useful technical tools for such issues (see, e.g., [13,18]).

Establishing the desired logarithmic Sobolev inequality (LSI) with the correct scaling in system size, however, can be quite challenging. In the present paper we provide a general criterion (Theorem 3) that may be applied to a variety of situations; with its help we then deduce a criterion for the hydrodynamic limit (Theorem 8). An interesting feature of the latter result is that it is entirely constructive, yielding explicit estimates for the deviation from hydrodynamic behavior – while the traditional approach cares only about identification of the limit.

The unity of these results comes not only from the fact that the LSI is used in the proof of the hydrodynamic limit, but also from the use of a two-scale method in both cases (a macroscopic and a microscopic scale, and no more). It is very natural in hydrodynamic limits to separate microscopic and macroscopic scales, but in the context of logarithmic Sobolev inequalities this is less common.

We apply these results to a well-known example from statistical physics which, although slightly academic, has been studied (as a kind of archetypal problem) by many authors. Namely, we consider a system of spins interacting by Kawasaki dynamics with a Ginzburg–Landau-type potential. The (conservative) Kawasaki dynamics forces us to work with the canonical ensemble in which the mean m is fixed. We derive two results: First, we prove an LSI that is uniform in m and has the optimal scaling in the system size; second, we establish convergence to the hydrodynamic limit. Neither the LSI nor the hydrodynamic limit are new results (see [4,14,18] and [10,24], respectively), but the proof is conceptually simpler than previous arguments known to us – and, as mentioned earlier, the final error estimates are quantitative.

The present work leaves open two natural questions. The first one is the convergence of the microscopic entropy to the hydrodynamic entropy. The second one is the local Gibbs behavior. To limit the size of this paper, both of these issues will be addressed in a separate work.

The plan of the paper is as follows: After recalling some background in Section 1, we present our main results in Section 2. The proof of the abstract LSI result and the abstract hydrodynamic limit result is given in Sections 3 and 4, respectively. The application to Kawasaki dynamics is given in Sections 5 and 6; see in particular Section 5.1 for set-up and preparations. Finally, we include a technical appendix.

Notation

- $\mathcal{P}(X)$ stands for the set of Borel probability measures on X.
- $\mathrm{Lip}(X)$ is the set of Lipschitz functions $X \to \mathbb{R}$.
- $\mathcal{L}^k$ is the k-dimensional Lebesgue measure.
- $\mathcal{H}^k$ is the k-dimensional Hausdorff measure.
- $T_x X$ is the tangent space to X at x (in this paper X will almost always be an affine Euclidean subspace of $\mathbb{R}^N$).
- ∇ stands for gradient, Hess for Hessian, $|\cdot|$ for norm, and $\langle\cdot,\cdot\rangle$ for inner product. When there are two spaces X and Y involved, the relevant space will be indicated with a subscript, e.g., $\langle\cdot,\cdot\rangle_Y$. When no space is indicated, by default the Euclidean space X with the ℓ^2 inner product is intended.
- A^t is the transpose of the operator A.
- $\mathrm{Ran}(A)$ is the range of the operator A.
- $\mathrm{osc}_X(\Psi) = \sup_X \Psi - \inf_X \Psi$ is the oscillation of Ψ.
- $\Phi(x) = x \log x$.
- C is a positive constant that may change from line to line, or even within a line.
- Z is a positive constant enforcing unit mass of a given probability measure.
- id_X is the identity map $X \to X$.
- $f_\# \mu = \mu \circ f^{-1}$ is the image measure of μ by f.
- LSI is an abbreviation for Logarithmic Sobolev Inequality.

1. Background

1.1. *Logarithmic Sobolev inequalities*

Definition 1 (LSI). *Let X be a Riemannian manifold. A probability measure $\mu \in \mathcal{P}(X)$ is said to satisfy an LSI with constant $\rho > 0$ (in short: $\mathrm{LSI}(\rho)$) if, for any locally Lipschitz, nonnegative function $f \in L^1(\mu)$,*

$$\int \Phi(f)\,\mathrm{d}\mu - \Phi\left(\int f\,\mathrm{d}\mu\right) \le \frac{1}{\rho}\int \frac{|\nabla f|^2}{2f}\,\mathrm{d}\mu.$$

This inequality is a powerful tool for studying particle systems. First of all, it implies convergence to equilibrium for the diffusion equation naturally associated to the measure μ. It also implies the spectral gap inequality via "linearization." Most importantly, LSI is well-adapted for developing estimates that are independent of the dimension of the system (see Criterion 1 below and the discussion following). Nicely written introductions to LSI can be found in [9,16,21]. As for their application to spin systems, among many references one can quote [4].

The following three fundamental criteria, stated here by order of increasing complexity, are the main results commonly used to derive LSI's:

Criterion 1 (Tensorization principle). *If $\mu_1 \in \mathcal{P}(X_1)$ and $\mu_2 \in \mathcal{P}(X_2)$ satisfy $\mathrm{LSI}(\rho_1)$ and $\mathrm{LSI}(\rho_2)$ respectively, then $\mu_1 \otimes \mu_2$ satisfies $\mathrm{LSI}(\min\{\rho_1, \rho_2\})$.*

Criterion 2 (Holley–Stroock perturbation lemma). *Let $\mu \in \mathcal{P}(X)$ satisfy $\mathrm{LSI}(\rho)$, and let $\delta\Psi : X \to \mathbb{R}$ be a bounded function. Let $\tilde{\mu} \in \mathcal{P}(X)$ be defined via*

$$\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}(x) = \frac{1}{Z}\exp\bigl(-\delta\Psi(x)\bigr).$$

Then $\tilde{\mu}$ satisfies $\mathrm{LSI}(\tilde{\rho})$, where

$$\tilde{\rho} = \rho \cdot \exp(-\mathrm{osc}_X\,\delta\Psi). \tag{1}$$

Criterion 3 (Bakry–Émery theorem). *Let X be a K-dimensional Riemannian manifold, let $H \in C^2(X)$, and let $\mu \in \mathcal{P}(X)$ be defined by*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{H}^K}(x) = \frac{1}{Z}\exp\bigl(-H(x)\bigr).$$

If there is a constant $\rho > 0$ such that Hess $H \geq \rho$, *or more explicitly*

$$\forall x \in X,\ \forall v \in T_x X, \quad \bigl\langle v, \operatorname{Hess} H(x) v\bigr\rangle \geq \rho |v|^2,$$

then μ satisfies LSI(ρ).

These facts are fundamental. The tensorization property is one of the main reasons why LSI's are suitable for large (or infinite) configuration spaces ([8], Remark 3.3). The Bakry–Émery theorem [1,20] is the simplest general sufficient condition for LSI. Taken together, the Holley–Stroock perturbation lemma [11] and Bakry–Émery theorem considerably extend the class of probability measures known to satisfy LSI.

Still, these three criteria are not always sufficient to treat situations of interest, as can be seen from the following basic problem, which was our initial motivation. Consider a probability measure μ on $\mathbb{R}^N$ (N large) under which all particles are independent save for one constraint, that the mean is fixed. (This represents an interest in a Gibbs state distribution for a physical system having one conservation law.) Let $\mu_1(\mathrm{d}x) = \exp(-\psi(x))\,\mathrm{d}x$ be the distribution of a free particle; then μ is an N-fold tensor product of μ_1, conditioned by the affine constraint $(1/N)\sum x_i = m$.

Let us try to establish an LSI for μ, with constants independent of N. If it were not for the conditioning, then the tensorization principle would apply. Given the mean constraint, our only hope is convexity. Indeed, if ψ were uniformly convex, then μ would take the form $\exp(-\Psi(x))\,\mathrm{d}\mathcal{H}^{N-1}(x)$ with Ψ uniformly convex (as the restriction of the uniformly convex function $\sum \psi(x_i)$); hence the Bakry–Émery principle would apply (as noticed in [13]). But if ψ is nonconvex this line of reasoning is doomed. So the problem lies in the combination of *conditioning* and *nonconvexity*.

In Theorem 3 below we shall present a new sufficient condition for LSI. Roughly speaking, it says that if one can decompose the system into a microscopic and macroscopic scale and prove LSI on each scale separately, then one can derive an LSI for the original measure. In Section 2.4, we will show how this theorem can be used to deduce a uniform LSI for the example just described. At first glance, the situation looks bad: We have replaced the problem of proving one LSI with the problem of proving *two* LSI's. The main insight is that by choosing the macroscopic size sufficiently large, there is some extra convexity to exploit: In the language of physics, the coarse-grained Hamiltonian is convex.

1.2. *Hydrodynamic limit*

The subject of hydrodynamic limits of particle systems is an old topic, reviewed in [12]. The field has mainly been developed via specific model problems, without achieving very general results – which is natural given the complexity of the subject. Still one usually identifies two main methods: the "GPV" method introduced in [10], based on the convergence of martingales and on entropy estimates; and Yau's entropy method [24], which is based on a sophisticated Gronwall-type estimate for a relative entropy functional. Yau's method is simpler and gives stronger results, but it makes stronger assumptions on the initial data (closeness to hydrodynamic behavior in the sense of relative entropy rather than in the sense of macroscopic observables).

Our method is intermediate between these two main strategies. On the one hand, we shall make rather weak assumptions on the behavior of the initial data; on the other hand, we shall use a Gronwall-type estimate for a well-chosen functional. The use of logarithmic Sobolev inequalities will lead us to make the rather stringent assumption of quadratic growth of the interaction potential at infinity (an LSI requires at least quadratic growth at infinity [15], Section 5.1), but this does not rule out the possibility of an extension of LSI techniques for potentials with subquadratic growth (as was done for instance in [22] for the convergence to equilibrium of the Fokker–Planck equation).

2. Main results

2.1. *Microscopic and macroscopic variables*

In the sequel X and Y will be two Euclidean spaces; think of X as the space of microscopic variables and of Y as the space of macroscopic variables. There is a linear operator $P : X \to Y$ which to every $x \in X$ associates the corresponding macroscopic profile $y = Px$. We shall assume

$$PNP^t = \mathrm{id}_Y \tag{2}$$

for some (large) $N \in \mathbb{N}$, which we think of as the size of the system, measured at microscopic scale.

Remark 2. *To motivate* (2), *think of the trivial case* $X = \mathbb{R}^N$ (*with its usual Euclidean structure*), $Y = \mathbb{R}$, *and* P *is the average*: $Px = (1/N)\sum x_i$. *Then* $P^t y = (1/N)(y, \dots, y)$ *and* (2) *clearly holds.*

The symmetric operator NP^tP is the orthogonal projection of X to $(\ker P)^\perp$. (Indeed $\ker(NP^tP) = \ker P$ and $(NP^tP)^2 = NP^t(PNP^t)P = NP^tP$.) This induces a decomposition of X into *macroscopic variables* in $Y \simeq (\ker P)^\perp$ and *microscopic fluctuations* in $\ker P$. Consider the probability measure on X

$$\mu(\mathrm{d}x) = \exp\bigl(-H(x)\bigr)\,\mathrm{d}x,$$

where $\mathrm{d}x$ is a shorthand for Lebesgue measure on X. The decomposition of variables introduces

(i) *A decomposition of measures:* Define $\bar\mu = P_\#\mu$ as the distribution of macroscopic variables, and let $\mu(\mathrm{d}x|y)$ be the conditional measure of x given $Px = y$. For each y, $\mu(\mathrm{d}y|x)$ is a probability measure on X, and we have

$$\mu(\mathrm{d}x) = \mu(\mathrm{d}x|y)\bar\mu(\mathrm{d}y), \tag{3}$$

or more explicitly: for any test function ξ, $\int \xi\,\mathrm{d}\mu = \int_Y (\int \xi(x)\mu(\mathrm{d}x|y))\bar\mu(\mathrm{d}y)$.

(ii) *A decomposition of gradients:* If f is a smooth function of x, its gradient ∇f can be decomposed into a macroscopic gradient and a fluctuation gradient:

$$\nabla^{\mathrm{macro}} f(x) = NP^tP\nabla f(x), \qquad \nabla^{\mathrm{fluct}} f(x) = \bigl(\mathrm{id}_X - NP^tP\bigr)\nabla f(x). \tag{4}$$

Indeed, $\mathrm{id}_X - NP^tP$ is the orthogonal projection onto $\ker P$, which is the tangent space to the fiber $\{Px = y\}$.

(iii) *An interaction between scales:* We will measure interactions between the microscopic and macroscopic scales via

$$\kappa := \max\Bigl\{\bigl\langle \mathrm{Hess}\,H(x)\cdot u, v\bigr\rangle;\, u \in \mathrm{Ran}\bigl(NP^tP\bigr),\ v \in \mathrm{Ran}\bigl(\mathrm{id}_X - NP^tP\bigr);\ |u| = |v| = 1\Bigr\}, \tag{5}$$

which we will require to be finite.

The macroscopic coarse-graining of the microscopic measure induces a natural coarse-graining of the microscopic Hamiltonian. Define

$$\bar H(y) := -\frac{1}{N}\log\biggl(\frac{\mathrm{d}\bar\mu}{\mathrm{d}y}\biggr), \tag{6}$$

so $\bar\mu(\mathrm{d}y) = \exp(-N\bar H(y))\,\mathrm{d}y$. Informally, we can express $\bar H$ as

$$\bar H(y) = -\frac{1}{N}\log\int_{\{Px=y\}} \exp\bigl(-H(x)\bigr)\,\mathrm{d}x.$$

This coarse-grained Hamiltonian will play a crucial role in our results (as in many other works such as [10,14]).

2.2. *An abstract two-scale criterion for LSI*

With the setup and notation of Section 2.1 we have the following result:

Theorem 3 (Two-scale logarithmic Sobolev inequality). *Let $\mu(\mathrm{d}x) = \exp(-H(x))\,\mathrm{d}x$ be a probability measure on X, and let $P : X \to Y$ satisfy* (2). *Assume that*

(i) *For κ given by* (5), *we have $\kappa < +\infty$*;
(ii) *There is $\rho > 0$ such that $\mu(\mathrm{d}x|y)$ satisfies* $\mathrm{LSI}(\rho)$ *for all y*;
(iii) *There is $\lambda > 0$ such that $\bar{\mu}$ satisfies* $\mathrm{LSI}(\lambda N)$.

Then μ satisfies $\mathrm{LSI}(\hat{\rho})$ *with*

$$\hat{\rho} := \frac{1}{2}\left(\rho + \lambda + \frac{\kappa^2}{\rho} - \sqrt{\left(\rho + \lambda + \frac{\kappa^2}{\rho}\right)^2 - 4\rho\lambda}\right) > 0. \tag{7}$$

Remark 4 (On the assumptions). *Assumption* (i) *says that the Hessian of the potential does not induce a strong coupling between the microscopic and macroscopic scales. Assumption* (ii) *says that the distribution of microscopic fluctuations satisfies a logarithmic Sobolev inequality, independently of the macroscopic state: explicitly, for any $y \in Y$ and any positive $f \in \mathrm{Lip}(X)$,*

$$\int \Phi\big(f(x)\big)\mu(\mathrm{d}x|y) - \Phi\left(\int f(x)\mu(\mathrm{d}x|y)\right) \le \frac{1}{\rho}\int \frac{|\nabla^{\mathrm{fluct}} f(x)|^2}{2f(x)}\mu(\mathrm{d}x|y).$$

Finally, assumption (iii) *says that the logarithmic Sobolev constant at macroscopic scale grows linearly with the system size* (*which is the natural scaling*).

Remark 5 (Zero interactions). *In the particular case that microscopic and macroscopic scales behave independently* ($\kappa = 0$), *Theorem* 3 *returns the factorization principle* $\hat{\rho} = \min\{\rho, \lambda\}$ (*cf. Criterion* 1 *in Section* 1.1).

Remark 6 (Checking the assumptions). *Assumption* (i) *is often easy to check*; *it suffices for instance that H has bounded Hessian. Then Theorem* 3 *leaves us with two logarithmic Sobolev inequalities*, *of very different natures*, *to verify. The idea exploited in the application in Section* 2.4 *is the following*: *As long as the microscopic scale is finite*, *the microscopic LSI can be proved by the usual methods. On the other hand by choosing the microscopic scale large*, *one averages over many variables. This averaging should result in a convexification of the macroscopic Hamiltonian $\bar{H}$, so that the Bakry–Émery theorem can be applied. This convexification phenomenon is well known in statistical mechanics*, *where it is used in conjunction with renormalization group methods. Theorem* 3 *as it is stated would correspond to just one step of the renormalization group*, *which in principle is reasonable as long as one stays away from phase transition.*

Remark 7. *In a different spirit but with similar ingredients*, *recent work of Blower and Bolley* [2], *Theorem* 1.3, *also proves LSI via a decomposition of the system into components.*

2.3. *An abstract result for the hydrodynamic limit*

Let A be a positive definite symmetric linear operator $X \to X$. We equip X with a (Gibbs) measure μ and we consider the (reversible) stochastic dynamics on X described by the time-evolution

$$\frac{\partial}{\partial t}(f\mu) = \nabla \cdot (A\nabla f\mu) \tag{8}$$

(to be understood in weak sense; that is, for any smooth test function ξ, one has $(\mathrm{d}/\mathrm{d}t)\int \xi(x) f(t,x)\mu(\mathrm{d}x) = -\int \nabla\xi(x) \cdot A\nabla f(t,x)\mu(\mathrm{d}x)$). Here both μ and $f\mu$ are probability measures on X and the unknown $f(t,x)$ is the microscopic density with respect to $\mu(\mathrm{d}x)$.

We are interested in the distribution $\bar{f}(t, y)\bar{\mu}(\mathrm{d}y)$ of $y = Px$ under $f(t, x)\mu(\mathrm{d}x)$; again we assume that P satisfies (2). Since $\bar{\mu} = P_{\#}\mu$, we have

$$\bar{f}(t, y) = \int f(t, x)\mu(\mathrm{d}x|y).$$

Define the coarse-grained operator $\bar{A} : Y \to Y$ by

$$(\bar{A})^{-1} = PA^{-1}NP^t. \tag{9}$$

Let $x(t)$ denote the process associated with the forward equation (8). We shall compare $Px(t)$, the projection of the microscopic dynamics, with the solution $\eta(t) \in Y$ of the deterministic equation

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\bar{A}\nabla_Y \bar{H}(\eta), \tag{10}$$

where ∇_Y stands for the gradient on Y. (Recall that the notion of gradient depends on the inner product on Y; see Section 5.1 for a specific example.) In general $Px(t) \neq \eta(t)$, because the coarse-graining operator P is not compatible with the kinetics, in the sense that $PA \neq \bar{A}P$.

The key ingredient for the hydrodynamic limit is that fluctuations are strongly penalized by the dynamics, in the sense that

$$\forall x \in X \quad \left|\left(\mathrm{id}_X - NP^tP\right)x\right|^2 \le \frac{C}{(\dim Y)^2}\langle x, Ax\rangle \quad \text{as} \dim Y \uparrow \infty$$

(cf. assumption (vi) below); this limit will correspond to considering finer and finer details of the microscopic dynamics. (The rate $1/(\dim Y)^2$ is natural for the application we have in mind, but the precise form is not important; any quantity that is $\mathrm{o}(1)$ as $(\dim Y) \uparrow \infty$ would do.)

Theorem 8 (Hydrodynamic limit). *Let $\mu(\mathrm{d}x) = \exp(-H(x))\,\mathrm{d}x$ be a probability measure on X and let $P : X \to Y$ satisfy (2). Assume that:*

(i) *For κ given by (5), we have $\kappa < +\infty$;*
(ii) *There is $\rho > 0$ such that $\mu(\mathrm{d}x|y)$ satisfies $\mathrm{LSI}(\rho)$ for all y;*
(iii′) *There is $\lambda > 0$ such that $\langle \tilde{y}, \operatorname{Hess} \bar{H}(y)\tilde{y}\rangle_Y \ge \lambda\langle \tilde{y}, \tilde{y}\rangle_Y$;*
(iv) *There is $\alpha > 0$ such that $\int |x|^2\mu(\mathrm{d}x) \le \alpha N$;*
(v) *There is $\beta > 0$ such that $\inf_{y\in Y} \bar{H}(y) \ge -\beta$.*

Define $M := \dim Y$ and let $A : X \to X$ be a symmetric linear operator such that:

(vi) *There is $\gamma > 0$ such that for all $x \in X$, $|(\mathrm{id}_X - NP^tP)x|^2 \le \gamma M^{-2}\langle x, Ax\rangle_X$.*

Let $f(t, x)$ and $\eta(t)$ solve (8) and (10) respectively, with respective initial data $f(0, \cdot)$ and η_0 satisfying

(vii) $\int f(0, x) \log f(0, x)\mu(\mathrm{d}x) \le C_1N$; $\bar{H}(\eta_0) \le C_2$.

Define

$$\Theta(t) := \frac{1}{2N}\int \left\langle \left(x - NP^t\eta(t)\right), A^{-1}\left(x - NP^t\eta(t)\right)\right\rangle f(t, x)\mu(\mathrm{d}x). \tag{11}$$

Then for any $T > 0$ we have, with $\hat{\rho}$ given by (7),

$$\begin{aligned}
&\max\left\{\sup_{0<t\le T}\Theta(t), \frac{\lambda}{2}\int_0^T \left(\int_Y \left|y - \eta(t)\right|_Y^2 \bar{f}(t, y)\bar{\mu}(\mathrm{d}y)\right)\mathrm{d}t\right\} \\
&\quad\le \Theta(0) + T\left(\frac{M}{N}\right) + \left(\frac{C_1\gamma\kappa^2}{2\lambda\rho^2}\right)\frac{1}{M^2} \\
&\qquad + \left[\sqrt{2T\gamma}\left(\alpha + \frac{2C_1}{\hat{\rho}}\right)^{1/2}\left(C_1^{1/2} + (C_2 + \beta)^{1/2}\right)\right]\frac{1}{M}.
\end{aligned}$$

Remark 9 (On the assumptions). *By the Bakry–Émery theorem, assumption* (iii$'$) *implies assumption* (iii) *from Theorem* 3. *Notice that assumption* (iii$'$) *also implies that* $\bar{H}$ *is bounded below for any finite* M, *so assumption* (v) *requires only that this lower bound is uniform in* M.

Remark 10 (LSI). *By the above remark, the assumptions of Theorem* 8 *imply in particular that all of the assumptions of Theorem* 3 *are satisfied, so* μ *satisfies* $\mathrm{LSI}(\hat{\rho})$ *for some* $\hat{\rho} > 0$. *We will use this fact in the proof of Theorem* 8.

Remark 11 (On the result). *Think of* $NP^t\eta(t)$ *as a microscopic state that is "purely hydrodynamic" and has* $\eta(t)$ *as its macroscopic profile. Then* $\Theta(t)$ *is a "weak" way to quantify the deviation of* x *from hydrodynamic behavior.* (*In slightly pedantic terms,* Θ *is, up to a factor* $(2N)^{-1}$, *the square of the quadratic Monge–Kantorovich–Wasserstein distance* [20,23] *between* $f\mu$ *and* $\delta_{NP^t\eta}$, *where distances in* X *are measured with the scalar product associated to* A^{-1}.)

The strong quadratic $L^2(Y)$ *norm*

$$\int_Y \left|y - \eta(t)\right|_Y^2 \bar{f}(t, y)\bar{\mu}(\mathrm{d}y)$$

is another way to measure the deviation from hydrodynamic behavior. (*It is the square of the quadratic Monge–Kantorovich–Wasserstein distance between* $P_\#(f\mu)$ *and* $\delta_\eta = P_\#(\delta_{NP^t\eta})$.) *A main difference between this functional and the preceding one is that here hydrodynamic fluctuations are killed by the action of* P, *while in the definition of* $\Theta(t)$ *they are penalized by the action of* A^{-1}.

The following corollary of Theorem 3 makes the hydrodynamic limit more explicit.

Corollary 12 (Propagation of hydrodynamic behavior). *Consider a sequence* $\{X_\nu, Y_\nu, P_\nu, A_\nu, \mu_\nu, f_{0,\nu}, \eta_{0,\nu}\}_{\nu=1}^\infty$ *of data satisfying the assumptions of Theorem* 8 *for every* ν *with uniform constants* $\lambda, \rho, \kappa, \alpha, \beta, \gamma, C_1, C_2$. *Suppose that*

$$M_\nu \uparrow \infty; \qquad N_\nu \uparrow \infty; \qquad \frac{N_\nu}{M_\nu} \uparrow \infty; \tag{12}$$

further assume that the initial microscopic data satisfy

$$\lim_{\nu\uparrow\infty} \frac{1}{N_\nu} \int \left(x - N_\nu P_\nu^t \eta_{0,\nu}\right) \cdot A_\nu^{-1}\left(x - N_\nu P_\nu^t \eta_{0,\nu}\right) f_{0,\nu}(x)\mu_\nu(\mathrm{d}x) = 0. \tag{13}$$

Then for any $T > 0$,

(a) *The microscopic variables are close to the solution of* (10) *in the weak norm induced by* A_ν^{-1}, *uniformly in* $t \in (0, T)$:

$$\lim_{\nu\uparrow\infty} \sup_{0\le t\le T} \frac{1}{N_\nu} \int \left(x - N_\nu P_\nu^t \eta\right) \cdot A_\nu^{-1}\left(x - N_\nu P_\nu^t \eta\right) f(t, x)\mu(\mathrm{d}x) = 0; \tag{14}$$

(b) *The macroscopic variables are close to the solution of* (10) *in the strong* $L^2(Y)$ *norm, in the time-integrated sense*:

$$\lim_{\nu\uparrow\infty} \int_0^T \int_Y \left|y - \eta(t)\right|_Y^2 \bar{f}(y)\bar{\mu}(\mathrm{d}y)\,\mathrm{d}t = 0. \tag{15}$$

Remark 13. *For Corollary* 12 *to be applicable in practice, one should have an explicit representation of the limiting behavior of* $\eta(t)$ *as* $\nu \uparrow \infty$, *after having embedded all spaces* Y_ν *into a single functional space. An example will be presented in the next subsection.*

2.4. *Applications to the Kawasaki dynamics*

Logarithmic Sobolev inequality

As an application of Theorem 3, we consider an L-periodic lattice system with continuous spin variables governed by a Ginzburg–Landau-type potential $\psi : \mathbb{R} \to \mathbb{R}$. We shall assume that

$$\psi(x) = \frac{1}{2}x^2 + \delta\psi(x), \qquad \|\delta\psi\|_{C^2(\mathbb{R})} < +\infty. \tag{16}$$

(Think for instance of a double-well potential with quadratic growth at infinity.) The *grand canonical* measure $\mu_N \in \mathcal{P}(\mathbb{R}^N)$ has density

$$\frac{\mathrm{d}\mu_N}{\mathrm{d}\mathcal{L}^N}(x) = \frac{1}{Z}\exp\left(-\sum_{i=1}^N \psi(x_i)\right),$$

where $N = L^d$ is the number of sites in a period cell.

Now we shall take into account constraints of *fixed mean spin*. Let $X_{N,m}$ be the $(N-1)$-dimensional hyperplane with mean $m \in \mathbb{R}$:

$$X_{N,m} = \left\{(x_1, \ldots, x_N) \in \mathbb{R}^N;\ \frac{1}{N}\sum_{i=1}^N x_i = m\right\}$$

equipped with the ℓ^2 inner product,

$$\langle x, \tilde{x}\rangle_{X_{N,m}} := \sum x_i \tilde{x}_i.$$

For given m, we define the probability measure $\mu_{N,m} \in \mathcal{P}(X_{N,m})$ as the *restriction* of μ_N to $X_{N,m}$. In other words,

$$\frac{\mathrm{d}\mu_{N,m}}{\mathrm{d}\mathcal{H}^{N-1}}(x) = \frac{1}{Z}1_{(1/N)\sum x_i = m}\exp\left(-\sum_{i=1}^N \psi(x_i)\right). \tag{17}$$

The measure $\mu_{N,m}$ is called the *canonical* ensemble. It gives the distribution of the random variables $x_1, \ldots, x_N$ conditioned on the event that their mean value is given by m, that is, $(1/N)\sum_{i=1}^N x_i = m$.

We will show that the following result can be deduced from Theorem 3:

Theorem 14. *Let ψ satisfy (16) and let $\mu_{N,m}$ be defined by (17). Then there exists $\rho > 0$ such that for any $N \in \mathbb{N}$ and any $m \in \mathbb{R}$, $\mu_{N,m}$ satisfies* LSI(ρ).

Remark 15 (From Glauber to Kawasaki). *Explicitly, the conclusion of Theorem 14 is that for any Lipschitz density function $f : X_{N,m} \to \mathbb{R}_+$, we have*

$$\int \Phi(f)\,\mathrm{d}\mu_{N,m} - \Phi\left(\int f\,\mathrm{d}\mu_{N,m}\right) \le \frac{1}{\rho}\int \frac{1}{2f}\sum_{i=1}^N \left(\frac{\partial f}{\partial x_i}\right)^2 \mathrm{d}\mu_{N,m}, \tag{18}$$

where f has been extended to be constant in the direction normal to $X_{N,m}$. This bound is given in terms of the Dirichlet form associated to the Glauber (mean-field) dynamics. Notice that geometry plays no role.

As already noted in [5], (18) *also implies a logarithmic Sobolev inequality for the Kawasaki dynamics via the discrete Poincaré inequality. For instance, in $d = 1$, if f satisfies $\sum_{i=1}^N \frac{\partial f}{\partial x_i} = 0$ (i.e., if f is constant normal to $X_{N,m}$), then we have*

$$\sum_{i=1}^N \left(\frac{\partial f}{\partial x_i}\right)^2 \le CN^2 \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_{i+1}}\right)^2, \tag{19}$$

which together with (18) *implies the Kawasaki bound*:

$$\int \Phi(f)\,\mathrm{d}\mu_{N,m} - \Phi\left(\int f\,\mathrm{d}\mu_{N,m}\right) \le \frac{CN^2}{\rho}\int \frac{1}{2f}\sum_{i=1}^{N}\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_{i+1}}\right)^2 \mathrm{d}\mu_{N,m}. \tag{20}$$

For a d-dimensional lattice the result is similar: *Take $N = L^d$, where L is the period*; *then we obtain* (20) *with N^2 replaced by L^2. This is the optimal scaling in d dimensions* [25].

Remark 16 (Coarse-graining). *In order to use Theorem 3, the main work will go into verifying assumption* (iii). *We will rely on the idea of coarse-graining from physics. To be more specific*: *Setting $X = X_{N,m}$ and $Y = Y_{M,m} := X_{M,m}$ with* (*say*) *$N = K \cdot M$, we will define the projection operator $P_{N,K} : X_{M,m} \to Y_{M,m}$ as $P_{N,K}x = y$, where each "block spin" is the average over the block*:

$$y_j = \frac{1}{K}\sum_{i=(j-1)K+1}^{jK} x_i.$$

The idea is that for sufficiently large block spin size K, the coarse-grained Hamiltonian is convex (*see Lemma* 29).

As mentioned earlier, the result in Theorem 14 is not new: the well-known Lu–Yau martingale method has already been used to prove LSI for the canonical ensemble with a nonconvex potential [4,14,18]. Specifically, LSI for a bounded perturbation of a Gaussian potential and Kawasaki dynamics is proved in [14] via the martingale method. An adaptation of the method in [4] extends the result to the (stronger) bound for Glauber dynamics. So our contribution here is not a new result but rather a new technique.

For completeness and to contrast with the method presented in this paper, we briefly summarize the martingale method. The first step is to establish LSI for the one-site marginals. Subsequently, one seeks a recursive relationship for the N-site LSI constant in terms of the $(N-1)$-site LSI constant. Turning to the conditional expectations

$$f_k := \mathbb{E}_{\mu_{N,m}}(f|x_1,\ldots,x_k),$$

one appeals to a Markovian decomposition of the relative entropy into a sum of terms of the form

$$a_k := \mathbb{E}_{\mu_{N,m}}(f_k \log f_k - f_{k-1}\log f_{k-1}),$$

each of which depends only on a single spin. After applying the single-site LSI to each term, one wants to conclude by bounding the derivatives of a_k in terms of the derivatives of f. The central ingredient involves estimating the covariance terms from the Markovian decomposition by a variance term and a gradient term. Clever but elementary estimates produce the desired recursive relation and complete the argument.

Our method is more simple-minded. The martingale method, with the one-site distributions, the control of covariances on large enough blocks, and the recursive relationship between the LSI constant on $(N-1)$-blocks and N-blocks, operates on several scales. Ours operates on just two: the coarse measure on the blocks, and the fine measure on the microscale. Moreover, we require just one thing from equilibrium statistical mechanics: the strict convexity of the limiting free energy. This is a natural object on which to rely; it is precisely the strict convexity of the limit that rules out phase transition.

Hydrodynamic limit

Next we shall consider the microscopic Kawasaki dynamics and go to the hydrodynamic limit. For simplicity, we restrict to $d = 1$. Consider the function ψ and the space and measure $X_{N,0}$ and $\mu_{N,0}$ from above. (This choice of setting the mean to 0 is arbitrary.) The microscopic Kawasaki dynamics is governed by the $N \times N$ matrix $A = (A_{ij})$ defined by

$$A_{ij} = N^2(-\delta_{i,j-1} + 2\delta_{i,j} - \delta_{i,j+1}). \tag{21}$$

(In the indices of the Kronecker symbols, by convention $N + 1 = 1$ and $0 = N$.)

We shall identify $X = X_{N,0}$ with the space $\bar{X}$ of piecewise constant functions on $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$:

$$\bar{X} = \left\{ \bar{x} : \mathbb{T}^1 \to \mathbb{R};\ \bar{x} \text{ is constant on } \left(\frac{j-1}{N}, \frac{j}{N} \right],\ j = 1, \dots, N \right\}.$$

By convention, the “step function associated to $x \in X$” denotes the step function $\bar{x} \in \bar{X}$ with

$$\bar{x}(\theta) = x_j, \quad \theta \in \left(\frac{j-1}{N}, \frac{j}{N} \right]. \tag{22}$$

Similarly, the “vector associated to $\bar{x} \in \bar{X}$” will denote the vector $x \in X$ with

$$x_j = \bar{x}\left(\frac{j}{N} \right).$$

The L^2 norm on $\bar{X}$ is not well adapted to study macroscopic profiles. Instead it will be convenient to use the H^{-1} norm defined as follows: If $f : \mathbb{T}^1 \to \mathbb{R}$ is a function (say locally integrable) with zero mean, define

$$\|f\|^2_{H^{-1}} = \int_{\mathbb{T}^1} w^2(\theta)\, \mathrm{d}\theta; \qquad w' = f, \qquad \int w(\theta)\, \mathrm{d}\theta = 0. \tag{23}$$

The closure of $\bar{X}$ in the H^{-1} norm is the usual negative Sobolev space $H^{-1}(\mathbb{T}^1)$.

The hydrodynamic limit of the Kawasaki dynamics is captured by:

Theorem 17. *Assume that ψ satisfies (16). Let $f = f(t, x)$ be a time-dependent probability density on $(X_N, \mu_{N,0})$ solving*

$$\frac{\partial}{\partial t}(f \mu_{N,0}) = \nabla \cdot (A \nabla f \mu_N),$$

where $f(0, \cdot) = f_0(\cdot)$ satisfies

$$\int f_0(x) \log f_0(x) \mu_{N,0}(\mathrm{d}x) \le CN \tag{24}$$

for some constant $C > 0$. Assume that

$$\lim_{N \uparrow \infty} \int \|\bar{x} - \zeta_0\|^2_{H^{-1}} f_0(x) \mu_{N,0}(\mathrm{d}x) = 0 \tag{25}$$

for some $\zeta_0 \in L^2(\mathbb{T}^1)$ (initial macroscopic profile) with $\int \zeta_0\, \mathrm{d}\theta = 0$. Then for any $T > 0$, we have

$$\lim_{N \uparrow \infty} \sup_{0 \le t \le T} \int \left\| \bar{x} - \zeta(t, \cdot) \right\|^2_{H^{-1}} f(t, x) \mu_{N,0}(\mathrm{d}x) = 0, \tag{26}$$

where ζ is the unique weak solution of the nonlinear parabolic equation

$$\begin{cases} \dfrac{\partial \zeta}{\partial t} = \dfrac{\partial^2}{\partial \theta^2} \varphi'(\zeta), \\ \zeta(0, \cdot) = \zeta_0, \end{cases} \tag{27}$$

and φ is the Cramér transform of ψ, i.e.

$$\varphi(m) = \sup_{\sigma \in \mathbb{R}} \left\{ \sigma m - \log \int_{\mathbb{R}} \exp\big(\sigma x - \psi(x)\big)\, \mathrm{d}x \right\}.$$

Remark 18. *The precise meaning of "weak solution" in the above statement will be made precise later* (*see Definition* 32).

Remark 19 (Nontrivial set of initial conditions). *In some sense*, *assumptions* (24) *and* (25) *compete with each other*, *since* (24) *requires that the initial data be "sufficiently random," while* (25) *requires that the initial data be close in* H^{-1} *to the deterministic quantity* ζ_0. *In light of* [17], *it seems likely that the solution with deterministic initial data could be shown to satisfy* (24) *after an initial layer in time*, *but we will not try to extend our results to this case here*. *We will however give the following simple example*, *to demonstrate that there are initial conditions satisfying both* (24) *and* (25).

Consider $\zeta_0 = 0$ and take as initial datum a "local Gibbs state" defined by

$$f(x,0) = \exp\left(\sum_{i=1}^{N} \delta\psi(x_i)\right) \frac{\int_X \exp(-\sum_{i=1}^{N-1} 1/2x_i^2 + \delta\psi(x_i))\mathcal{H}^{N-1}(\mathrm{d}x)}{\int_X \exp(-\sum_{i=1}^{N-1} 1/2x_i^2)\mathcal{H}^{N-1}(\mathrm{d}x)}.$$

Using (16), it is not hard to see that (24) is satisfied. To see that (25) is also satisfied, we will use the fact that the function w from (23) above is the antiderivative of f with the smallest L^2 norm. Therefore

$$\langle \bar{x}, \bar{x}\rangle_{H^{-1}} \le \int_{\mathbb{T}^1} \tilde{w}(\theta)^2 \,\mathrm{d}\theta, \quad \text{where } \tilde{w}(\theta) = \frac{1}{N}\sum_{j=1}^{i-1} x_j + x_i\left(\theta - \frac{i-1}{N}\right),$$

for $\theta \in [(i-1)/N, i/N)$ and $i = 1, \dots, N$. Directly calculating the L^2 norm of $\tilde{w}$ and substituting $f(x,0)\mu(\mathrm{d}x)$, we deduce that

$$\begin{aligned}
&\int_X \int_{\mathbb{T}^1} \tilde{w}(\theta)^2 \,\mathrm{d}\theta f(x,0)\mu_{N,0}(\mathrm{d}x) \\
&\quad = \frac{1}{N^3}\frac{1}{Z}\int_X \sum_{i=1}^{N}\left(\frac{1}{3}x_i^2 + x_i\sum_{j=1}^{i-1} x_j + \left(\sum_{j=1}^{i-1} x_j\right)^2\right)\exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2\right)\mathcal{H}^{N-1}(\mathrm{d}x) \\
&\quad \le \frac{1}{N^3}\frac{1}{Z}\int_X \sum_{i=1}^{N}\left(\frac{5}{6}x_i^2 + \frac{3}{2}\left(\sum_{j=1}^{i-1} x_j\right)^2\right)\exp\left(-\frac{1}{2}\sum_{i=1}^{N} x_i^2\right)\mathcal{H}^{N-1}(\mathrm{d}x).
\end{aligned} \tag{28}$$

Finally, applying the Poincaré inequality for the Gaussian measure on X to the functions $f_i = x_i$ and $g_i = \sum_{j=1}^{i-1} x_j$, it follows that the right-hand size goes to zero as $N \uparrow \infty$.

Let us briefly compare Theorem 17 to the existing literature. In spirit the theorem is very close to the result of [10], with two main technical differences. The first one is that we impose quadratic growth of ψ at infinity (in a strong sense) instead of just superlinearity. The second one is the expression of macroscopic determinism: The condition in [10] can be rewritten as

$$\forall \varphi \in C^2(\mathbb{T}^1),\ \forall \varepsilon > 0, \quad \mu_{N,0}\left[\left\{x \in X_{N,0};\ \left|\int_{\mathbb{T}^1} (\zeta_0 - \bar{x})\varphi\right| \ge \varepsilon\right\}\right] \underset{N\to\infty}{\longrightarrow} 0, \tag{29}$$

whereas we use condition (25). Condition (25) is only slightly stronger than condition (29) – in fact, the conditions are equivalent as soon as

$$\lim_{R\to\infty} \limsup_{N\to\infty} N^{-1} \int_{|x|^2 \ge NR} |x|^2 \mu_{N,0}(\mathrm{d}x) = 0;$$

see [23], Theorem 7.12, for related results – and has the advantage that it is expressed in terms of a single numeric quantity, the mean square H^{-1} norm between the microscopic profile and the limit. Our final result (26) is expressed in terms of the same mean square norm. This is not just a different way to present the results: Our method yields an explicit estimate of how the departure from hydrodynamic behavior (expressed in the mean square H^{-1} norm) evolves in time.

This approach is quite different from the one in [10]; it is reminiscent of the one used by Yau [24], who estimates the departure from hydrodynamic behavior in terms of relative information (or entropy). But Yau's method leads one to impose a stronger assumption about the hydrodynamic behavior of the initial data, namely that it should behave like a local Gibbs state in the sense of relative entropy (see [24] for explanations).

Theorem 17 will be obtained in two steps. First, the abstract Theorem 8 will be used with the spaces $X_{N,0}$, $Y_{M,0}$ and the projection P from above (see also Section 5.1); in this way we shall see that the behavior of the system is well described by a macroscopic equation of gradient type

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = -\bar{A}\nabla_Y \bar{H}(\eta), \tag{30}$$

where $\eta \in Y_{M,0}$, and $\bar{A}$, $\bar{H}$ are as before. Then η is identified with a step function $\bar{\eta}$, so that (30) describes an evolution in $L^2(\mathbb{T}^1)$; in a separate step, it is shown that this evolution approaches the solution of (27) as $M \uparrow \infty$.

3. Proof of the abstract criterion for LSI

This section is devoted to the proof of Theorem 3. The key is to estimate the gradient of the macroscopic density $\bar{f}(y) = \int f(x)\mu(\mathrm{d}x|y)$ in terms of the full gradient, separating the contribution of the macroscopic part and the contribution of the fluctuations.

In the next statement, ∇^{macro} and ∇^{fluct} are defined as in (4).

Proposition 20. *Under the assumptions* (i) *and* (ii) *of Theorem* 3, *for any C^1 positive function f on X one has*, *for any $y \in Y$ and for any $\tau \in (0,1)$*,

$$\frac{1}{N}\frac{|\nabla_Y \bar{f}(y)|^2}{\bar{f}(y)} \le \frac{1}{1-\tau}\left(\frac{\kappa^2}{\rho^2}\right)\int \frac{|\nabla^{\mathrm{fluct}} f(x)|^2}{f(x)}\mu(\mathrm{d}x|y) + \frac{1}{\tau}\int \frac{|\nabla^{\mathrm{macro}} f(x)|^2}{f(x)}\mu(\mathrm{d}x|y). \tag{31}$$

Before proving Proposition 20 we shall see how it can be used to prove Theorem 3.

Proof of Theorem 3. First, the additive property of the entropy implies

$$\begin{aligned}\int \Phi(f)\,\mathrm{d}\mu - \Phi\left(\int f\,\mathrm{d}\mu\right) &= \int\left[\int \Phi\big(f(x)\big)\mu(\mathrm{d}x|y) - \Phi\left(\int f(x)\mu(\mathrm{d}x|y)\right)\right]\bar{\mu}(\mathrm{d}y)\\ &\quad + \left[\int \Phi\big(\bar{f}(y)\big)\bar{\mu}(\mathrm{d}y) - \Phi\left(\int \bar{f}(y)\bar{\mu}(\mathrm{d}y)\right)\right].\end{aligned} \tag{32}$$

By assumption (ii),

$$\begin{aligned}&\int\left[\int \Phi\big(f(x)\big)\mu(\mathrm{d}x|y) - \Phi\left(\int f(x)\mu(\mathrm{d}x|y)\right)\right]\bar{\mu}(\mathrm{d}y)\\ &\quad \le \frac{1}{\rho}\int\left[\int \frac{|(\mathrm{id}_X - NP^tP)\nabla f(x)|^2}{2f(x)}\mu(\mathrm{d}x|y)\right]\bar{\mu}(\mathrm{d}y).\end{aligned} \tag{33}$$

By assumption (iii), $\bar{\mu}$ satisfies LSI(λN), so that

$$\begin{aligned}&\int \Phi\big(\bar{f}(y)\big)\bar{\mu}(\mathrm{d}y) - \Phi\left(\int \bar{f}(y)\bar{\mu}(\mathrm{d}y)\right)\\ &\quad \le \frac{1}{\lambda N}\int \frac{|\nabla_Y \bar{f}(y)|^2}{2\bar{f}(y)}\bar{\mu}(\mathrm{d}y)\end{aligned}$$

$$\overset{(31)}{\le} \frac{1}{\lambda(1-\tau)}\left(\frac{\kappa^2}{\rho^2}\right)\int\left(\int \frac{|(\mathrm{id}_X - NP^tP)\nabla f(x)|^2}{2f(x)}\mu(\mathrm{d}x|y)\right)\bar{\mu}(\mathrm{d}y)$$
$$+\frac{1}{\lambda\tau}\int\left(\int \frac{|NP^tP\nabla f(x)|^2}{2f(x)}\mu(\mathrm{d}x|y)\right)\bar{\mu}(\mathrm{d}y). \tag{34}$$

(Assumptions (i) and (ii) were used here via Proposition 20.)

The combination of (32), (33) and (34), plus the identity (3), gives

$$\int \Phi(f)\,\mathrm{d}\mu - \Phi\left(\int f\,\mathrm{d}\mu\right)$$
$$\le \left[\frac{1}{\rho} + \frac{1}{\lambda(1-\tau)}\left(\frac{\kappa^2}{\rho^2}\right)\right]\int\left(\int \frac{|(\mathrm{id}_X - NP^tP)\nabla f(x)|^2}{2f}\mu(\mathrm{d}x|y)\right)\bar{\mu}(\mathrm{d}y)$$
$$+\frac{1}{\lambda\tau}\int \frac{|NP^tP\nabla f(x)|^2}{2f(x)}\mu(\mathrm{d}x).$$

It follows that μ satisfies LSI($\hat{\rho}$) with

$$\frac{1}{\hat{\rho}} = \max\left\{\frac{1}{\lambda\tau}, \frac{1}{\rho} + \frac{1}{\lambda(1-\tau)}\frac{\kappa^2}{\rho^2}\right\}.$$

Optimizing in τ gives the desired result. □

The proof of Proposition 20 is based on two lemmas involving covariance. We recall that the μ-covariance of two functions $f, g \in L^2(\mu)$ is the real number

$$\mathrm{cov}_\mu(f,g) = \int fg\,\mathrm{d}\mu - \left(\int f\,\mathrm{d}\mu\right)\left(\int g\,\mathrm{d}\mu\right). \tag{35}$$

This formula can be extended in an obvious way to vector-valued functions g (just apply (35) to each component).

The first lemma is a computation of the conditional expectation of $P\nabla f$ in terms of the macroscopic gradient and a covariance term.

Lemma 21. *For any $f \in \mathrm{Lip}(X)$ and any $y \in Y$,*

$$\int P\nabla f(x)\mu(\mathrm{d}x|y) = \frac{1}{N}\nabla_Y \bar{f}(y) + P\,\mathrm{cov}_{\mu(\mathrm{d}x|y)}(f, \nabla H). \tag{36}$$

The second lemma is a bound on the covariance using the LSI.

Lemma 22. *Let $\mu \in \mathcal{P}(X)$ satisfy* LSI(ρ) *for some $\rho > 0$. Then for any two Lipschitz functions $f : X \to \mathbb{R}_+$ and $g : X \to \mathbb{R}$,*

$$\left|\mathrm{cov}_\mu(f,g)\right| \le \|\nabla g\|_{L^\infty(\mu)}\sqrt{\frac{2}{\rho}\left(\int f\,\mathrm{d}\mu\right)\left(\int \Phi(f)\,\mathrm{d}\mu - \Phi\left(\int f\,\mathrm{d}\mu\right)\right)}$$
$$\le \frac{\|\nabla g\|_{L^\infty(\mu)}}{\rho}\sqrt{\left(\int f\,\mathrm{d}\mu\right)\left(\int \frac{|\nabla f|^2}{f}\,\mathrm{d}\mu\right)}.$$

Lemma 21 is a straightforward calculation. Lemma 22 goes back to Bodineau and Helffer [3]. Self-contained proofs can be found in Ledoux [16] or Otto and Westdickenberg [19]. Below we shall present a short alternative proof based on the results in [20]. (See Remark 23 for comments.)

Proof of Lemma 21. By definition

$$\begin{aligned}\bar{f}(y) &= \int f(x)\mu(\mathrm{d}x|y)\\ &= \frac{1}{\int_{\{Px=0\}} \exp(-H(NP^t y+z))\,\mathrm{d}z} \int_{\{Px=0\}} f\big(NP^t y+z\big)\exp\big(-H\big(NP^t y+z\big)\big)\,\mathrm{d}z.\end{aligned}$$

Thus, for any $\tilde{y} \in TY$,

$$\begin{aligned}\nabla_Y \bar{f}(y)\cdot\tilde{y} &= N\int \nabla f(x)\cdot P^t\tilde{y}\mu(\mathrm{d}x|y) - N\int f(x)\nabla H(x)\cdot P^t\tilde{y}\mu(\mathrm{d}x|y)\\ &\quad - N\int f(x)\mu(\mathrm{d}x|y)\cdot\int\big(-\nabla H(x)\cdot P^t\tilde{y}\big)\mu(\mathrm{d}x|y)\\ &= N\bigg[\int P\nabla f(x)\mu(\mathrm{d}x|y) - \int f(x)P\nabla H(x)\mu(\mathrm{d}x|y)\\ &\quad + \int f(x)\mu(\mathrm{d}x|y)\int P\nabla H(x)\mu(\mathrm{d}x|y)\bigg]\cdot\tilde{y},\end{aligned}$$

which proves the result. □

Proof of Lemma 22. Without loss of generality we may assume $\int f\,\mathrm{d}\mu = 1$. Let W_p stand for the Monge–Kantorovich–Wasserstein distance of order p (see [23]): $W_p(\mu,\nu) = \inf(\mathbb{E}|U-V|^p)^{1/p}$, where U and V are random variables with respective law μ and ν. Then

$$\begin{aligned}\big|\mathrm{cov}_\mu(f,g)\big| &= \left|\int gf\,\mathrm{d}\mu - \int g\,\mathrm{d}\mu\right|\\ &\le \|\nabla g\|_{L^\infty} \sup_{\|\nabla\varphi\|_{L^\infty}\le 1}\left|\int \varphi f\,\mathrm{d}\mu - \int\varphi\,\mathrm{d}\mu\right|\\ &\le \|\nabla g\|_{L^\infty} W_1(f\mu,\mu) \qquad (37)\\ &\le \|\nabla g\|_{L^\infty} W_2(f\mu,\mu) \qquad (38)\\ &\le \|\nabla g\|_{L^\infty}\sqrt{\frac{2}{\rho}\int \Phi(f)\,\mathrm{d}\mu}, \qquad (39)\end{aligned}$$

where (37) comes from the Kantorovich–Rubinstein duality [23], Theorem 1.3, (38) from the Hölder inequality and (39) from [20], Theorem 1. □

Remark 23. *The proof of Lemma 22 in* [19], *Lemma* 1, *is longer but more elementary, and mimics the proof of the main result of* [20]. *The proof given above is shorter*: *It uses the main result of* [20] *directly.*

We can finally prove Proposition 20 and thus conclude the proof of Theorem 3.

Proof of Proposition 20. First note that by (2), we have that for any $x \in X$,

$$\big|NP^tPx\big|^2 = N|Px|^2. \qquad (40)$$

Then by using successively Lemma 21, the Young inequality $\langle a, b\rangle \leq (1-\tau)^{-1}|a|^2 + \tau^{-1}|b|^2$, (40), and Jensen's inequality (with the convex function $(a,b) \longmapsto |b|^2/a$), we discover that

$$\begin{aligned}
\frac{|\nabla_Y \bar f(y)|_Y^2}{\bar f(y)} &= \frac{N^2}{\bar f(y)}\left| -P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H) + P\int \nabla f(x)\mu(\mathrm{d}x|y)\right|_Y^2 \\
&\leq \frac{N^2}{1-\tau}\frac{1}{\bar f(y)}\left|P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|_Y^2 + \frac{N^2}{\tau}\frac{1}{\bar f(y)}\left|P\int \nabla f(x)\mu(\mathrm{d}x|y)\right|_Y^2 \\
&\overset{(40)}{=} \frac{N}{1-\tau}\frac{1}{\bar f(y)}\left|NP^tP\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|^2 \\
&\quad + \frac{N}{\tau}\frac{1}{\bar f(y)}\left|NP^tP\int \nabla f(x)\mu(\mathrm{d}x|y)\right|^2 \\
&\leq \frac{N}{1-\tau}\frac{1}{\bar f(y)}\left|NP^tP\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|^2 \\
&\quad + \frac{N}{\tau}\int \frac{|NP^tP\nabla f(x)|^2}{f(x)}\mu(\mathrm{d}x|y).
\end{aligned} \tag{41}$$

It remains only to estimate the first term on the right-hand side. We use assumption (ii) and Lemma 22:

$$\begin{aligned}
&\left|NP^tP\operatorname{cov}_{\mu(\mathrm{d}x|y)}\big(f,\nabla H(x)\big)\right|^2 \\
&\quad = \sup_{|\tilde x|\leq 1}\left[\operatorname{cov}_{\mu(\mathrm{d}x|y)}\big(f, NP^tP\nabla H(x)\cdot\tilde x\big)\right]^2 \\
&\quad = \sup_{|\tilde x|\leq 1}\left[\operatorname{cov}_{\mu(\mathrm{d}x|y)}\big(f, \nabla H(x)\cdot\big(NP^tP\tilde x\big)\big)\right]^2 \\
&\quad \leq \frac{1}{\rho^2}\left(\sup_{|\tilde x|\leq 1}\sup_x\left|\big(\mathrm{id}_X - NP^tP\big)\nabla\big(\nabla H(x)\cdot\big(NP^tP\tilde x\big)\big)\right|^2\right)\left(\int f(x)\mu(\mathrm{d}x|y)\right) \\
&\qquad \times\left(\int \frac{|(\mathrm{id}_X - NP^tP)\nabla f(x)|^2}{f(x)}\mu(\mathrm{d}x|y)\right).
\end{aligned} \tag{42}$$

Finally, we observe that

$$\begin{aligned}
&\sup_{|\tilde x|\leq 1}\sup_x\left|\big(\mathrm{id}_X - NP^tP\big)\nabla\big(\nabla H(x)\cdot\big(NP^tP\tilde x\big)\big)\right| \\
&\quad = \sup_{|\tilde x|\leq 1, |\tilde z|\leq 1}\sup_x\left\langle\big(\mathrm{id}_X - NP^tP\big)\nabla\big(\nabla H(x)\cdot\big(NP^tP\tilde x\big)\big), \tilde z\right\rangle \\
&\quad = \sup_{|\tilde x|\leq 1, |\tilde z|\leq 1}\sup_x\left\langle\nabla\big(\nabla H(x)\cdot\big(NP^tP\tilde x\big)\big), \big(\mathrm{id}_X - NP^tP\big)\tilde z\right\rangle \\
&\quad = \sup\Big\{\big\langle\nabla\big(\nabla H(x)\cdot u\big), v\big\rangle;\ |u|, |v|\leq 1,\ u\in\operatorname{Ran}\big(NP^tP\big), \\
&\qquad\qquad v\in\operatorname{Ran}\big(\mathrm{id}_X - NP^tP\big)\Big\} \\
&\quad \overset{(5)}{=} \kappa.
\end{aligned} \tag{43}$$

Substituting (43) into (42) and inserting the result into (41) gives (31). ☐

4. Proof of the abstract result for the hydrodynamic limit

In this section we prove Theorem 8. We begin by recording some a priori entropy estimates for the microscopic and macroscopic systems and a bound on the second moments.

Proposition 24 (Entropy and moment estimates). *If $f(t,x)$ and $\eta(t)$ satisfy the assumptions of Theorem 8, then for any $T<\infty$ we have*

$$\int \Phi\bigl(f(T,x)\bigr)\mu(\mathrm{d}x)+\int_0^T\left(\int\frac{\nabla f\cdot A\nabla f}{f}(t,x)\mu(\mathrm{d}x)\right)\mathrm{d}t=\int\Phi\bigl(f(0,x)\bigr)\mu(\mathrm{d}x); \tag{44}$$

$$\bar{H}\bigl(\eta(T)\bigr)+\int_0^T\left\langle\frac{\mathrm{d}\eta}{\mathrm{d}t},\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\right\rangle_Y\mathrm{d}t=\bar{H}\bigl(\eta(0)\bigr); \tag{45}$$

$$\left(\int|x|^2f(t,x)\mu(\mathrm{d}x)\right)^{1/2}\le\left(\frac{2}{\hat{\rho}}\int\Phi\bigl(f(0,x)\bigr)\mu(\mathrm{d}x)\right)^{1/2}+\left(\int|x|^2\mu(\mathrm{d}x)\right)^{1/2}. \tag{46}$$

Remark 25. *From (44) together with the positivity of the relative entropy and the positive definiteness of the matrix A, it follows in particular that*

$$\max\left\{\sup_{0\le t\le T}\left(\int\Phi\bigl(f(t,x)\bigr)\mu(\mathrm{d}x)\right),\int_0^T\left(\int\frac{\nabla f\cdot A\nabla f}{f}(t,x)\mu(\mathrm{d}x)\right)\mathrm{d}t\right\}$$
$$\le\int\Phi\bigl(f(0,x)\bigr)\mu(\mathrm{d}x). \tag{47}$$

The proof of the proposition is (for the most part) standard and deferred to the end of the section. With its help, we shall now prove Theorem 8.

Proof of Theorem 8. The proof of Theorem 8 relies on estimating the time-derivative of $\Theta(t)$ to get a Gronwall-type estimate in the end. It is decomposed into three steps.

Step 1. *Computation of* $(\mathrm{d}/\mathrm{d}t)\Theta(t)$. In this step we establish the exact formula

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int\frac{1}{2N}\bigl(x-NP^t\eta\bigr)\cdot A^{-1}\bigl(x-NP^t\eta\bigr)f(t,x)\mu(\mathrm{d}x)\\
&\quad=-\int_Y\bigl(\nabla_Y\bar{H}(y)-\nabla_Y\bar{H}(\eta)\bigr)\cdot(y-\eta)\bar{f}(t,y)\bar{\mu}(\mathrm{d}y)+\frac{M}{N}\\
&\qquad-\int_Y(y-\eta)\cdot P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\bar{\mu}(\mathrm{d}y)\\
&\qquad-\int\frac{1}{N}\bigl(\mathrm{id}_X-NP^tP\bigr)x\cdot\nabla f(t,x)\mu(\mathrm{d}x)\\
&\qquad-\int\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot PA^{-1}\bigl(\mathrm{id}_X-NP^tP\bigr)xf(t,x)\mu(\mathrm{d}x).
\end{aligned}\tag{48}$$

To prove (48), we use the definition of the stochastic evolution, the coarse-grained deterministic evolution, and the splitting $x=NP^tPx+(\mathrm{id}_X-NP^tP)x$ to obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int\frac{1}{2N}\bigl(x-NP^t\eta\bigr)\cdot A^{-1}\bigl(x-NP^t\eta\bigr)f\mu(\mathrm{d}x)\\
&\quad\overset{(8)}{=}-\int\frac{1}{N}A^{-1}\bigl(x-NP^t\eta\bigr)\cdot A\nabla f\mu(\mathrm{d}x)-\int P^t\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot A^{-1}\bigl(x-NP^t\eta\bigr)f\mu(\mathrm{d}x)
\end{aligned}$$

$$
\begin{aligned}
&= -\int P^t(Px-\eta)\cdot\nabla f\,\mu(\mathrm{d}x) - \int \frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot PA^{-1}NP^t(Px-\eta)f\,\mu(\mathrm{d}x)\\
&\quad -\int \frac{1}{N}\big(\mathrm{id}_X - NP^tP\big)x\cdot\nabla f\,\mu(\mathrm{d}x) - \int\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot PA^{-1}\big(\mathrm{id}_X - NP^tP\big)xf\,\mu(\mathrm{d}x)\\
&\overset{(9)}{=} -\int (Px-\eta)\cdot P\nabla f\,\mu(\mathrm{d}x) - \int \bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot(Px-\eta)f\,\mu(\mathrm{d}x)\\
&\quad -\int \frac{1}{N}\big(\mathrm{id}_X - NP^tP\big)x\cdot\nabla f\,\mu(\mathrm{d}x) - \int\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot PA^{-1}\big(\mathrm{id}_X - NP^tP\big)xf\,\mu(\mathrm{d}x)\\
&\overset{(3),(10)}{=} -\int (y-\eta)\cdot P\int\nabla f\,\mu(\mathrm{d}x|y)\bar{\mu}(\mathrm{d}y) + \int\nabla_Y\bar{H}(\eta)\cdot(y-\eta)\bar{f}\bar{\mu}(\mathrm{d}y)\\
&\quad -\int \frac{1}{N}\big(\mathrm{id}_X - NP^tP\big)x\cdot\nabla f\,\mu(\mathrm{d}x) - \int\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot PA^{-1}\big(\mathrm{id}_X - NP^tP\big)xf\,\mu(\mathrm{d}x).
\end{aligned}
$$

We keep the last three terms unchanged, and transform the first term according to Lemma 21:

$$
\begin{aligned}
&-\int (y-\eta)\cdot P\int\nabla f\,\mu(\mathrm{d}x|y)\bar{\mu}(\mathrm{d}y)\\
&\quad\overset{(36)}{=} -\frac{1}{N}\int (y-\eta)\cdot\nabla_Y\bar{f}\bar{\mu}(\mathrm{d}y) - \int (y-\eta)\cdot P\,\mathrm{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\bar{\mu}(\mathrm{d}y).
\end{aligned}
$$

It again remains to consider the first term on the right-hand side. Using integration by parts, we obtain

$$
\begin{aligned}
-\frac{1}{N}\int (y-\eta)\cdot\nabla_Y\bar{f}\bar{\mu}(\mathrm{d}y) &\overset{(6)}{=} \frac{1}{N}\int\nabla_Y\cdot y\bar{f}\bar{\mu}(\mathrm{d}y) - \int (y-\eta)\cdot\nabla_Y\bar{H}(y)\bar{f}\bar{\mu}(\mathrm{d}y)\\
&= \frac{(\dim Y)}{N} - \int (y-\eta)\cdot\nabla_Y\bar{H}(y)\bar{f}\bar{\mu}(\mathrm{d}y).
\end{aligned}
$$

This concludes Step 1.

Step 2. *An upper bound.* In this step we establish the following upper bound:

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int\frac{1}{2N}\big(x-NP^t\eta\big)\cdot A^{-1}\big(x-NP^t\eta\big)f(t,x)\mu(\mathrm{d}x) + \frac{\lambda}{2}\int_Y |y-\eta|_Y^2\bar{f}(t,y)\bar{\mu}(\mathrm{d}y)\\
&\quad\le \frac{M}{N} + \frac{\gamma\kappa^2}{2\lambda\rho^2}\frac{1}{M^2}\int\frac{1}{Nf(t,x)}\nabla f\cdot A\nabla f(t,x)\mu(\mathrm{d}x)\\
&\qquad + \gamma^{1/2}\frac{1}{M}\bigg(\int\frac{1}{N}|x|^2 f(t,x)\mu(\mathrm{d}x)\bigg)^{1/2}\\
&\qquad \times\bigg(\bigg(\int\frac{1}{Nf(t,x)}\nabla f\cdot A\nabla f(t,x)\mu(\mathrm{d}x)\bigg)^{1/2}\\
&\qquad + \bigg(\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\bigg)^{1/2}\bigg).
\end{aligned}
\tag{49}
$$

To establish (49), let us come back to the expression on the right-hand side of (48) and bound it term by term. We use assumption (iii′) to bound the first term:

$$
-\int\big(\nabla_Y\bar{H}(y) - \nabla_Y\bar{H}(\eta)\big)\cdot(y-\eta)\bar{f}\bar{\mu}(\mathrm{d}y) \le -\lambda\int |y-\eta|_Y^2\bar{f}\bar{\mu}(\mathrm{d}y).
\tag{50}
$$

The third term in (48) is controlled by Cauchy–Schwarz and (40):

$$\begin{aligned}
&\left|\int (y-\eta)\cdot P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\bar\mu(\mathrm{d}y)\right| \\
&\quad\le \left(\int |y-\eta|_Y^2\bar f\bar\mu(\mathrm{d}y)\cdot\int\frac{1}{\bar f}\left|P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|_Y^2\bar\mu(\mathrm{d}y)\right)^{1/2} \\
&\quad\overset{(40)}{=} \left(\int |y-\eta|_Y^2\bar f\bar\mu(\mathrm{d}y)\cdot\int\frac{1}{N\bar f}\left|NP^tP\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|^2\bar\mu(\mathrm{d}y)\right)^{1/2}.
\end{aligned}\tag{51}$$

Recalling assumption (i) together with (42) and (43) from the proof of Proposition 20, we have

$$\begin{aligned}
\left|NP^tP\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\right|^2 &\le \frac{\kappa^2}{\rho^2}\bar f\int\frac{1}{f}\left|\left(\mathrm{id}_X-NP^tP\right)\nabla f\right|^2\mu(\mathrm{d}x|y) \\
&\le \gamma\frac{\kappa^2}{\rho^2}\frac{1}{M^2}\bar f\int\frac{1}{f}\nabla f\cdot A\nabla f\mu(\mathrm{d}x|y),
\end{aligned}\tag{52}$$

where we have recalled assumption (vi). Substituting (52) into (51) and using Young's inequality, we find

$$\begin{aligned}
&\left|\int (y-\eta)\cdot P\operatorname{cov}_{\mu(\mathrm{d}x|y)}(f,\nabla H)\bar\mu(\mathrm{d}y)\right| \\
&\quad\le \left(\gamma\frac{\kappa^2}{\rho^2}\frac{1}{M^2}\int |y-\eta|_Y^2\bar f\bar\mu(\mathrm{d}y)\cdot\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\right)^{1/2} \\
&\quad\le \frac{\gamma}{2\lambda}\frac{\kappa^2}{\rho^2}\frac{1}{M^2}\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)+\frac{\lambda}{2}\int |y-\eta|_Y^2\bar f\bar\mu(\mathrm{d}y).
\end{aligned}\tag{53}$$

It remains to take care of the last two terms in (48). We begin with the following estimate:

$$\left(\mathrm{id}_X-NP^tP\right)x\cdot A^{-1}\left(\mathrm{id}_X-NP^tP\right)x\le\gamma\frac{1}{M^2}|x|^2.\tag{54}$$

Indeed,

$$\begin{aligned}
&\left(\mathrm{id}_X-NP^tP\right)x\cdot A^{-1}\left(\mathrm{id}_X-NP^tP\right)x \\
&\quad= x\cdot\left(\mathrm{id}_X-NP^tP\right)A^{-1}\left(\mathrm{id}_X-NP^tP\right)x \\
&\quad\le |x|\left|\left(\mathrm{id}_X-NP^tP\right)A^{-1}\left(\mathrm{id}_X-NP^tP\right)x\right| \\
&\quad\le |x|\left(\gamma\frac{1}{M^2}A^{-1}\left(\mathrm{id}_X-NP^tP\right)x\cdot\left(\mathrm{id}_x-NP^tP\right)x\right)^{1/2},
\end{aligned}$$

where assumption (vi) was used in the last step.

Now we can estimate the fourth term in (48) by means of Cauchy–Schwarz and (54):

$$\begin{aligned}
&\left|\int\frac{1}{N}\left(\mathrm{id}_X-NP^tP\right)x\cdot\nabla f\mu(\mathrm{d}x)\right| \\
&\quad\le\left(\int\frac{1}{N}\left(\mathrm{id}_X-NP^tP\right)x\cdot A^{-1}\left(\mathrm{id}_X-NP^tP\right)xf\mu(\mathrm{d}x)\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\right)^{1/2} \\
&\quad\le\left(\gamma\frac{1}{M^2}\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\cdot\int\frac{1}{N}|x|^2f\mu(\mathrm{d}x)\right)^{1/2}.
\end{aligned}\tag{55}$$

Similarly for the fifth term in (48), we have

$$\begin{aligned}
&\left|\int \frac{\mathrm{d}\eta}{\mathrm{d}t} \cdot PA^{-1}\big(\mathrm{id}_X - NP^tP\big)xf\mu(\mathrm{d}x)\right| \\
&\quad\le \left(\int P^t\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot A^{-1}NP^t\frac{\mathrm{d}\eta}{\mathrm{d}t}f\mu(\mathrm{d}x)\right)^{1/2}\left(\int \frac{1}{N}\big(\mathrm{id}_X - NP^tP\big)x\cdot A^{-1}\big(\mathrm{id}_X - NP^tP\big)xf\mu(\mathrm{d}x)\right)^{1/2} \\
&\quad\overset{(9),(54)}{\le} \left(\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\right)^{1/2}\left(\int\frac{\gamma}{NM^2}|x|^2f\mu(\mathrm{d}x)\right)^{1/2}. \qquad (56)
\end{aligned}$$

Substituting (50), (53), (55) and (56) into (48) gives (49).

Step 3. *Time-integration and conclusion.* Recalling the definition (11) and integrating (49) with respect to time, we observe that for any $T > 0$, we have

$$\begin{aligned}
&\max\left\{\sup_{t\in(0,T)}\Theta(t), \frac{\lambda}{2}\int_0^T\int_Y |y-\eta|_Y^2\bar{f}(t,y)\bar{\mu}(\mathrm{d}y)\,\mathrm{d}t\right\} \\
&\quad\le \Theta(0) + T\frac{M}{N} + \frac{\gamma\kappa^2}{2\lambda\rho^2}\frac{1}{M^2}\int_0^T\int\frac{1}{Nf(t,x)}\nabla f\cdot A\nabla f(t,x)\mu(\mathrm{d}x)\,\mathrm{d}t \\
&\qquad + \gamma^{1/2}\frac{1}{M}\int_0^T\left(\int\frac{1}{N}|x|^2f(t,x)\mu(\mathrm{d}x)\right)^{1/2}\left(\left(\int\frac{1}{Nf(t,x)}\nabla f\cdot A\nabla f(t,x)\mu(\mathrm{d}x)\right)^{1/2}\right. \\
&\qquad \left. + \left(\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\right)^{1/2}\right)\mathrm{d}t. \qquad (57)
\end{aligned}$$

The first and second terms on the right-hand side are in final form. The third term is the entropy production, and can be controlled by means of Proposition 24 (inequality (47)). Finally, the last time-integral can be estimated as follows:

$$\begin{aligned}
&\int_0^T\left(\int\frac{1}{N}|x|^2f\mu(\mathrm{d}x)\right)^{1/2}\left(\left(\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\right)^{1/2} + \left(\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\right)^{1/2}\right)\mathrm{d}t \\
&\quad\le\left(\int_0^T\int\frac{1}{N}|x|^2f\mu(\mathrm{d}x)\,\mathrm{d}t\right)^{1/2}\left(\left(\int_0^T\int\frac{1}{Nf}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\,\mathrm{d}t\right)^{1/2} + \left(\int_0^T\frac{\mathrm{d}\eta}{\mathrm{d}t}\cdot\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\,\mathrm{d}t\right)^{1/2}\right).
\end{aligned}$$

After applying Proposition 24 (together with Young's inequality), the right-hand side improves to

$$\begin{aligned}
&\left(2\int_0^T\left(\frac{1}{N}\int|x|^2\mu(\mathrm{d}x) + \frac{2}{N\hat{\rho}}\int\Phi\big(f(0,x)\big)\mu(\mathrm{d}x)\right)\mathrm{d}t\right)^{1/2} \\
&\quad\times\left(\left(\frac{1}{N}\int\Phi\big(f(0,x)\big)\mu(\mathrm{d}x)\right)^{1/2} + \big(\bar{H}\big(\eta(0)\big) - \bar{H}\big(\eta(T)\big)\big)^{1/2}\right).
\end{aligned}$$

Substituting the constants from the assumptions concludes the proof of Theorem 8. □

Proof of Proposition 24. The first two bounds are standard; we shall not be careful about regularity issues. Multiply (8) by $\xi = \log f$ and integrate by parts to get

$$\int\Phi\big(f(T,x)\big)\mu(\mathrm{d}x) - \int\Phi\big(f(0,x)\big)\mu(\mathrm{d}x) = -\int_0^T\int\frac{1}{f}\nabla f\cdot A\nabla f\mu(\mathrm{d}x)\,\mathrm{d}t.$$

Next, to verify (45) notice that

$$\int_0^T\left\langle\frac{\mathrm{d}\eta}{\mathrm{d}t},\bar{A}^{-1}\frac{\mathrm{d}\eta}{\mathrm{d}t}\right\rangle_Y\mathrm{d}t \overset{(10)}{=} -\int_0^T\left\langle\frac{\mathrm{d}\eta}{\mathrm{d}t},\nabla_Y\bar{H}(\eta)\right\rangle_Y\mathrm{d}t = \bar{H}(\eta_0) - \bar{H}\big(\eta(T)\big).$$

It remains to establish (46). According to Remark 10, μ satisfies LSI($\hat{\rho}$). By Lemma 26 below,

$$\left(\int |x|^2 f(t,x)\mu(\mathrm{d}x)\right)^{1/2} \le \left(\int |x|^2\mu(\mathrm{d}x)\right)^{1/2} + \left(\frac{2}{\hat{\rho}}\int \Phi\big(f(t,x)\big)\mu(\mathrm{d}x)\right)^{1/2},$$

and the conclusion follows since $\int \Phi(f(t,x))\,\mathrm{d}\mu(x)$ is nonincreasing in t. □

The next lemma was used in the proof of Proposition 24:

Lemma 26. *Let $\rho > 0$ and let $\mu \in \mathcal{P}(\mathbb{R}^N)$ satisfy* LSI(ρ). *Then for any probability density f on $(\mathbb{R}^N, \mu)$,*

$$\left(\int |x|^2 f(x)\mu(\mathrm{d}x)\right)^{1/2} \le \left(\int |x|^2\mu(\mathrm{d}x)\right)^{1/2} + \left(\frac{2}{\rho}\int \Phi\big(f(x)\big)\mu(\mathrm{d}x)\right)^{1/2}.$$

Remark 27. *This particular estimate seems to be new. We shall prove it by mimicking the heat semigroup argument in* [20]. *It is also possible to prove it more directly as a consequence of the results in* [20] (*and to generalize it to more general length structures than just the Euclidean space*).

Proof of Lemma 26. Let $P_t f$ be the diffusion semigroup defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int P_t f\xi\mu(\mathrm{d}x) = -\int \nabla P_t f\cdot\nabla\xi\mu(\mathrm{d}x) \quad \forall\xi(x)\ge 0, \tag{58}$$

$$P_0 f = f. \tag{59}$$

We observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \Phi(P_t f)\mu(\mathrm{d}x) \overset{(58)}{=} -\int \frac{1}{P_t f}|\nabla P_t f|^2\mu(\mathrm{d}x) \overset{\mathrm{LSI}(\rho)}{\le} -2\rho\int \Phi(P_t f)\mu(\mathrm{d}x),$$

so that

$$\int \Phi(P_t f)\mu(\mathrm{d}x) \le \exp(-2\rho t)\int \Phi(P_0 f)\mu(\mathrm{d}x),$$

and in particular,

$$\lim_{t\uparrow\infty} P_t f = 1. \tag{60}$$

We now have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\int |x|^2 P_t f\mu(\mathrm{d}x)\right)^{1/2} - \left(\frac{2}{\rho}\int \Phi(P_t f)\mu(\mathrm{d}x)\right)^{1/2}\right)\\
&\quad\overset{(58)}{=} -\left(\int |x|^2 P_t f\mu(\mathrm{d}x)\right)^{-1/2}\int x\cdot\nabla P_t f\mu(\mathrm{d}x)\\
&\qquad + \left(2\rho\int \Phi(P_t f)\mu(\mathrm{d}x)\right)^{-1/2}\int \frac{1}{P_t f}|\nabla P_t f|^2\mu(\mathrm{d}x)\\
&\quad\overset{\mathrm{LSI}(\rho)}{\ge} -\left(\int |x|^2 P_t f\mu(\mathrm{d}x)\right)^{-1/2}\left(\int |x|^2 P_t f\mu(\mathrm{d}x)\right)^{1/2}\left(\int \frac{1}{P_t f}|\nabla P_t f|^2\mu(\mathrm{d}x)\right)^{1/2}\\
&\qquad + \left(\int \frac{1}{P_t f}|\nabla P_t f|^2\mu(\mathrm{d}x)\right)^{-1/2}\int \frac{1}{P_t f}|\nabla P_t f|^2\mu(\mathrm{d}x)
\end{aligned}$$

$$= -\left(\int \frac{1}{P_t f}|\nabla P_t f|^2 \mu(\mathrm{d}x)\right)^{1/2} + \left(\int \frac{1}{P_t f}|\nabla P_t f|^2 \mu(\mathrm{d}x)\right)^{1/2}$$
$$= 0.$$

Integrating this inequality and recalling (59) and (60) leads to:

$$\left(\int |x|^2 f \mu(\mathrm{d}x)\right)^{1/2} \le \left(\int |x|^2 \mu(\mathrm{d}x)\right)^{1/2} + \left(\frac{2}{\rho}\int \Phi(f)\mu(\mathrm{d}x)\right)^{1/2}. \qquad \square$$

5. Application: LSI for the canonical ensemble

In this section we shall prove Theorem 14 (modulo a key technical proposition deferred to the Appendix). We begin by introducing the set-up and notation which will be used throughout this section and also in Section 6.

5.1. *Set-up and preparations*

In keeping with the abstract framework of Theorem 3, we set

$$X = X_{N,m} = \left\{(x_1, \dots, x_N) \in \mathbb{R}^N;\ \frac{1}{N}\sum_{i=1}^{N} x_i = m\right\},$$

equipped with the ℓ^2 inner product

$$\langle x, \tilde{x}\rangle_X = \sum_{i=1}^{N} x_i \tilde{x}_i. \tag{61}$$

We divide the N spins into M blocks and define the macroscopic variables to be the mean of each block. To fix the ideas, let us first assume that all blocks have the same size K, so that $N = MK$ (see Remark 30 below). Then the macroscopic variables form a set of M numbers that still have mean m. This motivates the choice of the macroscopic space as

$$Y = Y_{M,m} = \left\{(y_1, \dots, y_M);\ \frac{1}{M}\sum_{j=1}^{M} y_j = m\right\},$$

equipped with the L^2 inner product

$$\langle y, \tilde{y}\rangle_Y = \frac{1}{M}\sum_{j=1}^{M} y_j \tilde{y}_j. \tag{62}$$

(Recall that the notion of gradient and Hessian depends on the inner product; cf. (67) and (68) below.) The projection $P = P_{N,K} : X_{N,m} \to Y_{M,m}$ is defined by

$$P_{N,K}(x_1, \dots, x_N) = (y_1, \dots, y_M); \qquad y_j = \frac{1}{K}\sum_{i=(j-1)K+1}^{jK} x_i. \tag{63}$$

It is straightforward to check that $P_{N,K}$ satisfies (2).

For each $y \in Y_{M,m}$ we define

$$X_{N,K,y} = \{x \in X_{N,m};\ P_{N,K}x = y\},$$

the pre-image of y under the projection. This is the set of all microscopic profiles compatible with the macroscopic profile y. This decomposition induces the factorization

$$X_{N,K,y} = \bigotimes_{j=1}^{M} X_{K,y_j}$$

and $P_{N,K}$ factorizes on the fibers $X_{N,K,y}$.

If $\mu_{N,m}$ is given by (17), we can explicitly compute the logarithmic densities H and $\bar{H}$:

$$H(x) = \sum_{i=1}^{N} \psi(x_i) + \log Z; \tag{64}$$

$$\begin{aligned}\bar{H}(y) &= -\frac{1}{N}\left(\log \int_{X_{N,K,y}} \exp\bigl(-H(x)\bigr)\mathcal{H}^{N-M}(\mathrm{d}x) - \log \bar{Z}\right)\\ &= -\frac{1}{N}\sum_{j=1}^{M} \log\left(\int_{X_{K,y_j}} \exp\bigl(-H(x)\bigr)\mathcal{H}^{K-1}(\mathrm{d}x)\right) + \frac{1}{N}\log \bar{Z},\end{aligned}$$

where Z and $\bar{Z}$ are normalizing constants. If we further define

$$\psi_K(m) := -\frac{1}{K}\log\left(\int_{X_{K,m}} \exp\left(-\sum_{i=1}^{K}\psi(x_i)\right)\mathcal{H}^{K-1}(\mathrm{d}x)\right), \tag{65}$$

then we can express $\bar{H}$ as

$$\bar{H}(y) = \frac{1}{M}\sum_{i=1}^{M}\psi_K(y_i) + \frac{1}{N}\log \bar{Z}. \tag{66}$$

For future reference, we remark that according to (62) and (66), the gradient of $\bar{H}$, defined via

$$\mathrm{d}\bar{H}(y) = \bigl\langle \nabla_Y \bar{H}(y), \tilde{y}\bigr\rangle_Y \quad \text{for all } \tilde{y} \in Y$$

is given by

$$\bigl(\nabla_Y \bar{H}(y)\bigr)_i = \psi_K'(y_i). \tag{67}$$

Similarly, the Hessian is identified as

$$(\mathrm{Hess}_Y \bar{H})_{ij} = \psi_K''(y_i)\delta_{ij}. \tag{68}$$

As a final preparation, let us compute the conditional measure $\mu(\mathrm{d}x|y)$. For each y, the fiber $P_{N,K}^{-1}(y)$ is $(N-M)$-dimensional, and

$$\begin{aligned}\frac{\mathrm{d}\mu(\cdot|y)}{\mathrm{d}\mathcal{H}^{N-M}}(x) &= \frac{\exp(-H(x))}{\exp(-N\bar{H}(y))} = \exp\left(-\sum_{i=1}^{N}\psi(x_i) + K\sum_{j=1}^{M}\psi_K(y_j)\right)\\ &= \prod_{j=1}^{M}\exp\left(K\psi_K(y_j) - \sum_{i=(j-1)K+1}^{jK}\psi(x_i)\right).\end{aligned}$$

Hence $\mu(\mathrm{d}x|y)$ tensorizes on $\bigotimes_{j=1}^{M} X_{K,y_j}$: more explicitly,

$$\mu(\mathrm{d}x|y) = \bigotimes_{j=1}^{M}\left(\frac{\exp(-\sum_{i=(j-1)K+1}^{jK}\psi(x_i))}{Z_j}\mathcal{H}^{K-1}(\mathrm{d}x)\right) =: \bigotimes_{j=1}^{M}\mu_{N,K,j,y_j}(\mathrm{d}x), \tag{69}$$

where each μ_{N,K,j,y_j} is a probability measure on X_{K,y_j}.

5.2. *Proof of Theorem* 14

Theorem 14 will follow from Theorem 3 once we verify the assumptions of the latter. We will need two ingredients: The first ingredient gives a microscopic LSI for any finite K.

Lemma 28 (Microscopic LSI). *Consider the measure $\mu_{N,m}$ given by* (17) *with ψ satisfying* (16), *that is*:

$$\psi(x) = \frac{1}{2}x^2 + \delta\psi(x), \qquad \|\delta\psi\|_{C^2(\mathbb{R})} < +\infty.$$

Then for any $K < \infty$, for all $y \in Y_{M,m}$ the probability measures $\mu(\mathrm{d}x|y)$ on $X_{N,K,y}$ satisfy LSI with constant $\exp(-K \operatorname{osc}_{\mathbb{R}} \delta\psi)$.

The second ingredient will imply a macroscopic LSI for K sufficiently large. (It is a corollary of Proposition 31 below.)

Lemma 29 (Convexity of coarse-grained Hamiltonian). *Consider the measure $\mu_{N,m}$ given by* (17) *with ψ satisfying* (16). *There exist $K_0 < \infty$ and $\lambda > 0$ dependent only on ψ such that for any $K \ge K_0$ we have*

$$\langle \tilde{y}, \operatorname{Hess}\bar{H}(y)\tilde{y}\rangle_Y \ge \lambda\langle \tilde{y}, \tilde{y}\rangle_Y. \tag{70}$$

Proof of Theorem 14. We need only verify the assumptions of Theorem 3. Assumption (i) is an obvious consequence of (16). Assumption (ii) is given by Lemma 28 for any finite K. Finally, assumption (iii) for K sufficiently large follows from the combination of Lemma 29, the definition (6) of $\bar{H}$, and the Bakry–Émery theorem (cf. Criterion 3). □

Remark 30. *It remains only to discuss what happens when N/K is not an integer. This is easily handled by generalizing the construction, allowing for the M blocks to have different sizes*:

$$Px = (y_1, \ldots, y_M); \qquad y_j = \frac{1}{K_j}\sum_{i=K_1+\cdots+K_{j-1}+1}^{K_1+\cdots+K_j} x_i; \qquad \sum_{j=1}^{M} K_j = N.$$

Then the space Y is defined by $(1/M)\sum \alpha_j y_j = m$, where $\alpha_j = MK_j/N$.

With this generalized setting the whole proof goes through provided that the K_j are all large enough but uniformly bounded, and, say, $1 \le \alpha_j \le 2$. This is always feasible since any number $N \ge K$ can be decomposed into $N = K_1 + \cdots + K_M$ where each K_j satisfies $K \le K_j \le 2K$ and $1 \le MK_j/N \le 2$.

5.3. *Microscopic LSI*

The microscopic LSI follows easily from the factorization (69):

Proof of Lemma 28. According to (69) and the tensorization principle (cf. Criterion 1), it suffices to show that for every j,

$$\mu_{N,K,j,y_j} \text{ satisfies LSI with constant } \exp(-K \operatorname{osc}_{\mathbb{R}} \delta\psi). \tag{71}$$

According to (69) and (16), we have

$$\mu_{N,K,j,y_j}(\mathrm{d}x) = \frac{1}{Z_j} \exp\left[-\left(\sum_{i=(j-1)K+1}^{jK} \frac{1}{2}x_i^2 + \delta\Psi_j(x)\right)\right]\mathcal{H}^{K-1}(\mathrm{d}x),$$

$$\text{where } \delta\Psi_j(x) := \sum_{i=(j-1)K+1}^{jK} \delta\psi(x_i),$$

and clearly $\mathrm{osc}_{\mathbb{R}}(\delta\Psi_j) \le K\,\mathrm{osc}_{\mathbb{R}}(\delta\psi)$. Since the Gaussian measure satisfies LSI(1), the Holley–Stroock perturbation lemma (Criterion 2 in Section 1.1) gives (71). □

5.4. *Macroscopic LSI via convexification*

Recalling (68), Lemma 29 will follow immediately if we can show that

$$\forall m \in \mathbb{R}, \quad \psi_K''(m) \ge \lambda. \tag{72}$$

While the convexity cannot be true for all K (think that $\psi_1 = \psi$), we shall see that it does hold for K large enough (depending on ψ).

Let $x_1, \ldots, x_K$ be independent, identically distributed random variables with common law $Z^{-1}\exp(-\psi(x))\,\mathrm{d}x$. The joint law of $(x_1, \ldots, x_K)$ takes the form $Z_K^{-1}\exp(-\sum\psi(x_i))\mathcal{L}^K(\mathrm{d}x)$, so the mean value $m = (x_1 + \cdots + x_K)/K$ is distributed according to $f_K(m)\,\mathrm{d}m$, where

$$f_K(m) = \frac{1}{Z_K}\int_{X_{K,m}} \exp\left(-\sum_{i=1}^{K}\psi(x_i)\right)\mathcal{H}^{K-1}(\mathrm{d}x) = \frac{\exp(-K\psi_K(m))}{Z_K}.$$

Then the classical *Cramér theorem* in large deviation theory [6] asserts that for any $[a,b] \subset \mathbb{R}$,

$$-\frac{1}{K}\log\int_a^b f_K(m)\,\mathrm{d}m \underset{K\uparrow\infty}{\longrightarrow} \inf_{[a,b]}\varphi,$$

where φ is the so-called Cramér transform of $\exp(-\psi)$:

$$\varphi(m) = \sup_{\sigma\in\mathbb{R}}\left(\sigma m - \log\int_{\mathbb{R}}\exp\bigl(\sigma x - \psi(x)\bigr)\,\mathrm{d}x\right). \tag{73}$$

Since $(-1/K)\log f_K = \psi_K + (1/K)\log Z_K$, the following proposition will appear as a natural *local* refinement of Cramér's theorem:

Proposition 31 (Local Cramér theorem). *If $\psi:\mathbb{R}\to\mathbb{R}$ satisfies (16) and ψ_K is defined by (65) then*

$$\psi_K \underset{K\uparrow\infty}{\longrightarrow} \varphi \qquad \textit{in the uniform } C^2 \textit{ topology},$$

where φ is the Cramér transform of ψ, defined by (73).

(From here it is easy to deduce (72).) Although the result might be well known in certain circles (maybe in some equivalent version), for completeness we provide a proof of Proposition 31 in the Appendix, based on a C^2 local central limit theorem and the explicit representation

$$(\psi_N - \varphi)(m) = -\frac{1}{N}\log\left(\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}\right)(0).$$

Here $g_{N,m}$ denotes the distribution of $(1/\sqrt{N})\sum_{i=1}^N(x_i-m)$, where the x_i's are independent and distributed according to the common law $\mathrm{d}\mu^\sigma := Z^{-1}\exp(\sigma x-\psi(x))\,\mathrm{d}x$, and $\sigma=\sigma(m)$ is chosen so that the x_i's have mean m:

$$\int_{\mathbb{R}} x\mu^\sigma(\mathrm{d}x)=m. \tag{74}$$

(The idea of this "exponential" change of measure goes back to Cramér; see Section 5.5 below for details.)

5.5. *Statistical mechanics interpretation*

It might be enlightening to re-examine the preceding argument in terms of statistical mechanics. As discussed earlier (cf. Section 2.4), there are two statistical ensembles that play a natural role here: the *canonical ensemble* $\mu_{N,m}$ (mean density m fixed, exponential distribution of energy) and the *grand canonical ensemble* μ_N (density m not fixed, exponential distribution of energy). The one which is of direct relevance for us is the canonical ensemble: Our goal is to show the strict convexity of the associated *free energy*, which is by definition

$$\frac{1}{N}\log\left(\int_{X_{N,m}}\exp\left(\sum_{i=1}^N\bigl(\sigma x_i-\psi(x_i)\bigr)\right)\mathcal{H}^{N-1}(\mathrm{d}x)\right)\stackrel{(65)}{=}\sigma m-\psi_N(m). \tag{75}$$

As mentioned above (cf. (74)), Cramér's trick consists of changing the measure $\exp(-\psi(x))\,\mathrm{d}x$ into $\mathrm{d}\mu^\sigma=\exp(\sigma x-\psi(x))\,\mathrm{d}x$ with σ chosen so that the mean is m. This condition can be recast as

$$\frac{\mathrm{d}\varphi^*}{\mathrm{d}\sigma}(\sigma)=m, \tag{76}$$

where

$$\varphi^*(\sigma):=\log\int\exp\bigl(\sigma x-\psi(x)\bigr)\,\mathrm{d}x.$$

For each m there is a unique $\sigma=\sigma(m)$ solving (76). Indeed, as is well-known and easy to show (cf. Lemma 41 in Section A.2), φ^* is uniformly convex and C^1, so its derivative is continuous increasing. In fact m and σ are related by the usual equations of Legendre transform:

$$m=\bigl(\varphi^*\bigr)'(\sigma),\qquad \sigma=\varphi'(m),\qquad \varphi(m)+\varphi^*(\sigma)=m\sigma. \tag{77}$$

Consider now the *modified grand canonical ensemble* defined by

$$\mu_N^m=\frac{1}{Z}\exp\left(-\sum_{i=1}^N\bigl(\sigma x_i+\psi(x_i)\bigr)\right)\mathcal{L}^N(\mathrm{d}x). \tag{78}$$

As it did for μ_N, conditioning μ_N^m on $(1/N)\sum_{i=1}^N x_i=m$ gives rise to $\mu_{N,m}$ (adding a linear function to the Hamiltonian dows not affect $\mu_{N,m}$). What have we gained? In view of

$$\int_{\mathbb{R}^N}\left(\frac{1}{N}\sum_{i=1}^N x_i\right)\mu_N^m(\mathrm{d}x)\stackrel{(74)}{=}m,$$

the conditioning is expected to be less dramatic: μ_N^m should concentrate around the mean anyway by the usual law of large numbers. The principle of *equivalence of ensembles* says that $\mu_{N,m}$ and μ_N^m should be asymptotically close, in some sense, as $N\uparrow\infty$. Since the free energy of μ_N^m, given by

$$\log\int_{\mathbb{R}}\exp\bigl(\sigma x-\psi(x)\bigr)=\varphi^*(\sigma)=\sigma m-\varphi(m), \tag{79}$$

is strictly convex, does the free energy of the canonical ensemble inherit strict convexity for N large?

Statistical mechanics does suggest that one measure of the closeness of ensembles is the difference in free energies which, according to (75) and (79), is given by

$$\varphi(m) - \psi_N(m).$$

This is precisely the quantity that is controlled (in the C^2 topology) in Proposition 31.

To summarize: The local version of Cramér's theorem quantifies the equivalence of ensembles in the sense that it proves the asymptotic closeness, in the uniform C^2 topology, of the free energies.

This piece of information about the closeness of free energies is the only input from equilibrium statistical mechanics that we need in order to complete the proof of Theorem 14. In particular we do not need to invoke the closeness of low-dimensional marginals of μ_N^m and $\mu_{N,m}$ as in [5,10,14].

6. Application: Hydrodynamic limit for Kawasaki dynamics

6.1. *Set-up*: *functional spaces and projection*

In this section we shall use the same set-up as in Section 5.1. In particular, we will consider the spaces $X = X_{N,0}$ and $Y = Y_{M,0}$ with ℓ^2 inner product (61) and L^2 inner product (62), respectively. Moreover, we will project from X to Y using the projection operator $P = P_{N,K}$. For simplicity we shall assume that $N = KM$, where K is so large that ψ_K is uniformly convex (recall Section 5.4). The general case can be handled with arguments similar to those used in Remark 30.

6.2. *Proof of Theorem* 17

Let us start with a precise definition of the notion of weak solution that will be used.

Definition 32. *We will call $\zeta = \zeta(t,\theta)$ a weak solution of (27) on $[0,T]\times\mathbb{T}^1$ if*

$$\zeta \in L_t^\infty\big(L_\theta^2\big), \qquad \frac{\partial\zeta}{\partial t} \in L_t^2\big(H_\theta^{-1}\big), \qquad \varphi'(\zeta) \in L_t^2\big(L_\theta^2\big), \tag{80}$$

and

$$\left\langle \xi, \frac{\partial \zeta}{\partial t}\right\rangle_{H^{-1}} = -\int_{\mathbb{T}^1} \xi\varphi'(\zeta)\,\mathrm{d}\theta \quad \textit{for all } \xi \in L^2, \textit{ for almost every } t \in [0,T]. \tag{81}$$

As in Corollary 12, we will consider a sequence $\{M_\nu, N_\nu\}_{\nu=1}^\infty$ with

$$M_\nu \uparrow \infty, \qquad N_\nu \uparrow \infty, \qquad K_\nu = \frac{N_\nu}{M_\nu} \uparrow \infty. \tag{82}$$

To simplify notation we shall write just N, K, M for N_ν, K_ν, M_ν. We shall also not explicitly denote the dependence of $X, Y, P, \bar{A}, f_0$, and $\bar{H}$ on ν. Finally, we shall for simplicity abbreviate $\mu_{N,0} = \mu$ in this section.

Let $\bar{\eta}_0^\nu \in \bar{Y}$ be a step function approximation of ζ_0 with

$$\|\bar{\eta}_0^\nu - \zeta_0\|_{L^2} \xrightarrow[\nu\uparrow\infty]{} 0. \tag{83}$$

Since by assumption ζ_0 lies in $L^2(\mathbb{T}^1)$, we have

$$\|\bar{\eta}_0^\nu\|_{L^2} \le C. \tag{84}$$

Let $\eta_0^\nu \in Y$ be the vector associated to $\bar{\eta}_0^\nu$. Consider the solutions η^ν of

$$\frac{\mathrm{d}\eta^\nu}{\mathrm{d}t} = -\bar{A}\nabla_Y \bar{H}\big(\eta^\nu\big), \qquad \eta^\nu(0) = \eta_0^\nu. \tag{85}$$

The next proposition, proven in Section 6.4 below, is the key to pass from the abstract hydrodynamic limit (Corollary 12) to the "concrete" result of Theorem 17:

Proposition 33. *With the above notation, the step functions $\bar{\eta}^\nu$ converge strongly in $L^\infty(H^{-1})$ to the unique weak solution of*

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2}\varphi'(\zeta), \qquad \zeta(0,\cdot) = \zeta_0. \tag{86}$$

After these preparations, let us see how to prove Theorem 17.

Proof of Theorem 17. The proof is in two steps based respectively on Corollary 12, and on Proposition 33.

Step 1. *Abstract result.* In this step the goal is to show

$$\lim_{\nu\uparrow\infty}\sup_{t\in[0,T]}\int \langle \bar{x} - \bar{\eta}^\nu(t,\cdot), \bar{x} - \bar{\eta}^\nu(t,\cdot)\rangle_{H^{-1}} f(t,x)\mu(\mathrm{d}x) = 0. \tag{87}$$

This will be obtained as a consequence of Corollary 12. There are four assumptions that we need to check in order to apply this corollary:

(a) There exists $\alpha < \infty$ such that μ satisfies

$$\int |x|^2\mu(\mathrm{d}x) \le \alpha N. \tag{88}$$

(b) There exists $C < \infty$ such that the coarse-grained Hamiltonian satisfies

$$\inf_{y\in Y}\bar{H}(y) \ge -C, \tag{89}$$

$$\bar{H}(\eta_0^\nu) \le C \tag{90}$$

(c) There exists a universal constant $\gamma < \infty$ such that for P and A defined by (63) and (21) respectively, one has

$$\left|\left(\mathrm{id}_X - NP^tP\right)x\right|_X^2 \le \frac{\gamma}{M^2}\langle x, Ax\rangle_X. \tag{91}$$

(d) There exists $C < \infty$ such that for any $x \in X$, if $\bar{x}$ is the associated step function then

$$\frac{1}{C}\langle \bar{x}, \bar{x}\rangle_{H^{-1}} \le \frac{1}{N}\langle x, A^{-1}x\rangle_X \le C\langle \bar{x}, \bar{x}\rangle_{H^{-1}}; \tag{92}$$

moreover, if $\bar{x}$ is bounded in L^2, then

$$\left|\langle \bar{x}, \bar{x}\rangle_{H^{-1}} - \frac{1}{N}\langle x, A^{-1}x\rangle_X\right| \le \frac{C}{N}. \tag{93}$$

Taking estimates (a)–(d) for granted, let us verify the assumptions of Theorem 8:

- Assumptions (i)–(iii′) are a consequence of (16), Lemmas 28 and 29.
- Assumptions (iv), (v) and (vi) are the same as (88), (89) and (91), respectively.
- Assumption (vii) follows from (24) and (90).

So Theorem 8 applies. Moreover, (12) is given by (82). Hence, in order to invoke Corollary 12 we need only check (13).

To do so, we use estimate (92) to deduce

$$\begin{aligned}\frac{1}{N}\left(x - NP^t\eta_0^\nu\right)\cdot A^{-1}\left(x - NP^t\eta_0^\nu\right) &\le C\left\|\bar{x} - \bar{\eta}_0^\nu\right\|_{H^{-1}}^2 \\ &\le 2C\left(\|\bar{x} - \zeta_0\|_{H^{-1}}^2 + \left\|\zeta_0 - \bar{\eta}_0^\nu\right\|_{H^{-1}}^2\right),\end{aligned} \tag{94}$$

where we have observed $\overline{NP^t\eta_0^\nu} = \bar{\eta}_0^\nu$. The combination of (94), (83) and (25) gives in particular

$$\lim_{N\uparrow\infty}\int \frac{1}{N}\big(x - NP^t\eta_0^\nu\big)\cdot A^{-1}\big(x - NP^t\eta_0^\nu\big) f_0(x)\mu(\mathrm{d}x) = 0.$$

Hence, by Corollary 12, we have (14) which, by (92), gives (87).

Step 2. By combining Step 1 and Proposition 33, we obtain

$$\begin{aligned}
&\lim_{\nu\uparrow\infty}\sup_{t\in[0,T]}\int \|\bar{x} - \zeta(t,\cdot)\|_{H^{-1}}^2 f(t,x)\mu(\mathrm{d}x)\\
&\quad\le 2\lim_{\nu\uparrow\infty}\sup_{t\in[0,T]}\bigg(\int \|\bar{x} - \bar{\eta}^\nu(t,\cdot)\|_{H^{-1}}^2 f(t,x)\mu(\mathrm{d}x)\\
&\qquad + \int \|\bar{\eta}^\nu(t,\cdot) - \zeta(t,\cdot)\|_{H^{-1}}^2 f(t,x)\mu(\mathrm{d}x)\bigg)\\
&\quad = 2\lim_{\nu\uparrow\infty}\sup_{t\in[0,T]}\bigg(\int \|\bar{x} - \bar{\eta}^\nu(t,\cdot)\|_{H^{-1}}^2 f(t,x)\mu(\mathrm{d}x) + \|\bar{\eta}^\nu(t,\cdot) - \zeta(t,\cdot)\|_{H^{-1}}^2\bigg) = 0.
\end{aligned}$$

This concludes the proof of Theorem 17, modulo the proofs of estimates (88) to (93), and the proof of Proposition 33. □

6.3. *Proofs of estimates* (88)–(93)

Proof of (88). Since μ satisfies the logarithmic Sobolev inequality with constant ρ (cf. Theorem 14), it also satisfies the spectral gap inequality with the constant ρ, that is,

$$\int f(x)^2\mu(\mathrm{d}x) - \left(\int f(x)\mu(\mathrm{d}x)\right)^2 \le \frac{1}{\rho}\int \left|\nabla f(x)\right|^2\mu(\mathrm{d}x). \tag{95}$$

Setting $f(x) = x_i$ for any $i \in \{1,\dots,N\}$, we have

$$\int x_i^2\mu(\mathrm{d}x) = \int x_i^2\mu(\mathrm{d}x) - \left(\int x_i\mu(\mathrm{d}x)\right)^2 \overset{(95)}{\le} \frac{1}{\rho}\int \mu(\mathrm{d}x) = \frac{1}{\rho}.$$

Summing over i gives (88) with $\alpha = 1/\rho$. (Here we have used that the mean of x_i is zero. In general, the calculation gives $\alpha = 1/\rho + m^2$.) □

Proof of (89)–(90). Recall that $\bar{H}(\eta) = (1/M)\sum_{i=1}^M \psi_K(\eta_i) + (1/N)\log\bar{Z}$ where

$$\bar{Z} = \int_X \exp\left(-\sum_{i=1}^N \psi(x_i)\right)\mathcal{H}^{N-1}(\mathrm{d}x).$$

It is easy to see that (16) implies that $(1/N)\log\bar{Z}$ is bounded above and below, so we may without loss of generality assume that

$$\bar{H}(\eta) = \frac{1}{M}\sum_{i=1}^M \psi_K(\eta_i). \tag{96}$$

The lower bound (90) follows from the uniform convergence of ψ_K (cf. Proposition 31) and the strict convexity of φ.

For the upper bound (89), we recall from (77) that

$$\varphi''(m) = \frac{\mathrm{d}\sigma}{\mathrm{d}m}.$$

Together with Lemma 41 parts (i) and (ii), this implies

$$\frac{1}{C} \le \varphi''(m) \le C, \tag{97}$$

and in particular, we have

$$\varphi(m) \le C\left(1 + \frac{1}{2}m^2\right). \tag{98}$$

The uniform convergence of ψ_K (Proposition 31) together with (96) and (98) then gives

$$\bar{H}\big(\eta_0^\nu\big) \le C\big(1 + \langle \eta_0^\nu, \eta_0^\nu \rangle_Y\big) = C\big(1 + \|\bar{\eta}_0^\nu\|_{L^2}^2\big) \overset{(84)}{\le} C.$$

□

Proof of (91). Since $\mathrm{id}_X - N P^t P$ is in the kernel of P, each block of K spins has mean zero:

$$x_{(j-1)K+1} + \cdots + x_{jK} = 0.$$

Hence, each block satisfies the discrete Poincaré inequality

$$\sum_{i=(j-1)K+1}^{jK} x_i^2 \le \gamma K^2 \sum_{i=(j-1)K+1}^{jK-1} (x_i - x_{i+1})^2,$$

where $\gamma < \infty$ is a universal constant. Thus

$$\left|\big(\mathrm{id}_X - N P^t P\big)x\right|_X^2 \le \gamma K^2 \sum_{i=1}^{N} (x_i - x_{i+1})^2 \overset{(21)}{=} \gamma \frac{K^2}{N^2} \langle x, Ax \rangle_X = \gamma \frac{1}{M^2} \langle x, Ax \rangle_X.$$ □

Proof of (92)–(93). We can express the discrete norm as:

$$\frac{1}{N}\big\langle x, A^{-1}x \big\rangle_X = \frac{1}{N} \sum_{i=1}^{N} F_i^2, \quad \text{where } x_i = N(F_{i+1} - F_i) \text{ and } \sum_{i=1}^{N} F_i = 0. \tag{99}$$

We would like to estimate the H^{-1} norm in terms of F. For this recall from Definition 23 that

$$\langle \bar{x}, \bar{x} \rangle_{H^{-1}} = \int_{\mathbb{T}^1} w^2(\theta)\, \mathrm{d}\theta, \quad \text{where } w' = \bar{x} \text{ and } \int_{\mathbb{T}^1} w\, \mathrm{d}\theta = 0. \tag{100}$$

It is easy to check that w defined via

$$w(\theta) = F_i + N(F_{i+1} - F_i)\left(\theta - \frac{i-1}{N}\right) \quad \text{for } \theta \in \left[\frac{i-1}{N}, \frac{i}{N}\right)$$

satisfies (100). Hence

$$\begin{aligned} \langle \bar{x}, \bar{x} \rangle_{H^{-1}} &= \sum_{i=1}^{N} \int_0^{N^{-1}} \big(F_i + N(F_{i+1} - F_i)\theta\big)^2 \mathrm{d}\theta = \frac{1}{N} \sum_{i=1}^{N} \left(F_i^2 + (F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2\right) \\ &= \frac{1}{N}\big\langle x, A^{-1}x \big\rangle_X + \frac{1}{N} \sum_{i=1}^{N} \left((F_{i+1} - F_i)F_i + \frac{1}{3}(F_{i+1} - F_i)^2\right). \end{aligned} \tag{101}$$

Note that for the second term on the right-hand side we have

$$-\frac{2}{3N}\sum_{i=1}^{N}F_i^2 \le \frac{1}{N}\sum_{i=1}^{N}\left((F_{i+1}-F_i)F_i+\frac{1}{3}(F_{i+1}-F_i)^2\right)\le 0. \tag{102}$$

Using (102) in (101) gives

$$\langle\bar{x},\bar{x}\rangle_{H^{-1}} \le \frac{1}{N}\langle x, A^{-1}x\rangle_X \le 3\langle\bar{x},\bar{x}\rangle_{H^{-1}}.$$

Moreover, if $\bar{x}$ is bounded in L^2, then

$$\frac{1}{N}\sum_{i=1}^{N}x_i^2 = N\sum_{i=1}^{N}(F_{i+1}-F_i)^2 \le C \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N}F_i^2 \le C,$$

and (101) gives

$$\begin{aligned}
&\left|\langle\bar{x},\bar{x}\rangle_{H^{-1}} - \frac{1}{N}\langle x, A^{-1}x\rangle_X\right| \\
&\quad = \frac{1}{N}\left|\sum_{i=1}^{N}\left((F_{i+1}-F_i)F_i+\frac{1}{3}(F_{i+1}-F_i)^2\right)\right| \\
&\quad \le \frac{1}{N}\left(\sum_{i=1}^{N}(F_{i+1}-F_i)^2\sum_{i=1}^{N}F_i^2\right)^{1/2} + \frac{1}{3N}\sum_{i=1}^{N}(F_{i+1}-F_i)^2 \le \frac{C}{N}.
\end{aligned}$$

□

6.4. *Proof of Proposition 33*

In this section we prove Proposition 33, modulo certain auxiliary lemmas which are deferred to Section 6.5. For the sequence η^ν we have the bounds:

Lemma 34. *Consider the sequence $\{\eta^\nu\}_{\nu=1}^\infty$ of solutions of (85) subject to (84). There exists $C<\infty$ (independent of ν) such that*

$$\sup_{t\in[0,T]}\langle\eta^\nu(t),\eta^\nu(t)\rangle_Y \le C, \tag{103}$$

$$\int_0^T\left\langle\frac{\mathrm{d}\eta^\nu}{\mathrm{d}t}(t),(\bar{A})^{-1}\frac{\mathrm{d}\eta^\nu}{\mathrm{d}t}(t)\right\rangle_Y \mathrm{d}t \le C. \tag{104}$$

In particular, (103) implies that for the sequence of associated step functions $\bar{\eta}^\nu$, there is a subsequence such that

$$\bar{\eta}^\nu \rightharpoonup \eta_* \quad \text{weak-}{*}\text{ in } L^\infty(L^2) = \left(L^1(L^2)\right)^*$$

for some limit η_*. We will show that η_* is the unique weak solution of (86) by using the following four lemmas.

Lemma 35. *Consider a sequence $\{\eta^\nu\}_{\nu=1}^\infty$ of solutions of (85) satisfying (103) and (104). Consider any subsequence such that the associated step functions weak-$*$ converge in $(L^1(L^2))^*$ to a limit η_*. Then η_* satisfies (80). That is, on any bounded time interval we have*

$$\eta_* \in L^\infty(L^2), \qquad \frac{\partial\eta_*}{\partial t} \in L^2(H^{-1}), \qquad \varphi'(\eta_*) \in L^2(L^2).$$

Lemma 36 (Inequality formulation for convex potential). *Assume $\bar{H}$ is convex. Then η satisfies (85) if and only if*

$$\int_0^T \bar{H}(\eta)\beta(t)\,\mathrm{d}t \le \int_0^T \bar{H}(\eta+\xi)\beta(t)\,\mathrm{d}t - \int_0^T \big\langle \xi, (\bar{A})^{-1}\eta\big\rangle_Y \dot{\beta}(t)\,\mathrm{d}t \tag{105}$$

for all $\xi \in Y$ and smooth $\beta : [0,T] \to [0,\infty)$.

Similarly, assume φ is convex. Then ζ satisfies (81) if and only if

$$\begin{aligned}\int_0^T \int_{\mathbb{T}^1} \varphi\big(\zeta(t,\theta)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t \\ \le \int_0^T \int_{\mathbb{T}^1} \varphi\big(\zeta(t,\theta)+\xi(\theta)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t - \int_0^T \big\langle \xi(\cdot), \zeta(t,\cdot)\big\rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t\end{aligned} \tag{106}$$

for all $\xi \in L^2(\mathbb{T}^1)$ and smooth $\beta : [0,T] \to [0,\infty)$.

Lemma 37. *Suppose that the sequence $\{\eta^\nu\}_{\nu=1}^\infty$ satisfies (103), (104) and (105), and consider a subsequence such that*

$$\bar{\eta}^\nu \rightharpoonup \eta_* \quad \textit{weak-}{*} \textit{ in } L^\infty\big(L^2\big) = \big(L^1\big(L^2\big)\big)^*. \tag{107}$$

Let $\xi^\nu := \pi_\nu(\xi+\eta_*) - \eta^\nu$, where ξ is an arbitrary L^2 function and π_ν is the L^2-projection onto elements of Y.

Then we have

$$\text{(i)} \quad \lim_{\nu\uparrow\infty} \int_0^T \bar{H}\big(\eta^\nu(t)\big)\beta(t)\,\mathrm{d}t \ge \int_0^T \int_{\mathbb{T}^1} \varphi\big(\eta_*(t,\theta)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t,$$

$$\text{(ii)} \quad \lim_{\nu\uparrow\infty} \int_0^T \bar{H}\big(\eta^\nu(t)+\xi^\nu(t)\big)\beta(t)\,\mathrm{d}t = \int_0^T \int_{\mathbb{T}^1} \varphi\big(\eta_*(t,\theta)+\xi(\theta)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t,$$

$$\text{(iii)} \quad \lim_{\nu\uparrow\infty} \int_0^T \big\langle \xi^\nu(t), (\bar{A})^{-1}\eta^\nu(t)\big\rangle_Y \dot{\beta}(t)\,\mathrm{d}t = \int_0^T \big\langle \xi(\theta), \eta_*(t,\theta)\big\rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t. \tag{108}$$

Lemma 38 (Uniqueness). *There is at most one weak solution of* (86).

Proof of Proposition 33. According to Lemma 34, we can consider a subsequence such that

$$\bar{\eta}^\nu \rightharpoonup \eta_* \quad \text{weak-}{*} \text{ in } L^\infty\big(L^2\big) = \big(L^1\big(L^2\big)\big)^*$$

and strongly in $L^\infty(H^{-1})$. By Lemma 35, the limit η_* satisfies condition (80).

According to Lemma 36, η^ν satisfies

$$\int_0^T \bar{H}\big(\eta^\nu\big)\beta(t)\,\mathrm{d}t \le \int_0^T \bar{H}\big(\eta^\nu+\xi^\nu\big)\beta(t)\,\mathrm{d}t - \int_0^T \big\langle \xi^\nu, (\bar{A})^{-1}\eta^\nu\big\rangle_Y \dot{\beta}(t)\,\mathrm{d}t$$

with $\xi^\nu := \pi_\nu(\xi+\eta_*) - \eta^\nu$ as in Lemma 37. By applying (i), (ii) and (iii) from Lemma 37 to the inequality, we deduce that η_* satisfies (106); hence, by another application of Lemma 36, η_* satisfies (81).

Therefore η_* is a weak solution of (86), and since according to Lemma 38 there is only one weak solution, the full sequence $\{\bar{\eta}^\nu\}_{\nu=1}^\infty$ converges to the unique weak solution of (86). □

6.5. *Proofs of auxiliary lemmas*

Proof of Lemma 34. To see (103), we begin by computing

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\bar{H}\big(\eta^{\nu}\big) &= \big\langle \nabla_Y \bar{H}\big(\eta^{\nu}\big), \dot{\eta}^{\nu}\big\rangle_Y \\ &\overset{(85)}{=} -\big\langle \nabla_Y \bar{H}\big(\eta^{\nu}\big), \bar{A}\nabla_Y \bar{H}\big(\eta^{\nu}\big)\big\rangle_Y \le 0,\end{aligned} \tag{109}$$

so that

$$\bar{H}\big(\eta^{\nu}(t)\big) \le \bar{H}\big(\eta_0^{\nu}\big). \tag{110}$$

Combining (90) and the strict convexity of $\bar{H}$ (recall that K has been chosen large enough that $\bar{H}$ is strictly convex by Proposition 31), we conclude that there exists $C > 0$ such that

$$\bar{H}(\eta) \ge -C + \frac{1}{C}\langle \eta, \eta\rangle_Y, \tag{111}$$

which gives the uniform in time bound

$$\big\langle \eta^{\nu}, \eta^{\nu}\big\rangle_Y \overset{(111)}{\le} C\big(\bar{H}\big(\eta^{\nu}\big) + C\big) \overset{(110)}{\le} C\big(\bar{H}\big(\eta_0^{\nu}\big) + C\big) \overset{(89)}{\le} 2C^2.$$

To see (104), we substitute $\nabla_Y H(\eta^{\nu}) = -(\bar{A})^{-1}\dot{\eta}^{\nu}$ in (109) and integrate in time to deduce

$$\begin{aligned}\int_0^T \big\langle \dot{\eta}^{\nu}, (\bar{A})^{-1}\dot{\eta}^{\nu}\big\rangle_Y \,\mathrm{d}t &= -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\bar{H}\big(\eta^{\nu}\big)\,\mathrm{d}t \\ &= \bar{H}\big(\eta_0^{\nu}\big) - \bar{H}\big(\eta^{\nu}(t)\big) \overset{(89),(90)}{\le} 2C.\end{aligned}$$

□

Proof of Lemma 35. By weak lower semicontinuity, we have for all $t \in [0, T]$ that

$$\int_{\mathbb{T}^1} \eta_*^2 \,\mathrm{d}\theta \le \liminf_{\nu\uparrow\infty} \int_{\mathbb{T}^1} \big(\bar{\eta}^{\nu}\big)^2 \,\mathrm{d}\theta = \big\langle \eta^{\nu}, \eta^{\nu}\big\rangle_Y \overset{(103)}{\le} C,$$

which is the first estimate.

For the second estimate we employ (92) with $x = NP^t\dot{\eta}^{\nu}$. Recalling that $\overline{NP^t\eta^{\nu}} = \bar{\eta}^{\nu}$, we have

$$\begin{aligned}\int_0^T \big\langle \dot{\bar{\eta}}^{\nu}, \dot{\bar{\eta}}^{\nu}\big\rangle_{H^{-1}} \,\mathrm{d}t &\overset{(92)}{\le} \frac{C}{N}\int_0^T \big\langle NP^t\dot{\eta}^{\nu}, A^{-1}NP^t\dot{\eta}^{\nu}\big\rangle_X \,\mathrm{d}t \\ &\overset{(9)}{=} C\int_0^T \big\langle \dot{\eta}, \bar{A}^{-1}\dot{\eta}\big\rangle_Y \,\mathrm{d}t \overset{(104)}{\le} C.\end{aligned}$$

Again using weak lower semicontinuity, we have in particular that

$$\int_0^T \langle \dot{\eta}_*, \dot{\eta}_*\rangle_{H^{-1}} \,\mathrm{d}t \le C.$$

The third estimate follows from the first observation along with

$$\big|\varphi'(m)\big| \le C\big(1 + |m|\big),$$

which is a consequence of (97). □

Proof of Lemma 38. Consider the weak formulation:

$$\langle \dot{\zeta}, \xi \rangle_{H^{-1}} = -\int_{\mathbb{T}^1} \varphi'(\zeta)\xi \, \mathrm{d}\theta \quad \text{for all } \xi \in L^2, \text{for a.e. } t \in [0, T].$$

Suppose that ζ_1 and ζ_2 are two weak solutions. Then we have

$$\langle \dot{\zeta}_1 - \dot{\zeta}_2, \xi \rangle_{H^{-1}} = -\int_{\mathbb{T}^1} \big(\varphi'(\zeta_1) - \varphi'(\zeta_2)\big)\xi \, \mathrm{d}\theta.$$

Recall from the definition of weak solution that

$$\zeta_i \in L^\infty\big(L^2\big), \qquad \dot{\zeta}_i \in L^2\big(H^{-1}\big), \qquad \varphi'(\zeta_i) \in L^2\big(L^2\big) \quad \text{for } i = 1, 2.$$

Hence we may choose $\xi = \zeta_1 - \zeta_2$ as a test function and deduce for a.e. $t \in [0, T]$ that

$$\frac{\mathrm{d}}{\mathrm{d}t}\big\langle \zeta_1(t) - \zeta_2(t), \zeta_1(t) - \zeta_2(t)\big\rangle_{H^{-1}} = -2\int_{\mathbb{T}^1} \big(\varphi'(\zeta_1) - \varphi'(\zeta_2)\big)(\zeta_1 - \zeta_2) \, \mathrm{d}\theta \le 0,$$

by the convexity of φ. Hence $\zeta_1 = \zeta_2$. □

Proof of Lemma 36. We will show that (85) is equivalent to (105). The equivalence of (81) and (106) follows analogously.

First let us rewrite (85) in weak form as

$$\int_0^T \big\langle \xi, (\bar{A})^{-1}\eta \big\rangle_Y \dot{\beta}(t) \, \mathrm{d}t = \int_0^T \big\langle \xi, \nabla_Y \bar{H}(\eta) \big\rangle_Y \beta(t) \, \mathrm{d}t \tag{112}$$

for all $\xi \in Y$ and all smooth $\beta : [0, T] \to [0, \infty)$. We begin by showing that (112) implies (105). By convexity, we have

$$\big\langle \xi, \nabla_Y \bar{H}(\eta) \big\rangle_Y \le -\bar{H}(\eta) + \bar{H}(\eta + \xi). \tag{113}$$

Inserting (113) into (112) gives

$$\begin{aligned} &\int_0^T \big\langle \xi, (\bar{A})^{-1}\eta \big\rangle_Y \dot{\beta}(t) \, \mathrm{d}t \\ &\quad \le -\int_0^T \bar{H}(\eta)\beta(t) \, \mathrm{d}t + \int_0^T \bar{H}(\eta + \xi)\beta(t) \, \mathrm{d}t \end{aligned} \tag{114}$$

which, after rearranging terms, is (105).

To show that (105) implies (112), we substitute $\tilde{\xi} = \varepsilon\xi$ in Eq. (105), for some $\varepsilon > 0$ and ξ in Y. Dividing both sides by ε and rearranging terms, we have

$$\int_0^T \big\langle \xi, (\bar{A})^{-1}\eta \big\rangle_Y \dot{\beta}(t) \, \mathrm{d}t \le \int_0^T \frac{\bar{H}(\eta + \varepsilon\xi) - \bar{H}(\eta)}{\varepsilon} \beta(t) \, \mathrm{d}t.$$

Taking the limit $\varepsilon \downarrow 0$ returns

$$\int_0^T \big\langle \xi, (\bar{A})^{-1}\eta \big\rangle_Y \dot{\beta}(t) \, \mathrm{d}t \le \int_0^T \big\langle \xi, \nabla_Y \bar{H}(\eta) \big\rangle_Y \beta(t) \, \mathrm{d}t.$$

Repeating the process with $\tilde{\xi} = -\varepsilon\xi$ gives

$$\int_0^T \big\langle \xi, (\bar{A})^{-1}\eta \big\rangle_Y \dot{\beta}(t) \, \mathrm{d}t \ge \int_0^T \big\langle \xi, \nabla_Y \bar{H}(\eta) \big\rangle_Y \beta(t) \, \mathrm{d}t,$$

establishing (112). □

Proof of Lemma 37. As in the proof of (89)–(90), we may assume without loss that $\bar{H}$ is defined by (96), i.e., that

$$\bar{H}(\eta^\nu) = \frac{1}{M}\sum_{i=1}^{M}\psi_K(\eta_i^\nu). \tag{115}$$

Proof of (i): The uniform convergence $\psi_K \to \varphi$, the convexity of φ, and the consequent weak lower-semicontinuity give

$$\begin{aligned}
&\int_0^T \bar{H}(\eta^\nu)\beta(t)\,\mathrm{d}t \\
&\overset{(115)}{=} \int_0^T \int_{\mathbb{T}^1} \psi_K(\bar{\eta}^\nu)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t \\
&= \int_0^T \int_{\mathbb{T}^1} \varphi(\bar{\eta}^\nu)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t + \int_0^T \int_{\mathbb{T}^1} \big(\psi_K(\bar{\eta}^\nu) - \varphi(\bar{\eta}^\nu)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t \\
&\geq \int_0^T \int_{\mathbb{T}^1} \varphi(\eta_*)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t - \mathrm{o}(1)_{M\uparrow\infty} - \int_0^T \int_{\mathbb{T}^1} \sup_{\mathbb{R}}|\psi_K - \varphi|\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t \\
&\geq \int_0^T \int_{\mathbb{T}^1} \varphi(\eta_*)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t - \mathrm{o}(1)_{M\uparrow\infty} - \mathrm{o}(1)_{K\uparrow\infty}.
\end{aligned}$$

Proof of (ii): By choice of ξ^ν, we have $\eta^\nu + \xi^\nu = \pi_\nu(\eta_* + \xi)$, so that in particular

$$\bar{\eta}^\nu + \bar{\xi}^\nu \to \eta_* + \xi$$

strongly in L^2 for a.e. t as $M \uparrow \infty$. Because of the quadratic bounds on φ (cf. (97)), φ is continuous with respect to strong L^2 convergence, so that

$$\int_{\mathbb{T}^1} \varphi(\bar{\eta}^\nu + \bar{\xi}^\nu)\,\mathrm{d}\theta \to \int_{\mathbb{T}^1} \varphi(\eta_* + \xi)\,\mathrm{d}\theta$$

for a.e. t as $M \uparrow \infty$. We use the uniform in time bound

$$\left|\int_{\mathbb{T}^1} \varphi(\bar{\eta}^\nu + \bar{\xi}^\nu)\,\mathrm{d}\theta\right| \leq C\left(1 + \int_{\mathbb{T}^1} \left|\bar{\eta}^\nu + \bar{\xi}^\nu\right|^2\mathrm{d}\theta\right) \leq C\left(1 + \int_{\mathbb{T}^1} |\eta_* + \xi|^2\,\mathrm{d}\theta\right) \leq C$$

in the Dominated Convergence Theorem to conclude that

$$\lim_{M\uparrow\infty} \int_0^T \int_{\mathbb{T}^1} \varphi(\bar{\eta}^\nu + \bar{\xi}^\nu)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t = \int_0^T \int_{\mathbb{T}^1} \varphi(\eta_* + \xi)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t. \tag{116}$$

Together with the uniform convergence of ψ_K, this gives

$$\begin{aligned}
&\left|\int_0^T \bar{H}(\eta^\nu + \xi^\nu)\beta(t)\,\mathrm{d}t - \int_0^T \int_{\mathbb{T}^1} \varphi(\eta_* + \xi)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t\right| \\
&\overset{(115)}{=} \left|\int_0^T \int_{\mathbb{T}^1} \psi_K(\bar{\eta}^\nu + \bar{\xi}^\nu)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t - \int_0^T \int_{\mathbb{T}^1} \varphi(\eta_* + \xi)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t\right| \\
&\leq \left|\int_0^T \int_{\mathbb{T}^1} \big(\psi_K(\bar{\eta}^\nu + \bar{\xi}^\nu) - \varphi(\bar{\eta}^\nu + \bar{\xi}^\nu)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t\right| \\
&\quad + \left|\int_0^T \int_{\mathbb{T}^1} \big(\varphi(\bar{\eta}^\nu + \bar{\xi}^\nu) - \varphi(\eta_* + \xi)\big)\beta(t)\,\mathrm{d}\theta\,\mathrm{d}t\right|
\end{aligned}$$

$$\overset{(116)}{\leq} \left| \int_0^T \int_{\mathbb{T}^1} \sup_{\mathbb{R}} |\psi_K - \varphi| \beta(t)\,\mathrm{d}\theta\,\mathrm{d}t \right| + \mathrm{o}(1)_{M\uparrow\infty}$$
$$= \mathrm{o}(1)_{K\uparrow\infty} + \mathrm{o}(1)_{M\uparrow\infty}.$$

Proof of (iii): Recalling that $\xi^\nu = \pi_\nu(\eta_* + \xi) - \eta^\nu$, (iii) will follow from

$$\lim_{\nu\uparrow\infty} \int_0^T \left\langle \pi_\nu(\eta_*+\xi), (\bar{A})^{-1}\eta^\nu \right\rangle_Y \dot{\beta}(t)\,\mathrm{d}t = \int_0^T \langle \eta_*+\xi, \eta_* \rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t, \tag{117}$$

$$\lim_{\nu\uparrow\infty} \int_0^T \left\langle \eta^\nu, (\bar{A})^{-1}\eta^\nu \right\rangle_Y \dot{\beta}(t)\,\mathrm{d}t = \int_0^T \langle \eta_*, \eta_* \rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t. \tag{118}$$

Equation (117) is a consequence of two facts:

$$\left| \left\langle \pi_\nu(\eta_*+\xi), (\bar{A})^{-1}\eta^\nu \right\rangle_Y - \left\langle \bar{\pi}_\nu(\eta_*+\xi), \bar{\eta}^\nu \right\rangle_{H^{-1}} \right| \overset{(93)}{\leq} \frac{C}{N},$$

$$\lim_{\nu\uparrow\infty} \int_0^T \left\langle \bar{\pi}_\nu(\eta_*+\xi), \bar{\eta}^\nu \right\rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t = \int_0^T \langle \eta_*+\xi, \eta_* \rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t,$$

the second line following from the strong convergence of $\bar{\pi}_\nu(\eta_*+\xi)$ and $\bar{\eta}^\nu$ in $L^\infty(H^{-1})$.

For Eq. (118), on the other hand, it is enough to establish the identity

$$\left\langle \eta, (\bar{A})^{-1}\eta \right\rangle_Y = \langle \tilde{x}, \tilde{x} \rangle_{H^{-1}},$$
$$\text{where } x = NP^t\eta \text{ and } \tilde{x} = \sum_{i=1}^N \frac{x_i}{N}\delta_{i/N} \tag{119}$$

together with the convergence

$$\lim_{\nu\uparrow\infty} \int_0^T \left\langle \tilde{x}^\nu, \tilde{x}^\nu \right\rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t = \int_0^T \langle \eta_*, \eta_* \rangle_{H^{-1}} \dot{\beta}(t)\,\mathrm{d}t, \tag{120}$$

$$\text{where } x^\nu = NP^t\eta^\nu \text{ and } \tilde{x}^\nu = \sum_{i=1}^N \frac{x_i^\nu}{N}\delta_{i/N}. \tag{121}$$

Turning first to (119), we recall from (9) that

$$\left\langle \eta, (\bar{A})^{-1}\eta \right\rangle_Y = \left\langle \eta, PA^{-1}NP^t\eta \right\rangle_Y = \frac{1}{N}\left\langle NP^t\eta, A^{-1}NP^t\eta \right\rangle_X = \frac{1}{N}\left\langle x, A^{-1}x \right\rangle_X,$$

so that it is enough to show

$$\frac{1}{N}\left\langle x, A^{-1}x \right\rangle_X = \langle \tilde{x}, \tilde{x} \rangle_{H^{-1}}. \tag{122}$$

On the one hand, we have

$$\frac{1}{N}\left\langle x, A^{-1}x \right\rangle_X = \frac{1}{N}\sum_{i=1}^N F_i^2,$$
$$\text{where } x_i = N(F_{i+1} - F_i) \text{ and } \sum_{i=1}^N F_i = 0. \tag{123}$$

On the other hand, recall from Definition 23 that

$$\langle \tilde{x}, \tilde{x}\rangle_{H^{-1}} = \int_{\mathbb{T}^1} w^2(\theta)\, d\theta, \quad \text{where } w' = \tilde{x} \text{ and } \int_{\mathbb{T}^1} w\, d\theta = 0.$$

Because of the structure of $\tilde{x}$, we see that w is constant on $((j-1)/N, j/N)$ for $j = 1, \dots, N$. Denoting the constant values by $w_1, \dots, w_N$, respectively, we have that

$$\langle \tilde{x}, \tilde{x}\rangle_{H^{-1}} = \frac{1}{N}\sum_{i=1}^N w_i^2, \quad \text{where } w_{i+1} - w_i = \frac{x_i}{N} \text{ and } \sum_{i=1}^N w_i = 0. \tag{124}$$

Comparing (123) and (124), we conclude that $w_i = F_i$ for $i = 1, \dots, N$ and that (122) holds.

We turn now to the proof of (120). We remark first that $\tilde{x}^\nu$ weak-$*$ converges to η_* in $C([0,T]\times\mathbb{T}^1)^*$. Indeed, let $g \in C([0,T]\times\mathbb{T}^1)$ and notice that

$$\int_0^T \int_{\mathbb{T}^1} \tilde{x}^\nu g\, d\theta\, dt = \int_0^T \frac{1}{N}\sum_{i=1}^N x_i^\nu g\left(\frac{i}{N}\right) dt \underset{N\uparrow\infty}{\longrightarrow} \int_0^T \int_{\mathbb{T}^1} \eta_* g\, d\theta\, dt.$$

We will now show that $\tilde{u}^\nu := (\frac{\partial^2}{\partial\theta^2})^{-1}\tilde{x}^\nu$ is uniformly Hölder-1/2 in time and space, so that by the Arzela–Ascoli Theorem, a subsequence converges uniformly on $[0,T]\times\mathbb{T}^1$. By uniqueness of the limit, it follows that the limit must be $(\frac{\partial^2}{\partial\theta^2})^{-1}\eta_*$, and hence the full sequence converges. Together with the weak-$*$ convergence of $\tilde{x}^\nu$ this gives

$$\lim_{\nu\uparrow\infty} \int_0^T \int_{\mathbb{T}^1} \tilde{x}^\nu \tilde{u}^\nu \dot{\beta}(t)\, d\theta\, dt = \int_0^T \int_{\mathbb{T}^1} \eta_* \left(\frac{\partial^2}{\partial\theta^2}\right)^{-1} \eta_* \dot{\beta}(t)\, d\theta\, dt,$$

which is equivalent to (120).

Hence, it remains only to deduce the uniform bounds on $\tilde{u}^\nu$. In time, we use the uniform bound on $\frac{\partial}{\partial t}\tilde{x}^\nu$ in $L^2(H^{-1})$ which follows from

$$\begin{aligned}\int_0^T \left\langle \frac{\partial}{\partial t}\tilde{x}^\nu, \frac{\partial}{\partial t}\tilde{x}^\nu \right\rangle_{H^{-1}} dt &\overset{(122)}{=} \frac{1}{N}\int_0^T \left\langle \frac{\partial}{\partial t}x^\nu, A^{-1}\frac{\partial}{\partial t}x^\nu \right\rangle_X dt \\ &\overset{(121)}{=} \int_0^T \left\langle \frac{\partial}{\partial t}\eta^\nu, (\bar{A})^{-1}\frac{\partial}{\partial t}\eta^\nu \right\rangle_Y dt \overset{(104)}{\leq} C.\end{aligned}$$

We deduce a uniform bound on $\frac{\partial}{\partial t}\tilde{u}^\nu$ in $L^2(H^1)$, which by Sobolev imbedding implies

$$\int_0^T \sup_{\theta\in\mathbb{T}^1} \left|\frac{\partial}{\partial t}\tilde{u}^\nu\right|^2 dt \leq C,$$

and in particular,

$$\begin{aligned}\sup_{\theta\in\mathbb{T}^1} \left|u(t,\theta) - u(s,\theta)\right| &= \sup_{\theta\in\mathbb{T}^1} \left|\int_s^t \frac{\partial}{\partial\tau}\tilde{u}^\nu(\tau,\theta)\, d\tau\right| \\ &\leq \sup_{\theta\in\mathbb{T}^1} \left(\int_s^t \left|\frac{\partial}{\partial\tau}\tilde{u}^\nu(\tau,\theta)\right|^2 d\tau\right)^{1/2} \sqrt{|t-s|} \leq C\sqrt{|t-s|},\end{aligned}$$

which is the uniform Hölder-1/2 bound in time.

In space, we use the uniform bound on $\tilde{x}^\nu$ in $H^{-1}(\mathbb{T}^1)$:

$$\langle \tilde{x}^\nu, \tilde{x}^\nu \rangle_{H^{-1}} \overset{(122)}{=} \frac{1}{N}\langle x, A^{-1}x\rangle_X \overset{(92)}{\leq} C\langle \bar{x}, \bar{x}\rangle_{H^{-1}} \overset{(121)}{=} C\langle \bar{\eta}^\nu, \bar{\eta}^\nu\rangle_{H^{-1}}$$
$$\leq C\int_{\mathbb{T}^1} (\bar{\eta}^\nu)^2 \,\mathrm{d}\theta = C\langle \eta^\nu, \eta^\nu\rangle_Y \overset{(103)}{\leq} C$$

uniformly in time, which implies a uniform bound on $\tilde{u}^\nu$ in $H^1(\mathbb{T}^1)$. The uniform Hölder-$1/2$ bound in space follows from Sobolev imbedding. □

Appendix: Local Cramér theorem

The main result in this Appendix is Proposition 31. Some elements of the proof may also be found, for instance, in [7], Chapter XVI, [12], Appendix 2, [10], Section 3, and [14], p. 752 and Section 5.

A.1. *Proof of Proposition* 31

The proof hinges on Cramér's representation:

$$\exp\big(N\big(\varphi(m) - \psi_N(m)\big)\big) = \frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}(0), \tag{125}$$

which is easy to check by direct substitution and the Coarea Formula. The result will follow from uniform C^2 estimates for the right-hand side of (125).

Proof of Proposition 31. Below we will develop the uniform bounds

$$\frac{1}{C} \leq \frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}(0) \leq C, \qquad \left|\frac{\mathrm{d}^2}{\mathrm{d}m^2}\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}(0)\right| \leq C, \tag{126}$$

where C denotes a generic constant independent of both m and N. By interpolation, it will also follow that

$$\left|\frac{\mathrm{d}}{\mathrm{d}m}\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}(0)\right| \leq C.$$

Taking derivatives in (125) and applying the bounds, we deduce the uniform convergence of ψ_N, $\frac{\mathrm{d}\psi_N}{\mathrm{d}m}$ and $\frac{\mathrm{d}^2\psi_N}{\mathrm{d}m^2}$ as $N \uparrow \infty$.

To begin, recall that $g_{N,m}$ describes the distribution of a sum of independent variables. Hence $\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}$ can be written as a convolution. Defining the Fourier transform as

$$\mathcal{F}[f](\xi) := \int_{\mathbb{R}} \exp(\mathrm{i}\xi x) f(x)\mathcal{L}^1(\mathrm{d}x)$$

and recalling that convolution turns into multiplication under the Fourier transform, we can re-express the right-hand side of (125) as

$$\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}(0) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\left[\frac{\mathrm{d}g_{N,m}}{\mathrm{d}\mathcal{L}^1}\right](\xi)\mathcal{L}^1(\mathrm{d}\xi) = \frac{1}{2\pi}\int_{\mathbb{R}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi), \tag{127}$$

where

$$h(m,\xi) := \exp(-\mathrm{i}\xi m)\mathcal{F}\left[\frac{\mathrm{d}\mu_m}{\mathrm{d}\mathcal{L}^1}\right](\xi). \tag{128}$$

From the following useful representation,

$$\frac{d\mu_m}{d\mathcal{L}^1}(x) \overset{(78),(79)}{=} \exp\bigl(-\varphi^*(\sigma) + \sigma x - \psi(x)\bigr) \tag{129}$$

$$\overset{(77)}{=} \quad \exp\bigl(\varphi(m) + \sigma(x-m) - \psi(x)\bigr), \tag{130}$$

we can re-express (128) as

$$h(m,\xi) \overset{(130)}{=} \int_{\mathbb{R}} \exp\bigl(-\mathrm{i}\xi m + \mathrm{i}\xi x - \varphi^*(\sigma) + \sigma x - \psi(x)\bigr)\mathcal{L}^1(dx). \tag{131}$$

Using (127) and the nonnegativity of $g_{N,m}$, it follows that (126) is proved once we establish:

$$\frac{1}{C} \le \left| \int_{\mathbb{R}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) \right| \le C \tag{132}$$

and

$$\left| \frac{d^2}{dm^2} \int_{\mathbb{R}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) \right| \le C. \tag{133}$$

We will establish (132) and (133) by splitting the integrals into "inner" integrals over $\{N^{-1/2}|\xi| \le \delta\}$ and "outer" integrals over the complement. More precisely, on the one hand we show that there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$ and all $m \in \mathbb{R}$,

$$\left| \int_{\{N^{-1/2}|\xi| \le \delta\}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) \right| \le C, \tag{134}$$

$$\operatorname{Re} \int_{\{N^{-1/2}|\xi| \le \delta\}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) \ge 1/C, \tag{135}$$

$$\left| \frac{d^2}{dm^2} \int_{\{N^{-1/2}|\xi| \le \delta\}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) \right| \le C. \tag{136}$$

On the other hand, we will argue that for any $\delta > 0$, we have

$$\lim_{N\uparrow\infty} \int_{\{N^{-1/2}|\xi| \ge \delta\}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) = 0, \tag{137}$$

$$\lim_{N\uparrow\infty} \frac{d^2}{dm^2} \int_{\{N^{-1/2}|\xi| \ge \delta\}} h^N\bigl(m, N^{-1/2}\xi\bigr)\mathcal{L}^1(d\xi) = 0, \tag{138}$$

uniformly in m. The combination of (134)–(138) yields (132) and (133).

First consider the outer integrals. We control h and its derivatives using:

Lemma 39. *For h defined by* (131) *and any $\delta > 0$, there exists a positive constant C_δ* (*uniform in m*) *such that for all $|\xi| > \delta$:*

(i) $\bigl|h(m,\xi)\bigr| \le \dfrac{1}{1 + |\xi|/C_\delta}$,

(ii) $\left|\dfrac{\partial h}{\partial m}(m,\xi)\right| \le C_\delta |\xi|$,

(iii) $\left|\dfrac{\partial^2 h}{\partial m^2}(m,\xi)\right| \le C_\delta |\xi|^2$.

Lemma 39(i) implies (137):

$$
\begin{aligned}
&\left| \int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) \right| \\
&\quad = N^{1/2}\left| \int_{\{|\hat{\xi}|\geq\delta\}} h^N(m,\hat{\xi})\mathcal{L}^1(\mathrm{d}\hat{\xi}) \right| \\
&\quad \leq N^{1/2}\left(\frac{1}{1+\delta/C_\delta}\right)^{N-2} \int_{\{|\hat{\xi}|\geq\delta\}} \left(\frac{1}{1+|\hat{\xi}|/C_\delta}\right)^2 \mathcal{L}(\mathrm{d}\hat{\xi}) \underset{N\uparrow\infty}{\to} 0.
\end{aligned}
$$

For (138) we notice that

$$
\begin{aligned}
&\frac{\mathrm{d}^2}{\mathrm{d}m^2} \int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) \\
&\quad = \int_{\{N^{-1/2}|\xi|\geq\delta\}} N^2 h^{N-2}\big(m, N^{-1/2}\xi\big)\left(\frac{\partial h}{\partial m}\big(m, N^{-1/2}\xi\big)\right)^2 \mathcal{L}^1(\mathrm{d}\xi) \\
&\qquad + \int_{\{N^{-1/2}|\xi|\geq\delta\}} N h^{N-1}\big(m, N^{-1/2}\xi\big)\frac{\partial^2 h}{\partial m^2}\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi),
\end{aligned}
$$

so that by Lemma 39(ii) and (iii),

$$
\begin{aligned}
&\left| \frac{\mathrm{d}^2}{\mathrm{d}m^2} \int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) \right| \\
&\quad \leq C_\delta \int_{\{N^{-1/2}|\xi|\geq\delta\}} N^2 \big|h\big(m, N^{-1/2}\xi\big)\big|^{N-2}|\xi|^2\mathcal{L}^1(\mathrm{d}\xi) \\
&\quad = C_\delta N^{5/2} \int_{\{|\hat{\xi}|\geq\delta\}} \big|h(m,\hat{\xi})\big|^{N-2}|\hat{\xi}|^2\mathcal{L}^1(\mathrm{d}\hat{\xi}).
\end{aligned}
$$

We appeal once more to Lemma 39(i) to conclude

$$
\begin{aligned}
&\left| \frac{\mathrm{d}^2}{\mathrm{d}m^2} \int_{\{N^{-1/2}|\xi|\geq\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) \right| \\
&\qquad \leq C_\delta N^{5/2}\left(\frac{1}{1+\delta/C_\delta}\right)^{N-6} \int_{\{|\hat{\xi}|\geq\delta\}} \left(\frac{1}{1+|\hat{\xi}|/C_\delta}\right)^4 |\hat{\xi}|^2\mathcal{L}^1(\mathrm{d}\hat{\xi}) \underset{N\uparrow\infty}{\to} 0,
\end{aligned}
$$

establishing (138).

We now turn to the inner integrals. Since μ_m is a probability measure with mean m, we have

$$
\begin{aligned}
&h(m,0)=1, \qquad \frac{\partial h}{\partial \xi}(m,0)=0 \quad \text{and} \\
&-\frac{\partial^2 h}{\partial \xi^2}(m,0) = \int (x-m)^2\mu_m(\mathrm{d}x) = \mathrm{Var}(\mu_m) > 0.
\end{aligned}
\tag{139}
$$

According to Lemma 41(ii) in Section A.2, the variance of μ_m is bounded uniformly above and below:

$$
1/C \leq \mathrm{Var}(\mu_m) \leq C. \tag{140}
$$

It follows from the lower bound and Taylor's theorem (see also the proof of Lemma 40 in Section A.2) that there exists $h_2(m,\xi)$ defined on a uniform δ-neighborhood of $\xi = 0$ such that

$$
h(m,\xi) = \exp\big(-\xi^2 h_2(m,\xi)\big), \tag{141}
$$

and

$$h_2(m,0) = \mathrm{Var}(\mu_m). \tag{142}$$

The motivation for introducing h_2 is the formula

$$h^N\big(m, N^{-1/2}\xi\big) = \exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big). \tag{143}$$

The necessary control on h_2 is given by:

Lemma 40. *There exist $\delta > 0$ and $C < \infty$ (uniform in m) such that for $|\xi| \le \delta$ and all $m \in \mathbb{R}$:*

(i) $\left|\dfrac{\partial h_2}{\partial \xi}(m,\xi)\right| \le C,$

(ii) $\left|\dfrac{\partial h_2}{\partial m}(m,\xi)\right| \le C,$

(iii) $\left|\dfrac{\partial^2 h_2}{\partial m^2}(m,\xi)\right| \le C.$

Equipped with Lemma 40, we will now establish (134)–(135). In view of (143), we have

$$\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) = \int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big)\mathcal{L}^1(\mathrm{d}\xi). \tag{144}$$

According to (140) and Lemma 40(i), we have for $|\hat{\xi}| \le \delta$

$$\mathrm{Re}\, h_2(m,\hat{\xi}) \ge 1/C.$$

Thus, for $N^{-1/2}|\xi| \le \delta$, we have

$$\big|\exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big)\big| \le \exp\big(-\xi^2/C\big), \tag{145}$$

so that

$$\left|\int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big)\mathcal{L}^1(\mathrm{d}\xi)\right| \le C.$$

In view of (144), this proves (134).

The proof of (136) is similar. Applying (144), we have

$$\begin{aligned}
&\frac{\mathrm{d}^2}{\mathrm{d}m^2}\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi) \\
&\quad= \int_{\{N^{-1/2}|\xi|\le\delta\}} -\xi^2 \frac{\partial^2 h_2}{\partial m^2}\big(m, N^{-1/2}\xi\big)\exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big)\mathcal{L}^1(\mathrm{d}\xi) \\
&\qquad+ \int_{\{N^{-1/2}|\xi|\le\delta\}} \xi^4\left(\frac{\partial h_2}{\partial m}\right)^2\big(m, N^{-1/2}\xi\big)\exp\big(-\xi^2 h_2\big(m, N^{-1/2}\xi\big)\big)\mathcal{L}^1(\mathrm{d}\xi).
\end{aligned}$$

According to Lemma 40(ii) and (iii) and (145), this identity yields the estimate

$$\begin{aligned}
&\left|\frac{\mathrm{d}^2}{\mathrm{d}m^2}\int_{\{N^{-1/2}|\xi|\le\delta\}} h^N\big(m, N^{-1/2}\xi\big)\mathcal{L}^1(\mathrm{d}\xi)\right| \\
&\quad\le C\int_{\{N^{-1/2}|\xi|\le\delta\}} \big(\xi^2+\xi^4\big)\exp\big(-\xi^2/C\big)\mathcal{L}^1(\mathrm{d}\xi) \le C.
\end{aligned}$$

Finally, consider (135). It will be convenient to introduce h_3 via

$$h_2(m,\hat{\xi}) = h_2(m,0) + \hat{\xi} h_3(m,\hat{\xi}), \tag{146}$$

which, according to Taylor and Lemma 40(i), satisfies

$$\sup_{|\hat{\xi}|\le\delta} \left|h_3(m,\hat{\xi})\right| \le \sup_{|\hat{\xi}|\le\delta} \left|\frac{\partial h_2}{\partial \xi}(m,\hat{\xi})\right| \le C. \tag{147}$$

By the definition of h_3 in (146), we have

$$\begin{aligned}
&\exp\bigl(-\xi^2 h_2\bigl(m, N^{-1/2}\xi\bigr)\bigr) - \exp\bigl(-\xi^2 h_2(m,0)\bigr) \\
&\quad = \exp\bigl(-\xi^2 h_2(m,0)\bigr)\bigl(\exp\bigl(-N^{-1/2}\xi^3 h_3\bigl(m, N^{-1/2}\xi\bigr)\bigr) - 1\bigr) \\
&\quad \overset{(142)}{=} \exp\bigl(-\xi^2 \operatorname{Var}(\mu_m)\bigr)\bigl(\exp\bigl(-N^{-1/2}\xi^3 h_3\bigl(m, N^{-1/2}\xi\bigr)\bigr) - 1\bigr).
\end{aligned} \tag{148}$$

We use the fact:

$$\left|\exp(z) - 1\right| = \left|\sum_{j=1}^{\infty} \frac{z^j}{j!}\right| \le \sum_{j=1}^{\infty} \frac{|z|^j}{j!} = \exp\bigl(|z|\bigr) - 1,$$

with

$$z = -N^{-1/2}\xi^3 h_3(m,\hat{\xi})$$

to conclude from (148) that

$$\begin{aligned}
&\left|\exp\bigl(-\xi^2 h_2\bigl(m, N^{-1/2}\xi\bigr)\bigr) - \exp\bigl(-\xi^2 h_2(m,0)\bigr)\right| \\
&\quad \le \exp\bigl(-\xi^2 \operatorname{Var}(\mu_m)\bigr)\bigl(\exp\bigl(N^{1/2}|\xi|^3 \bigl|h_3\bigl(m, N^{-1/2}\xi\bigr)\bigr|\bigr) - 1\bigr).
\end{aligned}$$

Together with (140) and (147), this yields for ξ with $N^{-1/2}|\xi| \le \delta$:

$$\left|\exp\bigl(-\xi^2 h_2\bigl(m, N^{-1/2}\xi\bigr)\bigr) - \exp\bigl(-\xi^2 h_2(m,0)\bigr)\right| \le \exp\bigl(-\xi^2/C\bigr)\bigl(\exp\bigl(C\delta\xi^2\bigr) - 1\bigr).$$

Hence, for δ sufficiently small,

$$\begin{aligned}
&\left|\int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\bigl(-\xi^2 h_2\bigl(m, N^{-1/2}\xi\bigr)\bigr) - \exp\bigl(-\xi^2 h_2(m,0)\bigr)\mathcal{L}^1(\mathrm{d}\xi)\right| \\
&\quad \le \int_{\mathbb{R}} \exp\bigl(-\xi^2(1/C - C\delta)\bigr) - \exp\bigl(-\xi^2/C\bigr)\mathcal{L}^1(\mathrm{d}\xi) \\
&\quad = C\left(\frac{1}{\sqrt{1/C - C\delta}} - \frac{1}{\sqrt{1/C}}\right) \\
&\quad \le C\delta.
\end{aligned} \tag{149}$$

On the other hand, we have by (142) and (140) that

$$\exp\bigl(-\xi^2 h_2(m,0)\bigr) = \exp\bigl(-\xi^2 \operatorname{Var}(\mu_m)\bigr) \ge \exp\bigl(-\xi^2/C\bigr),$$

so that

$$\begin{aligned}
\int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\bigl(-\xi^2 h_2(m,0)\bigr)\mathcal{L}^1(\mathrm{d}\xi) &\ge \int_{\{N^{-1/2}|\xi|\le\delta\}} \exp\bigl(-\xi^2/C\bigr)\mathcal{L}^1(\mathrm{d}\xi) \\
&\ge 1/C - C\exp\bigl(-N^{1/2}\delta/C\bigr).
\end{aligned} \tag{150}$$

The combination of (149) and (150) yields

$$\operatorname{Re}\int_{\{N^{-1/2}|\xi|\le\delta\}}\exp\bigl(-\xi^2 h_2\bigl(m,N^{-1/2}\xi\bigr)\bigr)\mathcal{L}^1(\mathrm{d}\xi)\ge 1/C - C\bigl(\exp\bigl(-CN^{1/2}\delta\bigr)+\delta\bigr),$$

which establishes (135) for δ sufficiently small and N sufficiently large. □

A.2. *Proofs of lemmas*

Before turning to the proofs, we collect a few ingredients that we will use repeatedly. First, recall that by assumption, μ_m is a perturbation of a shifted Gaussian. To be precise, letting

$$\frac{\mathrm{d}g_\sigma}{\mathrm{d}\mathcal{L}^1}(x):=(2\pi)^{-1/2}\exp\biggl(-\frac{1}{2}(x-\sigma)^2\biggr), \tag{151}$$

we may write μ_m as

$$\frac{\mathrm{d}\mu_m}{\mathrm{d}\mathcal{L}^1}(x)=\frac{1}{Z}\exp\biggl(-\delta\psi(x)+\sigma x-\frac{1}{2}x^2\biggr)=\frac{1}{\tilde{Z}}\exp\biggl(-\delta\psi(x)-\frac{1}{2}(x-\sigma)^2\biggr),$$

and observe that

$$\exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\frac{\mathrm{d}g_\sigma}{\mathrm{d}\mathcal{L}^1}(x)\le\frac{\mathrm{d}\mu_m}{\mathrm{d}\mathcal{L}^1}(x)\le\exp(\operatorname{osc}_{\mathbb{R}}\delta\psi)\frac{\mathrm{d}g_\sigma}{\mathrm{d}\mathcal{L}^1}(x). \tag{152}$$

A second elementary but important observation is that the mean of a measure μ is optimal in the sense that for all $c\in\mathbb{R}$,

$$\begin{aligned}\int_{\mathbb{R}}(x-c)^2\mu(\mathrm{d}x)&=\int_{\mathbb{R}}x^2\mu(\mathrm{d}x)-2c\int_{\mathbb{R}}x\mu(\mathrm{d}x)+c^2\\&\ge\int_{\mathbb{R}}x^2\mu(\mathrm{d}x)-\biggl(\int_{\mathbb{R}}x\mu(\mathrm{d}x)\biggr)^2\\&=\int_{\mathbb{R}}\biggl(x-\int_{\mathbb{R}}y\mu(\mathrm{d}y)\biggr)^2\mu(\mathrm{d}x).\end{aligned} \tag{153}$$

Finally, we state and prove a lemma about the map between m and σ that is useful in the proofs of Lemmas 39 and 40. Here and below, we refer to the measure μ_m as μ^σ in order to emphasize the σ-dependence. The lemma will imply in particular that

$$\forall\sigma\in\mathbb{R},\quad\frac{\mathrm{d}^2\varphi^*}{\mathrm{d}\sigma^2}(\sigma)\ge\frac{1}{C}>0. \tag{154}$$

Lemma 41. *Consider the change of variables*

$$m=\frac{\mathrm{d}\varphi^*}{\mathrm{d}\sigma}(\sigma) \tag{155}$$

and the corresponding measure $\mu_m=\mu^\sigma\in\mathcal{P}(\mathbb{R})$ *with density*

$$\frac{\mathrm{d}\mu^\sigma}{\mathrm{d}\mathcal{L}^1}(x)=\exp\bigl(-\varphi^*(\sigma)+\sigma x-\psi(x)\bigr)$$

and mean $\int_{\mathbb{R}}x\mu^\sigma(\mathrm{d}x)=m$, *cf.* (77), (129) *and* (74). *Then*:

(i) *The first two derivatives of* m *are related to the moments of* μ^σ *as*:

$$\frac{\mathrm{d}m}{\mathrm{d}\sigma} = \frac{\mathrm{d}^2\varphi^*}{\mathrm{d}\sigma^2} = \int_{\mathbb{R}} (x-m)^2 \mu^\sigma(\mathrm{d}x),$$
$$\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2} = \frac{\mathrm{d}^3\varphi^*}{\mathrm{d}\sigma^3} = \int_{\mathbb{R}} (x-m)^3 \mu^\sigma(\mathrm{d}x).$$

(ii) *The moments of* μ^σ *satisfy the uniform bounds*:

$$\frac{1}{C} \le \int_{\mathbb{R}} (x-m)^2 \mu^\sigma(\mathrm{d}x) \le C,$$
$$\left| \int_{\mathbb{R}} (x-m)^3 \mu^\sigma(\mathrm{d}x) \right| \le C,$$
$$\int_{\mathbb{R}} (x-m)^4 \mu^\sigma(\mathrm{d}x) \le C.$$

(iii) *The second derivatives of the inverse map are uniformly bounded*:

$$\left| \frac{\mathrm{d}^2\sigma}{\mathrm{d}m^2} \right| \le C.$$

(iv) *The map is uniformly close to the identity*: $|\sigma - m| \le C$.

Proof of Lemma 41.

To show the equalities in (i), we first notice that for the variance we have from (74) and (129)

$$\begin{aligned}\frac{\mathrm{d}m}{\mathrm{d}\sigma} &= \frac{\mathrm{d}}{\mathrm{d}\sigma} \int_{\mathbb{R}} x \exp\big(-\varphi^*(\sigma) + \sigma x - \psi(x)\big) \mathcal{L}^1(\mathrm{d}x) \\ &= \int_{\mathbb{R}} x(x-m)\mu^\sigma(\mathrm{d}x) = \int_{\mathbb{R}} x^2 \mu^\sigma(\mathrm{d}x) - m^2 \\ &= \int_{\mathbb{R}} (x-m)^2 \mu^\sigma(\mathrm{d}x).\end{aligned} \tag{156}$$

Together with (77), this establishes the first equality of (i). For the second equality, we take a derivative in (156) and notice that because μ^σ has mean m,

$$\begin{aligned}\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2} &\overset{(155),(129)}{=} \int_{\mathbb{R}} (x-m)^3 \mu^\sigma(\mathrm{d}x) - 2\frac{\mathrm{d}m}{\mathrm{d}\sigma} \int_{\mathbb{R}} (x-m)\mu^\sigma(\mathrm{d}x) \\ &= \int_{\mathbb{R}} (x-m)^3 \mu^\sigma(\mathrm{d}x).\end{aligned}$$

Next we prove point (iv), which follows from

$$\begin{aligned}|\sigma - m|^2 &= \left| \int_{\mathbb{R}} (\sigma - x)\mu^\sigma(\mathrm{d}x) \right|^2 \le \int_{\mathbb{R}} (\sigma - x)^2 \mu^\sigma(\mathrm{d}x) \\ &\overset{(152)}{\le} \exp(\operatorname{osc}_{\mathbb{R}} \delta\psi) \int_{\mathbb{R}} (\sigma - x)^2 g_\sigma(\mathrm{d}x) \\ &\le \exp(\operatorname{osc}_{\mathbb{R}} \delta\psi).\end{aligned}$$

Turning to the first estimate in (ii), we observe that on the one hand,

$$\begin{aligned}\int_{\mathbb{R}}(x-m)^2\mu^\sigma(\mathrm{d}x) &\overset{(152)}{\geq} \exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}(x-m)^2 g_\sigma(\mathrm{d}x)\\ &\overset{(153)}{\geq} \exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}\left(x-\int_{\mathbb{R}} y g_\sigma(\mathrm{d}y)\right)^2 g_\sigma(\mathrm{d}x)\\ &= \exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi).\end{aligned}$$

On the other hand,

$$\begin{aligned}\int_{\mathbb{R}}(x-m)^2\mu^\sigma(\mathrm{d}x) &\overset{(153)}{\leq} \int_{\mathbb{R}}\left(x-\int_{\mathbb{R}} y g_\sigma(\mathrm{d}y)\right)^2 \mu^\sigma(\mathrm{d}x)\\ &\overset{(152)}{\leq} \exp(\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}\left(x-\int_{\mathbb{R}} y g_\sigma(\mathrm{d}y)\right)^2 g_\sigma(\mathrm{d}x)\\ &= \exp(\operatorname{osc}_{\mathbb{R}}\delta\psi).\end{aligned}$$

The bound in (ii) on the fourth moment follows from:

$$\begin{aligned}\int_{\mathbb{R}}(x-m)^4\mu^\sigma(\mathrm{d}x) &\leq C\left(\int_{\mathbb{R}}(x-\sigma)^4\mu^\sigma(\mathrm{d}x)+\int_{\mathbb{R}}(\sigma-m)^4\mu^\sigma(\mathrm{d}x)\right)\\ &\overset{(152)}{\leq} C\left(\exp(\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}(x-\sigma)^4 g_\sigma(\mathrm{d}x)+\int_{\mathbb{R}}(\sigma-m)^4\mu^\sigma(\mathrm{d}x)\right)\\ &\leq C,\end{aligned}$$

by (iv) and the definition (151) of g_σ. Hölder's inequality then implies the bound on the third moment.

Finally, (iii) follows immediately from (i), (ii), and:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}m^2}=\frac{\mathrm{d}}{\mathrm{d}\sigma}\left(\frac{\mathrm{d}^2\varphi^*}{\mathrm{d}\sigma^2}\right)^{-1}\frac{\mathrm{d}\sigma}{\mathrm{d}m}=-\frac{\mathrm{d}^3\varphi^*}{\mathrm{d}\sigma^3}\left(\frac{\mathrm{d}^2\varphi^*}{\mathrm{d}\sigma^2}\right)^{-3}.$$

□

Proof of Lemma 39. We prove (i) by splitting it into two pieces. First we bound h by a constant smaller than one, uniformly in m for ξ bounded away from zero. Then we show the decay for large ξ. By (131) and (78) we have

$$h(m,\xi)=\exp(-\mathrm{i}\xi m)\int_{\mathbb{R}}\exp(\mathrm{i}\xi x)\mu^\sigma(\mathrm{d}x).$$

Thus

$$\begin{aligned}\left|h(m,\xi)\right|^2 &= \left|\int_{\mathbb{R}}\cos(\xi x)\mu^\sigma(\mathrm{d}x)+\mathrm{i}\int_{\mathbb{R}}\sin(\xi x)\mu^\sigma(\mathrm{d}x)\right|^2\\ &= \left(\int_{\mathbb{R}}\cos(\xi x)\mu^\sigma(\mathrm{d}x)\right)^2+\left(\int_{\mathbb{R}}\sin(\xi x)\mu^\sigma(\mathrm{d}x)\right)^2\\ &= 1-\left(\int_{\mathbb{R}}\cos^2(\xi x)\mu^\sigma(\mathrm{d}x)-\left(\int_{\mathbb{R}}\cos(\xi x)\mu^\sigma(\mathrm{d}x)\right)^2\right)\\ &\quad -\left(\int_{\mathbb{R}}\sin^2(\xi x)\mu^\sigma(\mathrm{d}x)-\left(\int_{\mathbb{R}}\sin(\xi x)\mu^\sigma(\mathrm{d}x)\right)^2\right)\\ &=: 1-\operatorname{Var}_{\mu^\sigma}\big(\cos(\xi x)\big)-\operatorname{Var}_{\mu^\sigma}\big(\sin(\xi x)\big). \qquad (157)\end{aligned}$$

Therefore, to bound h by a constant smaller than one, we need to bound the variances away from zero. Recalling the elementary observations (152) and (153), we have:

$$\begin{aligned}
&\operatorname{Var}_{\mu^\sigma}\big(\cos(\xi x)\big)\\
&\quad=\int_{\mathbb{R}}\bigg(\cos(\xi x)-\int_{\mathbb{R}}\cos(\xi y)\mu^\sigma(\mathrm{d}y)\bigg)^2\mu^\sigma(\mathrm{d}x)\\
&\quad\geq\exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\int_{\mathbb{R}}\bigg(\cos(\xi x)-\int_{\mathbb{R}}\cos(\xi y)\mu^\sigma(\mathrm{d}y)\bigg)^2 g_\sigma(\mathrm{d}x)\\
&\quad\geq\exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\bigg(\int_{\mathbb{R}}\cos^2(\xi x)g_\sigma(\mathrm{d}x)-\bigg(\int_{\mathbb{R}}\cos(\xi x)g_\sigma(\mathrm{d}x)\bigg)^2\bigg).
\end{aligned}\tag{158}$$

Since the Fourier transform of a Gaussian is again Gaussian, the right-hand side of (158) can be computed explicitly. Looking at the second integral, we have

$$\begin{aligned}
&\bigg(\int_{\mathbb{R}}\cos(\xi x)g_\sigma(\mathrm{d}x)\bigg)^2\\
&\quad=\frac{1}{4}\bigg((2\pi)^{-1/2}\int_{\mathbb{R}}\big(\exp(\mathrm{i}\xi x)+\exp(-\mathrm{i}\xi x)\big)\exp\Big(-\frac{1}{2}(x-\sigma)^2\Big)\mathcal{L}^1(\mathrm{d}x)\bigg)^2\\
&\quad=\frac{1}{4}\bigg(\exp(\mathrm{i}\xi\sigma)(2\pi)^{-1/2}\int_{\mathbb{R}}\exp(\mathrm{i}\xi y)\exp\Big(-\frac{1}{2}y^2\Big)\mathcal{L}^1(\mathrm{d}y)\\
&\qquad+\exp(-\mathrm{i}\xi\sigma)(2\pi)^{-1/2}\int_{\mathbb{R}}\exp(-\mathrm{i}\xi y)\exp\Big(-\frac{1}{2}y^2\Big)\mathcal{L}^1(\mathrm{d}y)\bigg)^2\\
&\quad=\frac{1}{4}\bigg(\exp(\mathrm{i}\xi\sigma)\exp\Big(-\frac{1}{2}\xi^2\Big)+\exp(-\mathrm{i}\xi\sigma)\exp\Big(-\frac{1}{2}\xi^2\Big)\bigg)^2\\
&\quad=\frac{1}{4}\big(\exp(2\mathrm{i}\xi\sigma)\exp\big(-\xi^2\big)+\exp(-2\mathrm{i}\xi\sigma)\exp\big(-\xi^2\big)+2\exp\big(-\xi^2\big)\big)\\
&\quad=\frac{1}{2}\big(\cos(2\xi\sigma)+1\big)\exp\big(-\xi^2\big).
\end{aligned}$$

The second part of the right-hand side of (158) can be computed similarly. We get:

$$\int_{\mathbb{R}}\cos^2(\xi x)g_\sigma(\mathrm{d}x)=\frac{1}{2}\big(\cos(2\xi\sigma)\exp\big(-2\xi^2\big)+1\big).$$

Therefore,

$$\begin{aligned}
&\int_{\mathbb{R}}\cos^2(\xi x)g_\sigma(\mathrm{d}x)-\bigg(\int_{\mathbb{R}}\cos(\xi x)g_\sigma(\mathrm{d}x)\bigg)^2\\
&\quad=\frac{1}{2}\big(1-\exp\big(-\xi^2\big)\cos(2\xi\sigma)\big)\big(1-\exp\big(-\xi^2\big)\big)\\
&\quad\geq\frac{1}{2}\big(1-\exp\big(-\xi^2\big)\big)^2.
\end{aligned}$$

Inserting this into (158) and then (157) (the same inequality holds for $\operatorname{Var}_{\mu^\sigma}(\sin(\xi x))$), we obtain

$$\big|h(m,\xi)\big|^2\leq 1-\exp(-\operatorname{osc}_{\mathbb{R}}\delta\psi)\big(1-\exp\big(-\xi^2\big)\big)^2.$$

Hence for any $\delta > 0$ there exists a $C_\delta < \infty$ (uniform in m) such that

$$\big|h(m,\xi)\big| \le 1 - \frac{1}{C_\delta} \quad \text{for } |\xi| > \delta. \tag{159}$$

To complete the proof of Lemma 39(i), we need to establish decay of h for large $|\xi|$-values. This is done by an integration by parts argument; we have

$$\begin{aligned} h(m,\xi) &= \frac{1}{Z}\int_{\mathbb{R}} \exp(\mathrm{i}\xi x)\exp\big(\sigma x - \psi(x)\big)\mathcal{L}^1(\mathrm{d}x) \\ &= \frac{1}{Z}\int_{\mathbb{R}} \frac{1}{\mathrm{i}\xi}\exp(\mathrm{i}\xi x)\left(\sigma - x - \frac{\mathrm{d}\delta\psi}{\mathrm{d}x}(x)\right)\exp\left(\sigma x - \frac{1}{2}x^2 - \delta\psi(x)\right)\mathcal{L}^1(\mathrm{d}x). \end{aligned}$$

This yields the estimate

$$\begin{aligned} &\big|h(m,\xi)\big| \\ &\quad \le \frac{1}{|\xi|}\int_{\mathbb{R}}\left(|\sigma - x| + \left|\frac{\mathrm{d}\delta\psi}{\mathrm{d}x}(x)\right|\right)\mu^\sigma(\mathrm{d}x) \\ &\quad \overset{(152)}{\le} \frac{1}{|\xi|}\exp(\mathrm{osc}_{\mathbb{R}}\,\delta\psi)\int_{\mathbb{R}}\left(|\sigma - x| + \left|\frac{\mathrm{d}\delta\psi}{\mathrm{d}x}(x)\right|\right)g_\sigma(\mathrm{d}x) \\ &\quad \le \frac{1}{|\xi|}\exp(\mathrm{osc}_{\mathbb{R}}\,\delta\psi)\left((2\pi)^{-1/2}\int_{\mathbb{R}}|y|\exp\left(-\frac{1}{2}y^2\right)\mathcal{L}^1(\mathrm{d}y) + \sup_{\mathbb{R}}\left|\frac{\mathrm{d}\delta\psi}{\mathrm{d}x}(x)\right|\right). \end{aligned} \tag{160}$$

Since by elementary interpolation

$$\sup_{\mathbb{R}}\left|\frac{\mathrm{d}\delta\psi}{\mathrm{d}x}(x)\right| \le C\sup_{\mathbb{R}}\big|\delta\psi(x)\big|\sup_{\mathbb{R}}\left|\frac{\mathrm{d}^2\delta\psi}{\mathrm{d}x^2}(x)\right| \overset{(16)}{<} \infty,$$

we infer from (160) that

$$\big|h(m,\xi)\big| \le \frac{C}{|\xi|}. \tag{161}$$

The combination of (159) and (161) yields Lemma 39(i).

We turn now to the estimates for (ii) and (iii). Since $|\xi| \le \delta$, it suffices to prove

$$\begin{aligned} \left|\frac{\partial h}{\partial m}\right| &\le C\big(1 + |\xi|\big), \\ \left|\frac{\partial^2 h}{\partial m^2}\right| &\le C\big(1 + |\xi|^2\big). \end{aligned}$$

We appeal to the change of variables (153):

$$\frac{\partial h}{\partial m} = \frac{\partial h}{\partial\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}m} \quad \text{and} \quad \frac{\partial^2 h}{\partial m^2} = \frac{\partial^2 h}{\partial\sigma^2}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}m}\right)^2 + \frac{\partial h}{\partial\sigma}\frac{\mathrm{d}^2\sigma}{\mathrm{d}m^2}. \tag{162}$$

According to Lemma 41(i)–(iii), it thus suffices to prove

$$\begin{aligned} \left|\frac{\partial h}{\partial\sigma}\right| &\le C\big(1 + |\xi|\big), \\ \left|\frac{\partial^2 h}{\partial\sigma^2}\right| &\le C\big(1 + |\xi|^2\big). \end{aligned} \tag{163}$$

The starting point is formula (131):

$$h = \int_{\mathbb{R}} \exp\bigl(\mathrm{i}\xi(x-m)\bigr)\mu^{\sigma}(\mathrm{d}x) = \int_{\mathbb{R}} \exp\bigl(\mathrm{i}\xi x - \mathrm{i}\xi m - \varphi^{*}(\sigma) + \sigma x - \psi(x)\bigr)\mathcal{L}^{1}(\mathrm{d}x). \tag{164}$$

Using (155), we infer the identities

$$\frac{\partial h}{\partial \sigma} = \int_{\mathbb{R}} \left(-\mathrm{i}\xi \frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right) \exp\bigl(\mathrm{i}\xi(x-m)\bigr)\mu^{\sigma}(\mathrm{d}x), \tag{165}$$

$$\begin{aligned}\frac{\partial^2 h}{\partial \sigma^2} &= \int_{\mathbb{R}} \left(-\mathrm{i}\xi \frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2} - \frac{\mathrm{d}m}{\mathrm{d}\sigma}\right) \exp\bigl(\mathrm{i}\xi(x-m)\bigr)\mu^{\sigma}(\mathrm{d}x) \\ &\quad + \int_{\mathbb{R}} \left(-\mathrm{i}\xi \frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right)^2 \exp\bigl(\mathrm{i}\xi(x-m)\bigr)\mu^{\sigma}(\mathrm{d}x). \end{aligned} \tag{166}$$

By Jensen's inequality and (74), these yield the inequalities

$$\begin{aligned}\left|\frac{\partial h}{\partial \sigma}\right|^2 &\le \int_{\mathbb{R}} \left| - \mathrm{i}\xi \frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right|^2 \mu^{\sigma}(\mathrm{d}x) \\ &= \xi^2 \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right|^2 + \int_{\mathbb{R}} (x-m)^2 \mu^{\sigma}(\mathrm{d}x)\end{aligned}$$

and

$$\begin{aligned}\left|\frac{\partial^2 h}{\partial \sigma^2}\right| &\le |\xi| \left|\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2}\right| + \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right| + \int_{\mathbb{R}} \left| - \mathrm{i}\xi \frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right|^2 \mu^{\sigma}(\mathrm{d}x) \\ &= |\xi| \left|\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2}\right| + \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right| + \xi^2 \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right|^2 + \int_{\mathbb{R}} (x-m)^2 \mu^{\sigma}(\mathrm{d}x).\end{aligned}$$

Hence (163) follows from Lemma 41(i) and (ii). □

Proof of Lemma 40. Since $h(m,0) = 1$ and $\frac{\partial h}{\partial \xi}(m,0) = 0$, cf. (139), we may introduce by Taylor the complex-valued function $h_1(m,\xi)$ by

$$h(m,\xi) = 1 - \xi^2 h_1(m,\xi) = \exp\bigl(-\xi^2 h_2(m,\xi)\bigr), \tag{167}$$

so that

$$h_2(m,\xi) = \begin{cases} -\xi^{-2}\log\bigl(1 - \xi^2 h_1(m,\xi)\bigr), & \xi \neq 0, \\ h_1(m,0), & \xi = 0. \end{cases} \tag{168}$$

We claim that Lemma 40 is a consequence of the following bounds on h_1:

$$\bigl|h_1(m,\xi)\bigr| \le C, \qquad \left|\frac{\partial h_1}{\partial \xi}\right| \le C, \tag{169}$$

$$\left|\frac{\partial h_1}{\partial m}\right| \le C, \qquad \left|\frac{\partial^2 h_1}{\partial m^2}\right| \le C. \tag{170}$$

Indeed, Lemma 12(i) follows from (169) after rewriting (168) in the form:

$$h_2(m,\xi) = h_1(m,\xi) f\bigl(\xi^2 h_1(m,\xi)\bigr),$$

for the function $f(z) = -z^{-1}\log(1-z)$, which is smooth in a neighborhood of zero. Points (ii) and (iii) follow from

(170) via the chain rule applied to (168):

$$\frac{\partial h_2}{\partial m} = \frac{1}{1-\xi^2 h_1}\frac{\partial h_1}{\partial m}$$

and

$$\frac{\partial^2 h_2}{\partial m^2} = \frac{1}{1-\xi^2 h_1}\frac{\partial^2 h_1}{\partial m^2} + \frac{\xi^2}{(1-\xi^2 h_1)^2}\left(\frac{\partial h_1}{\partial m}\right)^2,$$

along with the fact that $|\xi| \le \delta$. As in the proof of Lemma 39, it will be convenient to consider derivatives with respect to σ instead of m. By (162) and Lemma 41(i)–(iii), we can establish (170) by showing

$$\left|\frac{\partial h_1}{\partial \sigma}\right| \le C\big(1+|\xi|\big) \quad \text{and} \quad \left|\frac{\partial^2 h_1}{\partial \sigma^2}\right| \le C\big(1+|\xi|^2\big). \tag{171}$$

In view of definition (167), which can be reformulated as

$$h_1(m,\xi) = \frac{1}{\xi^2}\int_0^{\xi}(\xi'-\xi)\frac{\partial^2 h}{\partial \xi^2}(m,\xi')\,\mathrm{d}\xi',$$

(169) and (171) are consequences of

$$\left|\frac{\partial^2 h}{\partial \xi^2}\right| \le C, \qquad \left|\frac{\partial^3 h}{\partial \xi^3}\right| \le C, \tag{172}$$

$$\left|\frac{\partial^3 h}{\partial \xi^2\,\partial \sigma}\right| \le C, \qquad \left|\frac{\partial^4 h}{\partial \xi^2\,\partial^2 \sigma}\right| \le C. \tag{173}$$

The estimates (172) are easily established. We infer from (164)

$$\frac{\partial^k h}{\partial \xi^k}(m,\xi) = \int_{\mathbb{R}}\big(\mathrm{i}(x-m)\big)^k \exp\big(\mathrm{i}\xi(x-m)\big)\mu^{\sigma}(\mathrm{d}x).$$

Thus, (172) follows from Lemma 41(ii).

For (173) we turn to (165) and (166), which we write as

$$\frac{\partial h}{\partial \sigma} = \int_{\mathbb{R}} a_1(\sigma,\xi)\mu^{\sigma}(\mathrm{d}x) \quad \text{and} \quad \frac{\partial^2 h}{\partial \sigma^2} = \int_{\mathbb{R}} a_2(\sigma,\xi)\mu^{\sigma}(\mathrm{d}x),$$

where we set for abbreviation:

$$a_1(\sigma,\xi) = \left(-\mathrm{i}\xi\frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right)\exp\big(\mathrm{i}\xi(x-m)\big),$$

$$a_2(\sigma,\xi) = \left(\left(-\mathrm{i}\xi\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2} - \frac{\mathrm{d}m}{\mathrm{d}\sigma}\right) + \left(-\mathrm{i}\xi\frac{\mathrm{d}m}{\mathrm{d}\sigma} + x - m\right)^2\right)\exp\big(\mathrm{i}\xi(x-m)\big).$$

Since for $|\xi| \le \delta$ we have

$$\left|\frac{\partial^2 a_1}{\partial \xi^2}\right| \le C\left(\left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right| + 1\right)\big(|x-m|^3 + 1\big)$$

and

$$\left|\frac{\partial^2 a_2}{\partial \xi^2}\right| \le C\left(\left|\frac{\mathrm{d}^2 m}{\mathrm{d}\sigma^2}\right| + \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right|^2 + \left|\frac{\mathrm{d}m}{\mathrm{d}\sigma}\right| + 1\right)\big(|x-m|^4 + 1\big),$$

(173) follows from Lemma 41(i) and (ii). □

Acknowledgement

We thank the referee for a careful reading and useful comments. C.V. thanks Elena Kosygina for kindly giving him a primer in the probabilistic study of hydrodynamic limits, and providing a preliminary version of [13]. This was the starting point for our first attempts to solve this problem, back in 1999 as C.V. was enjoying the hospitality of the University of Santa Barbara.

References

[1] D. Bakry and M. Émery. Diffusions hypercontractives. In *Sem. Probab. XIX. Lecture Notes in Mathematics* **1123** 177–206. Springer, Berlin, 1985. MR0889476

[2] G. Blower and F. Bolley. Concentration of measure on product spaces with applications to Markov processes. *Studia Math.* **175** (2006) 47–72. MR2261699

[3] Th. Bodineau and B. Helffer. On log Sobolev inequalities for unbounded spin systems. *J. Funct. Anal.* **166** (1999) 168–178. MR1704666

[4] D. Chafaï. Glauber versus Kawasaki for spectral gap and logarithmic Sobolev inequalities of some unbounded conservative spin systems. *Markov Process. Related Fields* **9** (2001) 341–362. MR2028218

[5] P. Caputo. Uniform Poincaré inequalities for unbounded conservative spin systems: The non-interacting case. *Stochastic Process. Appl.* **106** (2003) 223–244. MR1989628

[6] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, 2nd edn. *Appl. Math.* **38**. Springer, New York, 1998. MR1619036

[7] W. Feller. *An Introduction to Probability Theory and Its Applications*, *Vol. 2*. Wiley, New York, 1971. MR0270403

[8] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** (1975) 1061–1083. MR0420249

[9] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Sminaire de Probabilits, XXXVI*. *Lecture Notes in Mathematics* **1801** 1–134. Springer, Berlin, 2003. MR1971582

[10] M. Z. Guo, G. C. Papanicolaou and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Commun. Math. Phys.* **118** (1988) 31–59. MR0954674

[11] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.* **46** (1987) 1159–1194. MR0893137

[12] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*. Springer, Berlin, 1999. MR1707314

[13] E. Kosygina. The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.* **29** (2001) 1086–1110. MR1872737

[14] C. Landim, G. Panizo and H. T. Yau. Spectral gap and logarithmic Sobolev inequality for unbounded conservative spin systems. *Ann. Inst. H. Poincaré Probab. Statist.* **38** (2002) 739–777. MR1931585

[15] M. Ledoux. *The Concentration of Measure Phenomenon, Math. Surveys and Monographs* **89**. Amer. Math. Soc., Providence, RI, 2001. MR1849347

[16] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisted. In *Sem. Probab. XXXV*. *Lecture Notes in Mathematics* **1755** 167–194. Springer, Berlin. 2001. MR1837286

[17] S. L. Lu. Hydrodynamic scaling with deterministic initial configurations. *Ann. Probab.* **23** (1995) 1831–1852. MR1379170

[18] S. L. Lu and H. T. Yau. Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Commun. Math. Phys.* **156** (1993) 399–433. MR1233852

[19] F. Otto and M. G. Reznikoff. A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.* **243** (2007) 121–157. MR2291434

[20] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** (2000) 361–400. MR1760620

[21] G. Royer. *Une Initiation aux Inégalités de Sobolev Logarithmiques*. *Cours Spécialisés*. Soc. Math. de France, Paris, 1999. MR1704288

[22] G. Toscani and C. Villani. On the trend to equilibrium for some dissipative systems with slowly increasing a-priori bounds. *J. Stat. Phys.* **98** (2000) 1279–1309. MR1751701

[23] C. Villani. *Topics in Optimal Transportation*. *Graduate Texts in Mathematics* **58**. Amer. Math. Soc., Providence, RI, 2003. MR1964483

[24] H.-T. Yau. Relative entropy and hydrodynamics of Ginzburg–Landau models. *Lett. Math. Phys.* **22** (1991) 63–80. MR1121850

[25] H.-T. Yau. Logarithmic Sobolev inequality for lattice gases with mixing conditions. *Commun. Math. Phys.* **181** (1996) 367–408. MR1414837