Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 192, 1–57.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1248

A quantitative hydrodynamic limit of the Kawasaki dynamics*

Deniz Dizdar[†] Georg Menz[‡] Felix Otto[§] Tianqi Wu[¶]

Abstract

We derive a rate of convergence to the hydrodynamic limit of the Kawasaki dynamics for a one-dimensional lattice spin system as considered by Guo, Papanicolaou and Varadhan. We follow the two-scale approach of Grunewald, Villani, Westdickenberg, and the middle author. However, we use a different coarse-graining operator that allows us to leverage the gradient flow structure. As a consequence, we obtain a better convergence rate.

Keywords: two-scale approach; logarithmic Sobolev inequality; spin system; Kawasaki dynamics; canonical ensemble; coarse-graining; splines; Galerkin approximation.
MSC2020 subject classifications: Primary 60K35, Secondary 60J25; 82B21.
Submitted to EJP on February 23, 2024, final version accepted on November 29, 2024.
Supersedes arXiv:1807.09850v2.

Contents

*This research has been partially supported by NSF grant DMS-1407558. Georg Menz and Tianqi Wu want to thank the Max-Planck Institute for Mathematics in the Sciences, Leipzig, Germany, for financial support.
[†]Université de Montréal, Canada. E-mail: deniz.dizdar@umontreal.ca
[‡]UCLA, USA. E-mail: menz@math.ucla.edu
[§]Max Planck Institute for Mathematics in the Sciences, Germany. E-mail: Felix.Otto@mis.mpg.de
[¶]Technion, Israel. E-mail: wtq@campus.technion.ac.il

Part I

Hydrodynamic limit via a two-scale approach

1 Introduction

The broader context of this work is the derivation of scaling limits for lattice systems. Typically, such a result shows that under a suitable time-space re-scaling, a random evolution of a lattice system converges to a macroscopic evolution as the system size goes to infinity. Two different kinds of limits may be considered. In the hydrodynamic limit (a dynamical version of the law of large numbers), the limiting macroscopic evolution is deterministic and describes the typical macroscopic behavior of the system. In the fluctuation limit (a dynamical version of the central limit theorem), the limiting macroscopic evolution is random and describes the fluctuations around the hydrodynamic limit.

This work is devoted to the hydrodynamic limit of the Kawasaki dynamics of one-dimensional lattice systems of continuous, unbounded spins. The Kawasaki dynamics is a spin-exchange dynamics preserving the mean spin. In the hydrodynamic limit, it converges to a non-linear diffusion equation. On a qualitative level, this convergence was established in [17] using resolvent techniques and in [21] using convergence of martingales and entropy estimates. Our quantitative approach is closer to the [21] method in the sense that we use thermodynamically natural quantities like the relative entropy and its dissipation, and allow for non-convex single site potentials. As an alternative to the martingale method in [21], Lu and Yau introduced the entropy method in [26], which is based on a sophisticated Gronwall-type estimate for a relative entropy functional. This method is more straightforward and gives stronger results, but also 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). All those results were qualitative, and it is not apparent how to make them quantitative.

In the present work we develop a quantitative theory of the hydrodynamic limit of the Kawasaki dynamics by establishing convergence rates. The first step toward a quantitative theory was made in [19] by introducing the two-scale approach. For a detailed description of the two-scale approach, we refer to Section 3. In a nutshell, the two-scale approach introduces an additional mesoscopic scale in-between the microscopic and macroscopic scales. The hydrodynamic limit is then deduced in two steps, first showing the closeness of the stochastic microscopic dynamics to a carefully chosen, deterministic mesoscopic dynamics, and then showing the closeness of that mesoscopic dynamics to the macroscopic dynamics. In [19], the hydrodynamic limit is still deduced only on a qualitative level, but the main estimate for the first step is already quantitative, and in principle the second step could also be made quantitative with some numerical analysis, which overall would lead to a quantitative result on the hydrodynamic limit.

However, rather than completing the approach of [19], we instead proceed by applying the two-scale approach with a different choice of the mesoscopic scale. The reason is that the choice of the mesoscopic scale in [19] would result in error terms with a worse scaling in the system size compared to ours (for details see Remark 3.16 and Remark 4.5 below). More precisely, [19] defines the mesoscopic observables by projection onto piecewise constant functions. Due to the lack of regularity, the mesoscopic dynamics has to be defined in an unnatural way, and consequently one has to use a mixed Galerkin procedure, which is not optimal. In the present work, we define the mesoscopic observables by projection onto splines. Because the splines are smooth, the mesoscopic dynamics can be defined more naturally as the Galerkin approximation of the macroscopic dynamics, leading to better error estimates compared to [19]. On the other hand, because splines do not have a localized basis, deducing the main ingredients of the two-scale approach becomes more subtle.

The second motivation behind improving the estimates of [19] is to develop a quantitative theory of the fluctuation limit, which states that the fluctuations of the Kawasaki dynamics converge to the solution of a stochastic diffusion equation. As with the hydrodynamic limit, the fluctuation limit of the Kawasaki dynamics is well understood on a qualitative level (see for example [38, 40, 8, 11]), but there is no quantitative result. A possible line of attack would be to use the two-scale approach. The estimates of [19] for the distance between the microscopic and mesoscopic dynamics are too weak when using the scaling of the fluctuation limit. Because our error terms scale better, our estimates are still meaningful under this scaling (cf. Theorem 3.17).

Another interesting question in this setting is the convergence of the microscopic entropy to the hydrodynamic entropy. Again, this question is well understood from a qualitative point of view (cf. [24, 14]). With the tools provided here, one could hope to make the approach of Fathi [14] quantitative.

1.1 Connection to gradient flows

Deducing the hydrodynamic limit is more accessible if both the microscopic and macroscopic dynamics come from gradient flows, i.e. the evolution of each dynamics reduces some kind of energy in the fastest possible way via some dissipation mechanism (see e.g. [3, 36] for more details, examples, and further references). The main idea is that Γ-convergence of the energy functionals, together with the convergence of the dissipation mechanisms in the proper sense, yields the convergence of the associated gradient flows (see e.g. [35, 37, 28]). This new perspective was applied, for example, in the recent works [15, 29].

Hence, finding the appropriate gradient flow structure for the microscopic and macroscopic dynamics is beneficial. This task is non-trivial because different gradient flow structures could give rise to the same evolution equation. For example, it was pointed out in [30] that the porous medium equation may be seen both as a H^{-1}-gradient flow of functions and as a Wasserstein gradient flow of number densities. Studying this question led to the recently highlighted insight that the appropriate gradient flow structure arises from the large deviation principle of the underlying microscopic process (see e.g. [1, 2, 13, 16]), as was implicitly known before (see e.g. line (1.5) in [9]).

Let us illustrate the importance of selecting the appropriate gradient flow structure with two examples. The first example is the hydrodynamic limit for interacting Brownian particles on the circle (see [39]). The second example is the hydrodynamic limit of the Kawasaki dynamics on a one-dimensional lattice spin system (see e.g. [21]), which is studied in the present work. The two examples appear quite similar in that they both yield a porous medium type equation in the hydrodynamic limit. However, they differ significantly in terms of the underlying microscopic model: an interacting particle system in the first example and a spin system in the second one. The differences become even more apparent when studying the associated gradient flow structures.

The first example, the hydrodynamic limit of the interacting Brownian particles, can be interpreted as a convergence of gradient flows in the following way.

- In [39], on the microscopic level N Brownian particles interact on a circle S. The positions X_i of the particles are given by a coupled system of SDEs with repulsive interaction. Because the evolution is reversible, one can interpret the associated forward Kolmogorov equation as a gradient flow for the relative entropy functional w.r.t. to the Gibbs equilibrium measure in the Wasserstein space of probability measures (see [23]). Here, the inner metric in the Wasserstein distance is given by the Euclidean distance on S^N.

- Because the Brownian particles are indistinguishable, one considers the empirical distribution of the particles, which is obtained by "forgetting" the labels of the particles. The unlabelling of the particles naturally pushes forward the inner metric (see Section 4 in [30]). For the former, the inner metric as described in the first bullet point describes the displacement of the particles (Lagrangian description). For the latter, the inner metric is the discrete Wasserstein distance, i.e. the minimal displacement of the particles required to transport one empirical distribution into another (Eulerian description). The microscopic dynamics relevant for the hydrodynamic limit is then the associated projected evolution of the empirical distributions of the particles.
- As a consequence, one should view the porous medium equation obtained in the hydrodynamic limit as a Wasserstein gradient flow, namely the gradient flow of the macroscopic free energy on the Wasserstein space $M_1(S)$ of number densities on S.

It is worth noting that there are two "levels" of Wasserstein metrics involved here. The "inner" Wasserstein metric is associated with the "movement mechanism" of the dynamics itself, i.e. transporting empirical distributions (for the microscopic dynamics) or transporting number densities (for the macroscopic dynamics). The "outer" Wasserstein metric is associated with the stochastic fluctuations of the dynamics and becomes degenerate, as it is the nature of the hydrodynamic limit to be deterministic. The main takeaway is that the dissipation mechanism for the macroscopic gradient flow is induced by the underlying "movement mechanism" of the microscopic dynamics (in this case, the Wasserstein distance on $M_1(S)$).

Let us now turn to the second example: the interpretation of the hydrodynamic limit of the Kawasaki dynamics as a convergence of gradient flows.

- On the microscopic level, the spin system consists of N real-valued spins located on the discrete one-dimensional torus $\{1, \ldots, N\}$. The associated Hamiltonian for the spin values only has single-site potentials and no interaction term (see (2.1) below). The evolution of the spin values is governed via a coupled system of SDEs, called the Kawasaki dynamics (see (2.4) below). This means that a site can only change its spin by distributing the difference to its neighbors. This spin-exchange mechanism is mediated through the matrix A in (2.4). As in the first example, the associated forward Kolmogorov equation (see equation (2.10) below) has a gradient flow structure given by the relative entropy w.r.t. to the Gibbs equilibrium measure in the Wasserstein space of probability measures.
- Because a site can only reduce its energy via spin-exchange, the appropriate choice for the inner metric is the inner product $x \cdot A^{-1} y$. Because A is a second-order difference operator, $x \cdot A^{-1} y$ corresponds to a discrete H^{-1} metric. This illustrates another main difference between both examples. The interaction is not mediated by the Hamiltonian but by the dissipation mechanism of the dynamics.
- As a consequence, the porous medium equation obtained in the hydrodynamic limit of the Kawasaki dynamics should be considered as a (now continuum) H^{-1}-gradient flow.

In [19], those insights were applied to study the hydrodynamic limit. However, it was not completely carried out. Because of technical reasons, the authors of [19] did not choose a mesoscopic evolution with the natural $H^{-1}-$ gradient flow structure, leading to suboptimal estimates. In this work, we capitalize more on the idea of using the appropriate gradient flow structure and choose a mesoscopic evolution with the natural $H^{-1}-$ gradient flow structure. While this makes our proof more involved compared to [19], it leads to estimates with an improved scaling in the systems size N.

1.2 Notations and conventions

- We use the letter C to denote a universal generic constant $0 < C < \infty$ that is independent of the dimension N of the underlying lattice.
- We denote with $a \lesssim b$ that $a \leq Cb$. We denote $a \simeq b$ if $a \lesssim b$ and $b \lesssim a$.
- We denote with $a \cdot b$ and $|\cdot|$ the standard Euclidean inner product and norm on $\mathbb{R}^N$.
- Let X be a Euclidean space and $f : X \to \mathbb{R}$. Then we denote with ∇f and $\operatorname{Hess} f$ the gradient and Hessian inherited from the Euclidean structure of X.
- We use dx as a shorthand for the Hausdorff or Lebesgue measure of appropriate dimension.
- $\fint_a^b$ denotes an averaged integral over the interval $[a, b]$.
- $|\cdot|_{H^1}$ denotes the homogeneous H^1 norm.
- $[M] := \{1, 2, \ldots, M\}$. When indexing over $[M]$, we use the convention $0 = M$.
- $L^2(\mathbb{T})$ denotes the L^2 functions on the torus $\mathbb{T} = [0, 1]$.
- $L^2_0(\mathbb{T})$ denotes the L^2 functions on the torus $\mathbb{T} = [0, 1]$ *with mean zero.*

2 Setting and main result

2.1 Microscopic dynamics

We start with describing the Kawasaki dynamics on the microscopic lattice $\{1, \ldots, N\}$.

Definition 2.1 (Microscopic Hamiltonian H)**.** *The Hamiltonian $H : \mathbb{R}^N \to \mathbb{R}$ of the system is given by*

$$H(x) := H_N(x) = \sum_{n=1}^{N} \psi(x_n). \tag{2.1}$$

Here $\psi : \mathbb{R} \to \mathbb{R}$ is the single-site potential*, assumed to be of the form*

$$\psi(x) = \frac{1}{2} x^2 + ax + b + \delta\psi(x) \tag{2.2}$$

for some constants a, b and some function $\delta\psi$ that is bounded in $C^2(\mathbb{R})$, i.e.

$$\|\delta\psi\|_{L^\infty(\mathbb{R})} \leq C \text{ and } \|\delta\psi''\|_{L^\infty(\mathbb{R})} \leq C. \tag{2.3}$$

The function ψ may be non-convex and it helps to consider the case of a double-well potential (see Figure 1).

Definition 2.2 (Microscopic dynamics)**.** *The Kawasaki dynamics X_t is given by the solution of the SDE*

$$dX_t = -A\nabla H(X_t)dt + \sqrt{2A}dB_t. \tag{2.4}$$

Here B_t denotes a standard N-dimensional Brownian motion and $-A$ denotes the (centered) second-order difference operator for the periodic rescaled lattice $\left\{\frac{1}{N}, \ldots, 1\right\}$. More precisely, the operator A is given by the $N \times N$-matrix

$$A_{i,j} := N^2(-\delta_{i,j-1} + 2\delta_{i,j} - \delta_{i,j+1}). \tag{2.5}$$

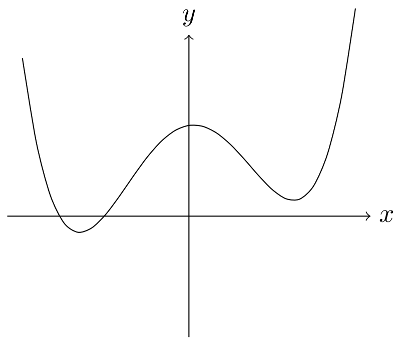

Figure 1: Double-well potential ψ.

It follows from the structure of the operator A that the Kawasaki dynamics (2.4) conserves the mean spin of the system. Hence, after a translation of the single-site potential ψ, we may restrict the state space $\mathbb{R}^N$ of the Kawasaki dynamics X_t to the hyperplane of zero mean

$$X_N := \left\{ x \in \mathbb{R}^N : \frac{1}{N} \sum_{i=1}^{N} x_i = 0 \right\}. \tag{2.6}$$

We endow the space X_N with the standard Euclidean inner product inherited from $\mathbb{R}^N$

$$\langle x, y \rangle_{X_N} := x \cdot y := \sum_{i=1}^{N} x_i y_i.$$

Additionally, the operator A is positive definite when restricted to X_N. Hence:

Definition 2.3 (Euclidean structures on X_N induced by A)**.** *The operator A induces a dual pair of inner products on the state space X_N, given by*

$$\langle x, y \rangle_A := x \cdot Ay, \text{ and } \langle x, y \rangle_{A^{-1}} := x \cdot A^{-1} y.$$

We denote by $|\cdot|_A$ and $|\cdot|_{A^{-1}}$ the corresponding norms on X_N.

The A/A^{-1}-Euclidean structures can be seen as a discrete version of the H^1/H^{-1} structures. In particular, we have the following well-known discrete analogue of the Poincaré inequality for functions with zero mean.

Lemma 2.4 (Discrete Poincaré inequality)**.** *There exists a constant $0 < C < \infty$ such that for all integers $N \geq 1$ and all $x \in X_N$,*

$$|x|^2 \leq CN^2 \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 \leq C|x|_A^2. \tag{2.7}$$

When the state space X_N is endowed with the A^{-1} inner product, the dynamics (2.4) can be written in the more suggestive form

$$dX_t = -\nabla_{A^{-1}} H(X_t) dt + \sqrt{2} dB_t^{A^{-1}}, \tag{2.8}$$

where $\nabla_{A^{-1}} := A\nabla$ denotes the gradient operation w.r.t. to the A^{-1} inner product and $B_t^{A^{-1}} := \sqrt{A} B_t$ denotes a Brownian motion on X_N having *identity covariance matrix* w.r.t. to the A^{-1} inner product.

As a standard result in the theory of stochastic processes (see for example [33]), the law of the process X_t at time t is characterized via the forward Kolmogorov equation.

Lemma 2.5 (Forward Kolmogorov equation)**.** *Assume that the law of initial condition X_0 is absolutely continuous w.r.t. the $N-1$ dimensional Hausdorff measure $\mathcal{L}^{N-1}$. Let μ denote the Gibbs measure on X_N associated to the Hamiltonian H, i.e. the measure μ is absolutely continuous w.r.t. the $N-1$-dimensional Hausdorff measure $\mathcal{L}^{N-1}$ with the Radon-Nikodym derivative given by*

$$\frac{d\mu}{d\mathcal{L}^{N-1}}(x) = \frac{1}{Z}\exp\left(-H(x)\right)\mathbb{1}_{x\in X_N}. \tag{2.9}$$

Then for all times $t>0$, the law $\rho(t)$ of the Kawasaki dynamics X_t (2.8) *is absolutely continuous w.r.t. the Gibbs measure μ, i.e. $\rho(t) = f(t)\mu$ for some $f(t)\in L^1(\mu)$, and is a weak solution of the Fokker-Planck equation*

$$\partial_t \rho_t = \nabla\cdot\left(\rho_t \nabla_{A^{-1}} H + \nabla_{A^{-1}}\rho_t\right) \tag{2.10}$$

in the sense that for any smooth test function $\xi : X_N \to \mathbb{R}$ it holds

$$\frac{d}{dt}\int \xi(x)\rho_t(dx) = \int -\nabla H\cdot\nabla_{A^{-1}}\xi\,\rho_t(dx) + \int \nabla\cdot\nabla_{A^{-1}}\xi\,\rho_t(dx).$$

In particular, the Gibbs measure μ is the unique stationary distribution of the Kawasaki dynamics (2.8)*.*

As a consequence of the forward Kolmogorov equation, the *relative entropy* of the law of X_t w.r.t. the Gibbs measure μ,

$$\mathrm{Ent}(\rho(t)|\mu) := \int f(t,x)\log f(t,x)\mu(dx),$$

decreases monotonically over time at the rate

$$\frac{d}{dt}\mathrm{Ent}\left(\rho(t)|\mu\right) = -\int \left|\nabla\log f(t,x)\right|_A^2\ \rho_t(dx). \tag{2.11}$$

The integral on the right hand side is the *Fisher information* for the Kawasaki dynamics, which differs from the standard Fisher information

$$\mathcal{I}_\mu(f(t)) := \int |\nabla\log f(t,x)|^2 f(t)\mu(dx)$$

only in the Euclidean structure being used. Hence, after we use the discrete Poincaré inequality in Lemma 2.4 to account for the different Euclidean structures on X_N, the rate of dissipation of the relative entropy is quantified by a *log-Sobolev inequality* (LSI) for the Gibbs measure μ by a standard Gronwall-type argument. In [19] it was shown that this rate of dissipation is independent of the system size N:

Proposition 2.6 (Uniform LSI for μ)**.** *The Gibbs measure μ given by* (2.9) *satisfies a LSI with constant $\hat\alpha>0$ uniform in the system size N. More precisely, for any nonnegative test function $g : X_N\to\mathbb{R}$ that satisfies $\int g(x)\mu(dx) = 1$, it holds that*

$$\mathrm{Ent}\left(g\mu|\mu\right) \le \frac{1}{2\hat\alpha}\mathcal{I}_\mu(g). \tag{2.12}$$

Remark 2.7 (Gradient flow structure of the microscopic dynamics)**.** The Fokker-Planck equation (2.10) can be written in the form

$$\partial_t\rho_t = \nabla\cdot\left(\rho_t\nabla_{A^{-1}}\frac{\delta E}{\delta\rho}(\rho_t)\right),$$

where $E(\rho)$ is the *microscopic free energy* of an ensemble $\rho = f\mu$,

$$E(\rho) := \int H d\rho + \int \rho \log \rho \, dx = \mathrm{Ent}(\rho|\mu),$$

and $\frac{\delta E}{\delta \rho} = H + \log \rho = \log f$ is its first variation. Consequently, on the level of probability densities on X_N, the Kawasaki dynamics X_t may be viewed as an A^{-1}-Wasserstein gradient flow for the convex energy functional $E(\cdot) = \mathrm{Ent}(\cdot|\mu)$, whose unique minimizer is the Gibbs measure μ. As expected for Wasserstein gradient flows, the energy functional E decreases over time at the rate

$$\frac{d}{dt} E(\rho_t) = -\int \left| \nabla_{A^{-1}} \frac{\delta E}{\delta \rho}(\rho) \right|_{A^{-1}}^2 d\rho_t.$$

We mention in passing the interesting fact that, using a gradient flow interpretation of the Fokker-Planck equation for the *Glauber dynamics*, it was shown in [32, Theorem 1] that the log-Sobolev inequality (2.12) implies

Corollary 2.8 (Talagrand's transportation inequality). *Let $\hat{\alpha}$ be the LSI constant in* (2.12) *for the Gibbs measure μ. Then for any probability measure ρ on X_N,*

$$\mathrm{Ent}\,(\rho|\mu) \geq \frac{\hat{\alpha}}{2} W_2^2(\rho, \mu). \tag{2.13}$$

Remark 2.9. The idea driving the proof of this fact in [32] is that the LSI bounds the energy functional $\mathrm{Ent}(\rho|\mu)$ from above by its Wasserstein gradient $I_\mu(f)$. By running the corresponding Wasserstein gradient flow on the initial datum ρ, this bound in turn implies the same energy functional is also bounded from below by the squared Wasserstein distance.

2.2 Hydrodynamic limit

The goal of the present work is to derive quantitative bounds on the hydrodynamic limit of the Kawasaki dynamics $X_t \in X_N$. Hydrodynamic limit means that as $N \to \infty$ the random dynamics X_t defined on the one-dimensional periodic lattice $\{1, 2, \cdots, N\}$ converges to a deterministic dynamics ζ_t on the one-dimensional torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Towards this end, we embed the spaces X_N into the space $L_0^2(\mathbb{T})$ of square-integrable functions of mean zero, by identifying the vector $x \in X_N$ with its corresponding step function on the interval $[0, 1]$.

Convention 1. *Given $x \in X_N$, we identify it with the step function*

$$x(\theta) = x_j, \qquad \theta \in \left[\frac{j-1}{N}; \frac{j}{N}\right).$$

Then the space X_N is identified with the space of piecewise constant functions on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with mean 0, i.e.

$$X_N = \left\{ x : \mathbb{T} \longrightarrow \mathbb{R}; \;\; x \text{ is constant on } \left[\frac{j-1}{N}; \frac{j}{N}\right), \;\; j = 1, .., N, \text{ and } \int_0^1 x(\theta) d\theta = 0 \right\}. \tag{2.14}$$

With this identification, $X_N \subset L_0^2(\mathbb{T})$ and inherits the L^2 inner product, which is related to the standard Euclidean inner product on X_N by a rescaling

$$\langle x, y \rangle_{L^2} = \frac{1}{N} x \cdot y.$$

It turns out the L^2 norm is not well-suited to describe the hydrodynamic limit since it is too sensitive to local fluctuations. Therefore we endow the embedded space $X_N \subset L_0^2(\mathbb{T})$ with the weaker homogeneous H^{-1}-norm, which is natural in light of the alternative form of the Kawasaki dynamics in (2.8) and the analogy between A^{-1}-norm and H^{-1}-norm.

Definition 2.10 (H^{-1}-norm)**.** *If $f : \mathbb{T} \to \mathbb{R}$ is a locally integrable function with mean zero, then*

$$||f||_{H^{-1}}^2 := \int_{\mathbb{T}} w(\theta)^2 d\theta, \qquad w' = f, \quad \int_{\mathbb{T}} w(\theta) d\theta = 0.$$

We now describe the limiting macroscopic dynamics ζ_t.

Definition 2.11 (Macroscopic free energy)**.** *The macroscopic free energy $\mathcal{H} : L^2(\mathbb{T}) \to \mathbb{R}$ is given by*

$$\mathcal{H}(\zeta) = \int_{\mathbb{T}} \varphi(\zeta(\theta)) d\theta,$$

where the function $\varphi : \mathbb{R} \to \mathbb{R}$ is the Cramér transform *of the single-site potential ψ, given by*

$$\varphi(m) = \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \log \int_{\mathbb{R}} \exp\left(\sigma z - \psi(z)\right) dz \right).$$

Accordingly, $\nabla \mathcal{H}(\zeta) = \varphi'(\zeta)$ in the variational sense:

$$\frac{d}{d\varepsilon} \mathcal{H}(\zeta + \varepsilon \xi) = \langle \varphi'(\zeta), \xi \rangle_{L^2} \quad \text{for any } \xi \in L^2(\mathbb{T}). \tag{2.15}$$

In particular, the macroscopic free energy $\mathcal{H}$ is convex. Indeed, the integrand $\varphi(m)$ is defined as the Legendre transform of the smooth function $\psi^* : \mathbb{R} \to \mathbb{R}$ given by

$$\psi^*(\sigma) := \log \int_{\mathbb{R}} \exp\left(\sigma z - \psi(z)\right) dz, \tag{2.16}$$

which is the *log partition function* associated to the linearly shifted potential $\psi(z) - \sigma z$. It turns out that the perturbed quadratic form (2.2) of ψ implies that ψ^* is strongly convex and bounded in C^2 (see e.g. Lemma 8.5 below or [19, Lemma 41]). These properties are then transferred to the conjugate function φ by the Legendre transform:

Lemma 2.12. *There exists constants $0 < \lambda < \Lambda < \infty$ such that*

$$0 < \lambda \le \varphi''(\theta) \le \Lambda < \infty \quad \textit{for all } \theta \in \mathbb{R}. \tag{2.17}$$

Up to a change in the linear term in the potential ψ, we may assume $(\psi^*)'(0) = 0$ and therefore $\varphi'(0) = 0$. Up to a change in the constant term in ψ, we may also assume $\psi^*(0) = 0$. After applying the Legendre transform, this means we conveniently have that

Assumption 2.13. *The function φ satisfies $\varphi(0) = \varphi'(0) = 0$. Consequently, the macroscopic free energy $\mathcal{H}(\zeta)$ is minimized at $\zeta = 0$ with $\mathcal{H}(0) = 0$.*

Definition 2.14 (Macroscopic dynamics)**.** *The macroscopic dynamics $\zeta(t, \cdot)$ is the unique weak solution of the nonlinear heat equation*

$$\frac{\partial \zeta}{\partial t} = \frac{\partial^2}{\partial \theta^2} \varphi'(\zeta) \tag{2.18}$$

with initial condition $\zeta(0, \cdot) := \zeta_0$. The precise formulation is deferred to Definition 2.17 at the end of this section.

Remark 2.15 (Gradient flow structure of the macroscopic dynamics)**.** The nonlinear heat equation (2.18) can be written in the form

$$\partial_t \zeta = -\nabla_{H^{-1}} \mathcal{H}(\zeta),$$

where $\nabla_{H^{-1}}$ is the gradient mapping of the first variation $\frac{\delta \mathcal{H}}{\delta \zeta}$ w.r.t. to the H^{-1} inner product (rather than w.r.t. to the L^2 inner product as in the formulation of (2.15)). Consequently, the macroscopic dynamics may be viewed as a H^{-1}-gradient flow for the convex energy functional $\mathcal{H}$, which therefore monotonically decreases over time at the rate

$$\frac{d}{dt}\mathcal{H}(\zeta_t) = -|\nabla_{H^{-1}}\mathcal{H}(\zeta)(t)|^2_{H^{-1}} = -|\varphi'(\zeta_t)|^2_{H^1}. \tag{2.19}$$

Now, let us formulate the main result of this work.

Theorem 2.16 (Quantitative hydrodynamic limit for the Kawasaki dynamics)**.** *We assume that the single-site potential ψ satisfies* (2.2) *and* (2.3)*. Let μ denote the Gibbs measure given by* (2.9) *and let $\rho(t) = f(t)\mu$ denote the law of the Kawasaki dynamics X_t (cf. Lemma 2.5). We assume that the initial law $\rho(0) = f(0)\mu$ of X_0 has bounded microscopic entropy in the sense that for some constant $0 < C_{\mathrm{Ent}} < \infty$,*

$$\mathrm{Ent}(\rho(0)|\mu) := \int f(0,x)\log f(0,x)\mu(dx) \le C_{\mathrm{Ent}} N. \tag{2.20}$$

Let ζ_t be the deterministic dynamics described by equation (2.18)*. Then there is a constant $0 < C < \infty$ depending only on the constants appearing in* (2.3) *such that for any $T > 0$,*

$$\sup_{0\le t\le T} \mathbb{E}|X_t - \zeta_t|^2_{H^{-1}} \le C\,\mathbb{E}|X_0 - \zeta_0|^2_{H^{-1}} + \frac{C}{N^{\frac{2}{3}}}\Big[T + C_{\mathrm{Ent}} + |\zeta_0|^2_{L^2} + 1\Big]. \tag{2.21}$$

The statement of Theorem (2.16) is a quantitative version of the hydrodynamic limit. In [19], only the error from comparing the microscopic scale to a mesoscopic scale was explicit. That error scaled in [19] like $\frac{1}{\sqrt{N}}$.

We finish this section by giving the precise formulation of equation (2.18) that describes the limiting macroscopic dynamics.

Definition 2.17. *We call $\zeta(t,\theta)$ a weak solution of (2.18) on $[0,T]\times\mathbb{T}$ if*

$$\zeta \in L^\infty_t(L^2_\theta), \qquad \frac{\partial \zeta}{\partial t} \in L^2_t(H^{-1}_\theta), \qquad \varphi'(\zeta) \in L^\infty_t(L^2_\theta)$$

and

$$\left\langle \xi, \frac{\partial \zeta}{\partial t} \right\rangle_{H^{-1}} = -\left\langle \xi, \varphi'(\zeta)\right\rangle_{L^2} \qquad \text{for all } \xi \in L^2_0(\mathbb{T}), \quad \text{for a.e. } t\in[0,T]. \tag{2.22}$$

Here, $L^\infty_t(L^2_\theta)$ (resp. $L^2_t(H^{-1}_\theta)$) is the set of functions $\zeta : [0,T]\times\mathbb{T} \longrightarrow \mathbb{R}$ such that $\int_{\mathbb{T}} \zeta(t,\theta)d\theta = 0$ and $||\zeta(t,\cdot)||_{L^2}$ (resp. $||\zeta(t,\cdot)||_{H^{-1}}$) is essentially bounded in t (resp. in $L^2([0,T])$).

3 The two-scale approach

We will use the two-scale approach from [19] to deduce Theorem 2.16. The main idea is to introduce a *mesoscopic dynamics* on an intermediate scale between the microscopic dynamics (2.4) and the macroscopic dynamics (2.18). The hydrodynamic limit is then deduced in two steps:

- In the first step, one deduces the convergence of the microscopic dynamics to the mesoscopic dynamics (see Theorem 3.17 from below).
- In the second step, one deduces the convergence of the mesoscopic dynamics to the macroscopic dynamics (see Theorem 3.18 from below).

The most important ingredient in the two-scale approach is the correct definition of the mesoscopic dynamics. It emerges from projecting the microscopic observables onto mesoscopic observables using a coarse-graining operator P. Let us first explain how this was done in [19]. Recall that an element $x \in X_N$ is identified with a step function on the torus $\mathbb{T} = [0, 1]$ that is piecewise constant with value x_n on the intervals (cf. (2.14))

$$\left[\frac{n-1}{N}, \frac{n}{N}\right), \quad n = 1, 2, \ldots, N.$$

In [19], the coarse-graining operator P was chosen to be the projection of X_N in $L^2(\mathbb{T})$ onto the space of step functions that are piecewise constant on the intervals

$$\left[\frac{m-1}{M}, \frac{m}{M}\right), \quad m = 1, 2, \ldots, M$$

where $M \in \mathbb{N}$ is chosen to be smaller than N. More specifically, one decomposes the lattice $\{1, 2, \ldots, N\}$ into M-many blocks $B(m)$ of size K, i.e. $N = MK$, and

$$B(m) = \{m(K-1)+1, \ldots, mK\} \qquad \text{for } 1 \leq m \leq M.$$

The coarse-graining operator $P : X_N \to \mathbb{R}^M$ in [19] is then given by

$$P(x) = \left(\frac{1}{K}\sum_{i\in B(1)} x_i, \ldots, \frac{1}{K}\sum_{i\in B(M)} x_i\right).$$

The main difference of this work compared to [19] is that the operator P is now defined as the L^2 projection onto *splines* of degree 2 instead (see Definition 3.2 from below). Because spline functions of degree 2 are $C^1(\mathbb{T})$, the mesoscopic variables are more regular compared to [19]. This has two important advantages:

- In the first step of the two-scale approach, namely showing the convergence of the microscopic dynamics to the mesoscopic dynamics (see Theorem 3.17 below), we get a better error estimate compared to [19, Theorem 8].
- The second step of the two-scale approach, namely deducing the convergence of the mesoscopic dynamics to the macroscopic dynamics, becomes significantly easier (see Theorem 3.18 below). Instead of a mixed method we can apply a direct Galerkin approximation method.

However, there is a trade-off compared to the argument of [19]. For showing the convergence of the microscopic dynamics to the mesoscopic dynamics, one needs certain technical ingredients, among which are the uniform strong convexity of the coarse-grained Hamiltonian (Theorem 3.9 below) and a uniform logarithmic Sobolev inequality (LSI) for the conditional Gibbs measures (Theorem 4.11 below). Deducing these results becomes more complicated compared to [19].

Remark 3.1. Both the uniform LSI and the convexity of the coarse-grained Hamiltonian were originally provided in Deniz Dizdar's diploma thesis [12]. The main estimate to deduce the convergence of the microscopic dynamics to the mesoscopic dynamics (see Theorem 4.2 from below) was also deduced there.

3.1 A coarse-grained dynamics

We now build up the notion of the mesoscopic dynamics by *coarse-graining* the relevant features of the microscopic dynamics X_t.

Definition 3.2 (The coarse-graining operator P)**.** *For $M \in \mathbb{N}$, let $Y = Y_M$ be the space of spline functions of degree L with mean zero on the torus $\mathbb{T} = [0,1]$ corresponding to the mesh $\left\{\frac{m}{M}\right\}_{m\in[M]}$. That is*

$$Y_M := \left\{ y \in C^{L-1}(\mathbb{T}) \,\middle|\, \forall\, m \in [M] : y|_{\left(\frac{m-1}{M}, \frac{m}{M}\right)} \text{ polynomial of degree} \leq L, \text{ and } \int_0^1 y(\theta)d\theta = 0 \right\}.$$

In this work, we choose the degree of the splines in Y_M to be $L = 2$. We endow Y_M with the inner product inherited from $L^2(\mathbb{T})$. We define the coarse-graining operator $P : L^2(\mathbb{T}) \to Y_M$ as the L^2-orthogonal projection onto Y_M.

The following basic facts show that splines serve as good approximations for deducing the hydrodynamic limit in the H^{-1} norm.

Lemma 3.3 (Penalization of fluctuations by a strong norm)**.** *For any function $\zeta \in L^2_0(\mathbb{T})$,*

$$|\zeta - P\zeta|_{H^{-1}} \lesssim \frac{1}{M}|\zeta - P\zeta|_{L^2} \lesssim \frac{1}{M^2}|\zeta|_{H^1}, \tag{3.1}$$

$$|P\zeta|_{H^1} \lesssim |\zeta|_{H^1}, \qquad |P\zeta|_{H^{-1}} \lesssim |\zeta|_{H^{-1}}. \tag{3.2}$$

The proof of Lemma 3.3 is given in Section 6, where we gather and prove facts about splines. The core idea is a Poincaré inequality on an internal length scale $\frac{1}{M}$.

Having embedded the microscopic space X_N into $L^2(\mathbb{T})$ as a subspace of step functions (cf. (2.14)), we may apply the coarse-graining operator $P : L^2(\mathbb{T}) \to Y_M$ to it. On the space X_N, P acts by projecting step functions onto splines. We will need to work with the adjoint of this projection operation.

Definition 3.4 (Adjoint of the operator P)**.** *We define $P^t : Y_M \to X_N$ to be the unique linear operator satisfying*

$$\langle Px, y\rangle_{L^2} = x \cdot P^t y \quad \forall\, x \in X_N,\, y \in Y_M.$$

Equivalently, $NP^t : Y_M \to X_N$ is the unique linear operator satisfying

$$\langle Px, y\rangle_{L^2} = \langle x, NP^t y\rangle_{L^2} \quad \forall\, x \in X_N,\, y \in Y_M.$$

The factor N appears due to different Euclidean structures on X_N.

Remark 3.5. Since orthogonal projections between two subspaces of a Hilbert space are adjoints of each other, the L^2 adjoint NP^t of the projection operator $P : X_N \to Y_M$ is an L^2-orthogonal projection of Y_M onto X_N. Moreover, since X_N consists of step functions that are piecewise constant on intervals of length $1/N$, NP^t acts on splines in Y_M by taking their average over these intervals.

Because the spline space $Y_M \not\subseteq X_N$ for spline degree $L \geq 1$, the back-and-forth projection $PNP^t \neq \mathrm{id}_{Y_M}$ in general (cf. assumption (2) in [19]). However, once the block size $K = N/M$ is large enough, the microscopic space X_N will have enough resolution to fully describe the splines in Y_M, and PNP^t will become close to the identity operator.

Lemma 3.6. *It holds that*

$$\|PNP^t - \mathrm{id}_{Y_M}\| = O\left(\frac{M^2}{N^2}\right). \tag{3.3}$$

In particular, if $K = \frac{N}{M}$ is large enough, then $PNP^t : Y_M \to Y_M$ is invertible.

The proof of Lemma 3.3 is given in Section 6. From now on, we assume $N = KM$ for $K \in \mathbb{N}$ large enough so that $PNP^t : Y_M \to Y_M$ is invertible. In particular, this means the coarse-graining operator $P : X_N \to Y_M$ has full range and the orthogonal projection $NP^t : Y_M \to X_N$ is an embedding. Hence:

Definition 3.7 (Disintegration of the canonical ensemble μ)**.** *The operator P induces a decomposition of the Gibbs measure μ into a family of conditional measures $\mu(dx|y) := \mu(dx|Px = y)$ on the fibers $P^{-1}(y) \subset X_N$ and a marginal measure $\bar{\mu}(dy)$ on the image Y_M, in the sense that*

$$\int g(x)\mu(dx) = \int \int g(x)\mu(dx|y)\bar{\mu}(dy)$$

for any test function $g : X_N \to \mathbb{R}$.

More explicitly, the conditional measure $\mu(dx|y)$ is a probability measure of the form

$$\mu(dx|y) = \frac{1}{Z} \mathbb{1}_{\{Px=y\}}(x) \ \exp\left(-H(x)\right) \mathcal{L}^{N-M}(dx) \tag{3.4}$$

where $\mathcal{L}^{N-M}$ denotes the $N-M$-dimensional Hausdorff measure on the affine subspace $P^{-1}(y) \subset X_N$. The marginal measure $\bar{\mu}$ is a probability measure of the form

$$\bar{\mu}(dy) = \frac{1}{Z} \exp\left(-N\bar{H}(y)\right) dy,$$

where $\bar{H}$ is the *coarse-grained Hamiltonian* given by (3.5) below and dy is the Hausdorff measure on Y_M.

Definition 3.8 (Coarse-grained Hamiltonian $\bar{H}$)**.** *The coarse-grained Hamiltonian $\bar{H} : Y_M \to \mathbb{R}$ is given by*

$$\bar{H}(y) := -\frac{1}{N} \log \int_{\{x \in X_N : Px = y\}} \exp\left(-H(x)\right) \mathcal{L}^{N-M}(dx), \tag{3.5}$$

where $\mathcal{L}^{N-M}$ denotes the $N-M$-dimensional Hausdorff measure.

It follows from a short calculation that the gradient of $\bar{H}$ is also a coarse-grained version of the gradient of H:

$$NP^t \nabla \bar{H}(y) = \mathbb{E}_{\mu(dx|y)} \nabla H(x), \tag{3.6}$$

which serves as a crucial link between the microscopic and mesoscopic dynamics. The main advantage of the coarse-grained Hamiltonian $\bar{H}$ over the original microscopic Hamiltonian H is a convexification resulting from averaging over large blocks, which is a well-known phenomenon in statistical mechanics and will be central to our analysis.

Theorem 3.9 (Uniform strong convexity of $\bar{H}$)**.** *There are constants $0 < \lambda, \Lambda, K^* < \infty$ such that for all $K \geq K^*$, M and all $y \in Y_M$ it holds*

$$2\lambda \, \mathrm{id}_{Y_M} \leq \operatorname{Hess} \bar{H}(y) \leq 2\Lambda \, \mathrm{id}_{Y_M}$$

in the sense of quadratic forms.

Remark 3.10. Theorem 3.9 will be proven in Section 8. It should be compared to the similar statement of Lemma 29 in [19]. The situation here is more subtle. In [19], the mesoscopic observables are also piecewise constant functions and therefore local functions. In contrast, the mesoscopic observables in our setting are given by continuous splines which are non-local functions. This introduces additional interactions between blocks. We work around this obstacle by first deducing the strong convexity for mesoscopic observables that are piecewise polynomials of degree L, or *discontinuous Galerkin functions* in the jargon of numerical analysis, and then transferring the result back to the spline space Y_M.

Besides the coarse-grained Hamiltonian $\bar{H}$, we also need a coarse-grained version of the second-order difference operator $-A$.

Definition 3.11 (Coarse-grained operator $\bar{A}$)**.** *The coarse-grained second-order difference operator $-\bar{A}$ is defined by*

$$\bar{A} := PANP^t.$$

In particular, the coarse-grained operator $\bar{A}$ inherits the positive definiteness of the operator A. Hence:

Definition 3.12 (Euclidean structures on Y_M induced by $\bar{A}$)**.** *The operator $\bar{A}$ induces a dual pair of inner products on the spline space Y_M*

$$\langle y, z\rangle_{\bar{A}} := \langle y, \bar{A}z\rangle_{L^2} \quad \textit{and} \quad \langle y, z\rangle_{\bar{A}^{-1}} := \langle y, \bar{A}^{-1}z\rangle_{L^2}.$$

We denote by $|\cdot|_{\bar{A}}$ and $|\cdot|_{\bar{A}^{-1}}$ the corresponding norms on Y_M.

The definition of $-\bar{A}$ as a coarse-graining of the second-order difference operator $-A$ suggests that it is a discrete version of the second derivative adapted to the spline space Y_M (see Lemma 5.4 below for a precise statement). Indeed, it turns out that the $\bar{A}$ and $\bar{A}^{-1}$ norms are equivalent to the H^1 and H^{-1} norm on Y_M, respectively.

Lemma 3.13. *There exists an integer K^* such that for all $K \geq K^*, M$ and all $y \in Y_M$,*

$$|y|_{\bar{A}} \simeq |y|_{H^1} \quad \textit{and} \quad |y|_{\bar{A}^{-1}} \simeq |y|_{H^{-1}}. \tag{3.7}$$

The proof of Lemma 3.13 is given in Section 6, where we gather and prove facts about splines. This is where we need the degree of the splines in Y_M to be at least $L \geq 1$.

We are now ready to introduce the mesoscopic dynamics.

Definition 3.14 (Mesoscopic dynamics)**.** *The mesoscopic dynamics η_t on Y_M is given by a solution of the ordinary differential equation*

$$\frac{d}{dt}\eta_t = -\bar{A}\nabla\bar{H}(\eta_t). \tag{3.8}$$

Remark 3.15 (Gradient flow structure of the mesoscopic dynamics)**.** The mesoscopic dynamics may be viewed as a $\bar{A}^{-1}$-gradient flow for the energy functional $\bar{H}$,

$$\frac{d}{dt}\eta_t = -\nabla_{\bar{A}^{-1}}\bar{H}(\eta_t), \tag{3.9}$$

where $\nabla_{\bar{A}^{-1}} := \bar{A}\nabla$ denotes the gradient operation w.r.t. to the $\bar{A}^{-1}$ inner product. The strong convexity of the energy functional $\bar{H}$ then implies the convergence of all trajectories irrespective of the starting point. More precisely, if η_t and $\tilde{\eta}_t$ are two solutions of (3.8), then

$$\begin{aligned} \frac{d}{dt}\frac{1}{2}|\eta_t - \tilde{\eta}_t|^2_{\bar{A}^{-1}} &= -\langle \eta_t - \tilde{\eta}_t, \nabla\bar{H}(\eta_t) - \nabla\bar{H}(\tilde{\eta}_t)\rangle_{L^2} \\ &\leq -2\lambda\,|\eta_t - \tilde{\eta}_t|^2_{L^2}. \end{aligned} \tag{3.10}$$

Before moving on, let us take a closer look how the gradient operation $\nabla_{\bar{A}^{-1}}$ on Y_M is related to the gradient operation $\nabla_{A^{-1}}$ on X_N: for any test function $\xi : Y_M \to \mathbb{R}$, we have

$$\nabla_{A^{-1}}(\xi \circ P) = AP^t\bar{A}^{-1}(\nabla_{\bar{A}^{-1}}\xi) \circ P, \tag{3.11}$$

and for any $x, v \in X_N$,

$$v \cdot \nabla_{A^{-1}}\,\xi(Px) = \langle v, ANP^t\bar{A}^{-1}\,\nabla_{\bar{A}^{-1}}\,\xi(Px)\rangle_{L^2}. \tag{3.12}$$

It is straightforward to check that the operator norm of $ANP^t\bar{A}^{-1}$ appearing in (3.12) blows up if one projects onto piecewise constant functions or piecewise linear functions that are $C^0(\mathbb{T})$ (i.e. splines of degree $L=0$ or $L=1$). However, we do get a good control if we project onto splines of degree $L=2$ (see Lemma 4.13 below). This observation was the original motivation to consider the two-scale approach with the coarse-graining operator P given by Definition 3.2.

Remark 3.16. In [19], the coarse-graining operator P was defined as the L^2-orthogonal projection onto piecewise constant functions and one worked around the problem that operator $ANP^t\bar{A}^{-1}$ is unbounded by using a less straight-forward definition of $\bar{A}$ as $\bar{A}^{-1} := PA^{-1}NP^t$. That choice led to a sub-optimal error when comparing the microscopic to the mesoscopic evolution (see also Remark 4.5 below). Choosing splines of degree $L>2$ does not improve the error derived with our method further. For a more detailed explanation, see Remark 3.19 below.

3.2 Convergence of the dynamics

Now, we state the first ingredient of the two-scale approach.

Theorem 3.17 (Convergence of the microscopic to the mesoscopic dynamics)**.** *Under the same assumption as in Theorem* (2.18)*, let η denote the solution of the mesoscopic equation* (3.8)*. Then*

$$\sup_{0\le t\le T} \mathbb{E}|X_t-\eta_t|^2_{H^{-1}} \lesssim \mathbb{E}|X_0-\eta_0|^2_{H^{-1}} + \frac{T}{K} + \frac{1}{M^2}\left(C_{\mathrm{Ent}}+1\right), \tag{3.13}$$

where C_{Ent} is given by (2.20)*.*

We prove Theorem 3.17 in Section 4. The error term $\frac{T}{K}$ on the right hand side of (3.13) comes from comparing the stochastic microscopic dynamics to the deterministic mesoscopic dynamics. Its scaling corresponds to what one would expect from the decay of variance in the weak law of large numbers, if we had chosen to project onto piecewise constant functions, in which case y would be a vector whose entries are means of K weakly correlated random variables and η would be interpreted as the expectation of this vector.

Now, let us state the second ingredient in the two-scale approach.

Theorem 3.18 (Convergence of the mesoscopic to the macroscopic dynamics)**.** *Let η denote the solution of the mesoscopic dynamics* (3.8) *and let ζ denote the solution of the macroscopic dynamics* (2.18)*. Then*

$$\sup_{0\le t\le T}|\zeta_t-\eta_t|^2_{H^{-1}} + \int_0^T |\zeta_s-\eta_s|^2_{L^2}ds \lesssim |\zeta_0-\eta_0|^2_{H^{-1}} + \frac{T}{K} + \left(\frac{1}{K^2}+\frac{1}{M^2}\right)|\zeta_0|^2_{L^2}.$$

We prove Theorem 3.18 in Section 5. For the proof we adapt a standard method from numerical analysis, in which the mesoscopic evolution (3.8) is interpreted as a Galerkin approximation of the macroscopic evolution (2.18). The non-standard part of the argument is that when comparing (3.8) to (2.18) one gets two additional sources of errors. One source of error comes from approximating the Euclidean structure $\langle\cdot,\cdot\rangle_{H^{-1}}$ by the Euclidean structure $\langle\cdot,\bar{A}^{-1}\cdot\rangle_{L^2}$. The other source of error comes from approximating the gradient of the macroscopic free energy $\mathcal{H}$ by the gradient of the coarse-grained Hamiltonian $\bar{H}$.

We are now ready to give the proof of Theorem 2.16.

Proof of Theorem 2.16. We choose the initial condition of the mesoscopic dynamics η given by (3.8) to be $\eta_0 = P\zeta_0$. Combining Theorem 3.17 and Theorem 3.18 yields the estimate

$$\begin{aligned}\sup_{0\le t\le T}\mathbb{E}|X_t-\zeta_t|^2_{H^{-1}} &\le \sup_{0\le t\le T} 2\mathbb{E}|X_t-\eta_t|^2_{H^{-1}}+\sup_{0\le t\le T}2|\eta_t-\zeta_t|^2_{H^{-1}}\\ &\lesssim \mathbb{E}|X_0-P\zeta_0|^2_{H^{-1}}+|\zeta_0-P\zeta_0|^2_{H^{-1}}\\ &\quad+\frac{T}{K}+\frac{1}{M^2}(C_{\mathrm{Ent}}+1)+\left(\frac{1}{K^2}+\frac{1}{M^2}\right)|\zeta_0|^2_{L^2}.\end{aligned}$$

Applying (3.1) and (3.2), and choosing $K=M^2$ yields the desired estimate (2.21). □

Remark 3.19. We do not know whether the rate of convergence in Theorem 2.16 is optimal. What we could say instead is that the rate of convergence *derived from the two-scale approach* cannot be improved further by choosing splines of higher order for the mesoscopic scale. To see this, first notice from comparing the error scaling in K and M in Theorem 3.17 and Theorem 3.18 that the “bottleneck” of our method is the former, which compares the microscopic to mesoscopic dynamics. The main ingredient of this error estimate comes from Theorem 4.2 below:

$$\mathbb{E}|PX_0-\eta_0|^2_{\bar{A}^{-1}}+\frac{2T}{K}+2C\,\frac{C_{\mathrm{Ent}}}{M^2}.$$

In the constant C above (see Remark 4.4 below), the only place where splines of order 2 are required is a uniform upper bound for the operator norm of $ANP^t\bar{A}^{-1}$ (see Lemma 4.13 below and also the earlier Remark 3.16). Since it always holds that $\|ANP^t\bar{A}^{-1}\|\ge 1$ (simply compare with $P(ANP^t\bar{A}^{-1})=\bar{A}\bar{A}^{-1}=\mathrm{Id}$), the order of magnitude for this key estimate is *already* optimal. Thus, choosing splines of order higher than 2 will not improve the scaling of error derived in Theorem 4.2. Intuitively, the scaling in K is from central limit theorem over block of size K, and the scaling in M is from a discrete Poincaré inequality on length scale $1/M$ (see Lemma 4.10 below, which is a discrete analogue of the earlier Lemma 3.3), which we cannot expect to improve further.

4 Convergence of microscopic dynamics to mesoscopic dynamics

The proof of Theorem 3.17 is quite complex. Before proceeding to the rigorous argument let us give some heuristics. Theorem 3.17 states that the stochastic microscopic evolution given by the Kawasaki dynamics in (2.8), i.e.

$$dX_t=-\nabla_{A^{-1}}H(X_t)dt+\sqrt{2}dB_t^{A^{-1}},$$

is close in the H^{-1}-norm to the mesoscopic deterministic dynamics given by (3.9), i.e.

$$\frac{d}{dt}\eta=-\nabla_{\bar{A}^{-1}}\bar{H}(\eta). \tag{4.1}$$

- The first observation needed is that because the H^{-1}-norm is a weak norm (i.e. it involves integration, see Definition 2.10) one can control the difference between X_t and the projected process PX_t in this norm (see Lemma 3.3). Hence, it suffices to show that the stochastic evolution

  $$dPX_t=-P\nabla_{A^{-1}}H(X_t)dt+\sqrt{2}PdB_t^{A^{-1}} \tag{4.2}$$

  is close to the deterministic mesoscopic dynamics (4.1).
- Because the operator P takes averages over blocks of size K, the noise term $\sqrt{2}PdB_t^{A^{-1}}$ of the projected Kawasaki dynamics (4.2) should vanish as $K\to\infty$ by the law of large numbers. It is left to show that the dynamics

  $$\frac{d}{dt}PX_t=-P\nabla_{A^{-1}}H(X_t) \tag{4.3}$$

  is close to the mesoscopic dynamics (4.1).

- By the coarse-graining relation (3.6) one sees that the mesoscopic dynamics (4.1) can be written as

$$\frac{d}{dt}\eta_t = -P\,\mathbb{E}_\mu\left[\nabla_{A^{-1}} H(x) \mid Px = \eta_t\right], \tag{4.4}$$

  where the expectation is taken with respect to the canonical ensemble μ conditioned on the mesoscopic profile given by η_t.
- Let us recall that μ is also the equilibrium distribution of the Kawasaki dynamics (2.4) (see Lemma 2.5). The nearest-neighbor interaction of the spins mediated by the matrix A means the Kawasaki dynamics X_t equilibrates faster on smaller spatial scales, so we expect that the dynamics (4.3) and (4.4) are close if the blocks are much smaller than the overall system size N, in other words $\frac{K}{N} = \frac{1}{M} \to 0$. In the rigorous argument, this fact will be quantified with the help of a uniform LSI for conditional measures which characterizes the speed of the convergence to equilibrium (see Theorem 4.11 below).

Let us turn now to the rigorous proof of Theorem 3.17. The first ingredient of the proof is an estimate of the second moment of X_t in L^2 norm, which controls the difference in H^{-1} norm between X_t and the projected dynamics PX_t by Lemma 3.3.

Proposition 4.1 (Second moment estimate)**.** *Under assumption* (2.20)*, the Kawasaki dynamics satisfies that*

$$\mathbb{E}|X_t|^2 \le 2\left(\frac{2}{\hat\alpha}\operatorname{Ent}(\rho(0)|\mu) + \mathbb{E}_\mu|x|^2\right) \le \frac{N}{\hat\alpha}\left(4C_{\mathrm{Ent}} + 2\right). \tag{4.5}$$

Proof. This was shown as part of Proposition 24 in [19] using the dissipation of relative entropy for *Glauber dynamics*, where the authors also mentioned it may be derived directly from the results of [32]. We give a short proof based on this suggestion. The first inequality in (4.5) can be restated as

$$W_2^2(\rho(t),\delta_0) \le 2\left(\frac{2}{\hat\alpha}\operatorname{Ent}(\rho(0)|\mu) + W_2^2(\mu,\delta_0)\right),$$

where W_2 denotes the L^2-Wasserstein distance and δ_0 is the Dirac measure supported at $0 \in X_N$. To show this, we first apply a triangle inequality for Wasserstein distance, followed by a Young's inequality, to see that

$$W_2^2(\rho(t),\delta_0) \le 2\left(W_2^2(\rho(t),\mu) + W_2^2(\mu,\delta_0)\right).$$

It remains to bound the Wasserstein distance $W_2(\rho(t),\mu)$ by the initial entropy $\operatorname{Ent}(\rho(0)|\mu)$. This is the main step and is provided by the *Talagrand's inequality* (2.13) and the monotonic decrease of relative entropy in time. The second inequality (4.5) then follows from the assumption (2.20) on initial entropy and a *Poincaré inequality* for μ applied to the coordinate functions x_i,

$$\mathbb{E}_\mu\left(\sum_{i=1}^N x_i^2\right) = \sum_{i=1}^N \operatorname{Var}_\mu(x_i) \le \frac{1}{\hat\alpha}\sum_{i=1}^N \mathbb{E}_\mu|\nabla_X x_i|^2 = \frac{1}{\hat\alpha}(N-1),$$

which is a direct corollary of the LSI (2.12) for μ. □

In light of the equivalence between the H^{-1} norm and $\bar A^{-1}$ norm (Lemma 3.13), it remains to control the difference between the projected microscopic dynamics PX_t and the mesoscopic dynamics η_t in $\bar A^{-1}$ norm. This is provided by the following estimate, which constitutes the main part of the proof of Theorem 3.17.

Theorem 4.2. *Under the same assumptions as in Theorem 3.17, there is an integer K^* and $\lambda > 0$ such that for all $K \geq K^*$ and any finite time $T > 0$ it holds*

$$\frac{1}{2}\sup_{0\leq t\leq T}\mathbb{E}|PX_t - \eta_t|^2_{\bar{A}^{-1}} + \lambda\int_0^T \mathbb{E}|PX_t - \eta_t|^2_{L^2}\ dt \ \leq \mathbb{E}|PX_0 - \eta_0|^2_{\bar{A}^{-1}} + \frac{2T}{K} + 2C\,\frac{C_{\mathrm{Ent}}}{M^2}. \tag{4.6}$$

Remark 4.3. The estimate (4.6) also shows that the projected Kawasaki dynamics (4.2) is close to the mesoscopic dynamics (4.1) using a time-integrated strong norm. This is reminiscent of the well-known phenomenon of parabolic improvement in numerical analysis.

Remark 4.4. The universal constant $0 < C < \infty$ in Theorem 4.2 is given by $C = \frac{\kappa^2\gamma}{4\sigma^2\lambda\alpha^2}$, where the constants κ, λ, γ, σ, and α are given by:

- $\kappa := \|\operatorname{Hess} H\|$, which is bounded independently of N by the assumption (2.1), (2.2) and (2.3);
- 2λ the lower bound on $\operatorname{Hess}\bar{H}$ as in Theorem 3.9 from below;
- α is the constant of the logarithmic Sobolev inequality (LSI) from Theorem 4.11 from below;
- σ is the constant from Lemma 4.13 from below;
- γ the constant from Lemma 4.10 from below.

Remark 4.5. Theorem 4.2 was first derived in Dizdar's diploma thesis [12]. We present below a more streamlined derivation that makes clear the underlying gradient flow structure. Theorem 4.2 should be compared with Theorem 8 in [19]. They arrive at a similar bound for the deviation from hydrodynamic behavior with an additional term scaling like M^{-1}. As mentioned before this additional error term occurs due to their choice of the coarse-graining operator P as the projection onto piecewise constant functions and the different definition of $\bar{A}$.

Theorem 4.2 will be proven in Section 4.1. We finish this section with a quick derivation of Theorem 3.17 based on the ingredients above.

Proof of Theorem 3.17. Combining Proposition 4.1 and Theorem 4.2 together with Lemma 3.3 and Lemma 3.13,

$$\begin{aligned}
\mathbb{E}|X_t - \eta_t|^2_{H^{-1}} &\leq 2\,\mathbb{E}|X_t - PX_t|^2_{H^{-1}} + 2\,\mathbb{E}|PX_t - \eta_t|^2_{H^{-1}} \\
&\lesssim \frac{1}{M^2}\mathbb{E}|X_t|^2_{L^2} + \mathbb{E}|PX_t - \eta_t|^2_{\bar{A}^{-1}} \\
&\lesssim \frac{1}{M^2}\frac{1}{\hat{\alpha}}(C_{\mathrm{Ent}} + 1) + \mathbb{E}|PX_0 - \eta_0|^2_{\bar{A}^{-1}} + \frac{T}{K} + C\frac{C_{\mathrm{Ent}}}{M^2} \\
&\lesssim \mathbb{E}|PX_0 - \eta_0|^2_{H^{-1}} + \frac{T}{K} + \frac{1}{M^2}(C_{\mathrm{Ent}} + 1).
\end{aligned}$$

This verifies the estimate (3.13). □

4.1 Proof of Theorem 4.2

Before we proceed, let us first introduce an conditioning for the Kawasaki dynamics X_t on the level of mesoscopic profiles, analogous to the disintegration of the canonical ensemble μ from Definition 3.7.

Definition 4.6 (Disintegration of the law ρ_t of the Kawasaki dynamics)**.** *The operator P induces a decomposition of the law ρ_t of the Kawasaki dynamics into a family of conditional measures $\rho_t(dx|y) := \rho_t(dx|Px = y)$ on the fibers $P^{-1}(y) \subset X_N$ and a marginal measure $\bar\rho_t(dy)$ on the image Y_M, in the sense that*

$$\int g(x)\rho_t(dx) = \int\int g(x)\rho_t(dx|y)\bar\rho_t(dy) \tag{4.7}$$

for any test function $g : X_N \to \mathbb{R}$.

We also need a decomposition of the microscopic observables in X_N into parts corresponding to mesoscopic profiles and microscopic fluctuations.

Definition 4.7 (Orthogonal decomposition of the state space X_N)**.** *The operator P induces an orthogonal decomposition of the state space X_N into a subspace corresponding to mesoscopic profiles, $X_N^\perp := \operatorname{im} NP^t$, and a subspace corresponding to microscopic fluctuations around these profiles, $X_N^\parallel := (\operatorname{im} NP^t)^\perp = \ker P \cap X_N$, as*

$$X_N \ni x = x_\parallel \oplus x_\perp \in X_N^\parallel \oplus X_N^\perp.$$

For a smooth function $f : X_N \to \mathbb{R}$, its gradient $\nabla_\parallel$ along $X_N^\parallel$ and its gradient $\nabla_\perp$ along $X_N^\perp$ are given by $\nabla_\parallel f = (\nabla f)_\parallel$ and $\nabla_\perp f = (\nabla f)_\perp$.

The starting point of the analysis is an equation of time evolution for the difference between the projected dynamics PX_t and the mesoscopic dynamics η_t, provided by the forward Kolmogorov equation (2.10) for the microscopic dynamics X_t:

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{2}\,\mathbb{E}\,|PX_t - \eta_t|^2_{\bar A^{-1}}\\
&= \int \left(\frac{d}{dt} - \nabla H(x)\cdot\nabla_{A^{-1}} + \nabla\cdot\nabla_{A^{-1}}\right)\left(\frac{1}{2}|Px-\eta_t|^2_{\bar A^{-1}}\right)\rho_t(dx)\\
&= -\int \left\langle \frac{d}{dt}\eta_t, Px-\eta_t\right\rangle_{\bar A^{-1}}\rho_t(dx) - \int \langle \nabla H(x), ANP^t\bar A^{-1}(Px-\eta_t)\rangle_{L^2}\,\rho_t(dx)\\
&\quad + \int \nabla\cdot(AP^t\bar A^{-1}(Px-\eta_t))\,\rho_t(dx)
\end{aligned} \tag{4.8}$$

where we used the relations (3.11) and (3.12) between $\nabla_{A^{-1}}$ and $\nabla_{\bar A^{-1}}$.

Let us first look at the last integral in (4.8). This is a purely *entropic* term coming from the projected Brownian motion $PdB_t^{A^{-1}}$ (see (4.2)), whose covariance matrix can be easily calculated to be $\frac{\mathrm{id}_Y}{N}$ w.r.t. the $\bar A^{-1}$ inner product. Indeed, the integrand evaluates to

$$\operatorname{tr}_X(AP^t\bar A^{-1}P) = \operatorname{tr}_Y(PAP^t\bar A^{-1}) = \operatorname{tr}_Y\left(\frac{\mathrm{id}_Y}{N}\right) = \frac{\dim Y}{N}.$$

This is a constant $\frac{M-1}{N} \approx \frac{1}{K}$ that scales like the variance of the average of K i.i.d. random variables. It accounts for the discrepancy that the Kawasaki dynamics (2.4) has noise whereas the mesoscopic dynamics (3.8) is deterministic.

Having disposed of the noise term, we now use the definition of the dynamics η_t and apply disintegration of measure (4.7) to rewrite the time evolution equation as

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}\,\mathbb{E}|PX_t-\eta_t|^2_{\bar A^{-1}} &= \frac{M-1}{N} - \int \langle -\nabla \bar H(\eta_t), y-\eta_t\rangle_{L^2}\,\bar\rho_t(dy)\\
&\quad - \int \langle \mathbb{E}_{\rho_t(dx|y)}\nabla H(x), ANP^t\bar A^{-1}(y-\eta_t)\rangle_{L^2}\,\bar\rho_t(dy).
\end{aligned}$$

Notice that the inner product inside the first integrand is almost in the form (3.10), except missing the counterpart

$$\langle \nabla \bar{H}(y), y - \eta_t \rangle_{L^2}.$$

Remarkably, the missing mesoscopic energy gradient may be supplied by the average microscopic energy gradient, when the *microscopic fluctuations* of the dynamics X_t around the mesoscopic profile y has *reached stochastic equilibrium*. Namely, the coarse-graining relation (3.6) between ∇H and $\nabla \bar{H}$ implies that

$$\nabla \bar{H}(y) = \bar{A}^{-1} P A \ \mathbb{E}_{\mu(dx|y)} \nabla H(x). \tag{4.9}$$

Moreover, the operator $\bar{A}^{-1}PA : X_N \to Y_M$ in (4.9) is exactly the L^2 adjoint of the operator $ANP^t\bar{A}^{-1} : Y_M \to X_N$ in (3.11). These observations lead to the rearranged equation

$$\begin{aligned} \frac{d}{dt}\frac{1}{2}\,\mathbb{E}|PX_t - \eta_t|^2_{\bar{A}^{-1}} = \frac{M-1}{N} &- \mathbb{E}\,\langle \nabla \bar{H}(PX_t) - \nabla \bar{H}(\eta_t), PX_t - \eta_t\rangle_{L^2} \\ &- \int \langle \mathbb{E}_{\rho_t(dx|y)} \nabla H(x) - \mathbb{E}_{\mu(dx|y)} \nabla H(x), ANP^t \bar{A}^{-1}(y - \eta_t)\rangle_{L^2}\, \bar{\rho}_t(dy). \end{aligned} \tag{4.10}$$

The first expectation can now be estimated by the uniform convexity of $\bar{H}$ from Theorem 3.9:

$$-\mathbb{E}\,\langle \nabla \bar{H}(PX_t) - \nabla \bar{H}(\eta_t), PX_t - \eta_t\rangle_{L^2} \le -2\lambda\, \mathbb{E}\,|PX_t - \eta_t|^2_{L^2}. \tag{4.11}$$

It remains to estimate the second integral on the right hand side of (4.10). After taking the operator norm of $ANP^t\bar{A}^{-1}$ into account (see Lemma 4.13 below), this integral quantifies how far away the conditional measure $\rho_t(dx|y)$ is from the conditional Gibbs measure $\mu(dx|y)$ through the *mean difference*

$$\mathbb{E}_{\rho_t(dx|y)} \nabla H(x) - \mathbb{E}_{\mu(dx|y)} \nabla H(x).$$

This quantity turns out to be controlled by a bound on the operator norm of $\operatorname{Hess} H$ and a logarithmic Sobolev inequality (LSI) of the conditional Gibbs measure $\mu(dx|y)$.

Proposition 4.8 (Mean difference estimate)**.** *Let $\kappa := \|\operatorname{Hess} H\|$. Suppose $\mu(dx|y)$ given by* (2.9) *satisfies a LSI with constant $\alpha > 0$ in the sense of* (4.14) *below. Then we have:*

$$|\mathbb{E}_{\rho_t(dx|y)} \nabla H(x) - \mathbb{E}_{\mu(dx|y)} \nabla H(x)|^2 \ \le\ \frac{\kappa^2}{\alpha^2} \int |\nabla_{||} \log f(t,x)|^2\, \rho_t(dx|y), \tag{4.12}$$

where $\nabla_{||}$ is the gradient along the subspace of fluctuations defined in Definition 4.7.

Remark 4.9. This is a well-known result, e.g. contained in Lemma 22 in [19] in the form of a covariance estimate. The proof given there starts with using the Kantorovich-Rubinstein duality to bound the mean difference by the Wasserstein distance. By Theorem 1 in [32], the Wasserstein distance is bounded by the relative entropy (i.e. a *Talagrand's inequality* holds) provided a log-Sobolev inequality holds, which in turn bounds the relative entropy by the Fisher information.

Moreover, the following discrete analogue of (3.1) allows us to pass from the Fisher information involving $\nabla_{||}$ in (4.12) to the full Fisher information for the Kawasaki dynamics in (2.11).

Lemma 4.10 (Penalization of fluctuations by spin-exchange)**.** *There exists constant $\gamma > 0$ such that for $x \in X_N$*

$$|x_{||}|^2 = |x - x_\perp|^2 \ \le\ \frac{\gamma}{M^2} x \cdot Ax, \tag{4.13}$$

where $x_{||}, x_\perp$ denote the fluctuation and mesoscopic parts of x, respectively, as defined in Definition 4.7.

The proof of Lemma 4.10 is given in Section 6, where we gather and prove facts about splines. It remains to establish a uniform log Sobolev inequality:

Theorem 4.11 (Uniform LSI for $\mu(dx|y)$)**.** *The conditional measure $\mu(dx|y)$ given by* (3.4) *satisfies a LSI with constant $\alpha > 0$ uniform in the system size N and the mesoscopic profile y. More precisely, this means that for any nonnegative test function $g : X_N \to \mathbb{R}$ that satisfies $\int g(x)\mu(dx|y) = 1$, it holds that*

$$\operatorname{Ent}\left(g\mu(dx|y)|\mu(dx|y)\right) \leq \frac{1}{2\alpha}\int \frac{|\nabla_{||}g(x)|^2}{g(x)}\mu(dx|y), \tag{4.14}$$

where $\nabla_{||}$ is the gradient along the subspace of fluctuations defined in Definition 4.7.

Remark 4.12. Theorem 4.11 should be compared to [19, Theorem 14], where a similar statement was deduced for the case $L = 0$ using the *two-scale criterion* for LSI (see Lemma 10.7 below). As with the proof of Theorem 3.9, because blocks are not independent for splines in Y_M, we cannot directly apply the two-scale criterion and have to take a detour through the space of discontinuous Galerkin functions.

The proof of Theorem 4.11 is given in Section 8. To get an overall estimate for the integral involving the mean difference, we also need to control the operator norm of $ANP^t\bar{A}^{-1}$, which measures the compatibility of projecting and taking second differences.

Lemma 4.13 (Interchanging second-order difference with coarse-graining)**.** *There exists a universal constant $\sigma > 0$ and an integer K^* such that for all $K \geq K^*, M$ and all $y \in Y_M$ it holds*

$$|ANP^t\bar{A}^{-1}y|_{L^2} \leq \frac{1}{\sigma}|y|_{L^2}. \tag{4.15}$$

The proof of Lemma 4.13 is given in Section 6, where we gather and prove facts about splines. We are now ready to prove Theorem 4.2.

Proof of Theorem 4.2. Applying the convexity estimate (4.11) to the evolution equation (4.10) and using Lemma 4.13 and Cauchy-Schwarz on the last integral yields that

$$\begin{aligned} &\frac{d}{dt}\frac{1}{2}\,\mathbb{E}|PX_t - \eta_t|^2_{\bar{A}^{-1}} + 2\lambda\,\mathbb{E}|PX_t - \eta_t|^2_{L^2} \\ &\leq \frac{1}{K} + \int \frac{1}{\sigma}|y - \eta_t|_{L^2}\left|\mathbb{E}_{\rho_t(dx|y)}\nabla H(x) - \mathbb{E}_{\mu(dx|y)}\nabla H(x)\right|_{L^2}\bar{\rho}_t(dy), \end{aligned} \tag{4.16}$$

where we accounted for different Euclidean structures on X_N. By Lemma 4.8, Lemma 4.10 and the observation (2.11), we have

$$\int |\mathbb{E}_{\rho_t(dx|y)}\nabla H(x) - \mathbb{E}_{\mu(dx|y)}\nabla H(x)|^2_{L^2}\,\bar{\rho}_t(dy) \leq -\frac{\kappa^2}{\alpha^2}\frac{\gamma}{M^2}\frac{1}{N}\frac{d}{dt}\operatorname{Ent}\left(\rho_t|\mu\right).$$

Applying Young's inequality and using this estimate, the integral on the right hand side of (4.16) is bounded by

$$\lambda\,\mathbb{E}|PX_t - \eta_t|^2_{L^2} - \frac{1}{4\lambda}\frac{\kappa^2\gamma}{\sigma^2\alpha^2}\frac{1}{M^2}\frac{1}{N}\,\frac{d}{dt}\operatorname{Ent}\left(\rho_t|\mu\right). \tag{4.17}$$

Putting the upper bound (4.17) back into (4.16) and integrating over the time interval $[0, T]$ yields the desired estimate (4.6). □

5 Convergence of mesoscopic dynamics to macroscopic dynamics

In this section we state the proof of Theorem 3.18. We need to show that the mesoscopic evolution (3.8)

$$\frac{d}{dt}\eta_t = -\nabla_{\bar{A}^{-1}}\bar{H}(\eta_t) = -\bar{A}\nabla\bar{H}(\eta_t)$$

converges to the macroscopic evolution (2.18)

$$\frac{\partial}{\partial t}\zeta_t = -\nabla_{H^{-1}}\mathcal{H}(\zeta_t) = \frac{\partial^2}{\partial\theta^2}\varphi'(\zeta_t).$$

Formally, this means that one has to exchange the coarse-grained operator $-\bar{A}$ with the second derivative operator $\frac{\partial^2}{\partial\theta^2}$ and the gradient of the coarse-grained Hamiltonian $\nabla\bar{H}$ with the gradient of the macroscopic free energy $\nabla\mathcal{H} = \varphi'$.

- The first exchange is plausible because $-\bar{A}$ is a coarse-grained version of the second-order difference operator $-A$.
- The second exchange represents a passage from microscopic free energy to macroscopic free energy, which makes sense from a thermodynamic perspective. It is essentially the consequence of a (local) Cramér theorem: $\bar{H}$ is a coarse-graining of the microscopic Hamiltonian H with single-site potential ψ, while φ is the Cramér transform of the same ψ.

The proof of Theorem 3.18 is inspired by the Galerkin approximation scheme, a well-known method in numerical analysis. First, we need to show the macroscopic dynamics ζ_t is close to the projected dynamics $P\zeta_t$. Because the H^{-1} norm is a weak norm, this difference is controlled by the spline estimates in Lemma 3.3 together with the following a priori energy estimates.

Lemma 5.1. *Let ζ_t denote the macroscopic dynamics given by* (2.22)*. Then it holds that*

$$\sup_{0\le t\le T}|\zeta_t|^2_{L^2} \le \frac{\Lambda}{\lambda}|\zeta_0|^2_{L^2} \quad \textit{and} \quad \int_0^\infty |\zeta_t|^2_{H^1}\,dt \le \frac{\Lambda}{\lambda^2}|\zeta_0|^2_{L^2}.$$

Proof. The convexity and C^2 estimates of φ in Lemma 2.12 yields

$$\lambda|\zeta|^2_{L^2} \le \mathcal{H}(\zeta) \le \Lambda|\zeta|^2_{L^2} \quad \text{and} \quad \lambda|\zeta|_{H^1} \le |\varphi'(\zeta)|_{H^1} \le \Lambda|\zeta|_{H^1}.$$

With the help of these estimates, we integrate the dissipation (2.19) of the macroscopic free energy $\mathcal{H}$ w.r.t. the H^{-1} gradient flow structure to find that

$$\sup_{0\le t\le T}|\zeta_t|^2_{L^2} \le \sup_{0\le t\le T}\frac{1}{\lambda}\mathcal{H}(\zeta_t) = \frac{1}{\lambda}\mathcal{H}(\zeta_0) \le \frac{\Lambda}{\lambda}|\zeta_0|^2_{L^2},$$
$$\int_0^\infty |\zeta_t|^2_{H^1}dt \le \frac{1}{\lambda^2}\int_0^\infty |\varphi'(\zeta)|^2_{H^1} \le \frac{1}{\lambda^2}\mathcal{H}(\zeta_0) \le \frac{\Lambda}{\lambda^2}|\zeta_0|^2_{L^2}.$$ □

Remark 5.2. Lemma 5.1 may be compared with Proposition 4.1. The proofs for both are ultimately based on how the gradient flow structure of an underlying dynamics dissipates its associated energy functional.

It remains to show that the mesoscopic dynamics η_t is close to the projected dynamics $P\zeta_t$. Because of Lemma 3.13, it is more convenient to work with the $\bar{A}^{-1}$ norm instead of the H^{-1} norm. Differentiating in time and using the definition of the dynamics, we have

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}|\eta_t - P\zeta_t|^2_{\bar{A}^{-1}} &= \left\langle \frac{d}{dt}\eta_t - \frac{d}{dt}P\zeta_t, \eta_t - P\zeta_t \right\rangle_{\bar{A}^{-1}} \\ &= \left\langle -\nabla_{\bar{A}^{-1}}\bar{H}(\eta_t) - P\frac{\partial^2}{\partial\theta^2}\varphi'(\zeta_t), \eta_t - P\zeta_t \right\rangle_{\bar{A}^{-1}},\end{aligned}$$

where we used that $\frac{d}{dt}P\zeta_t = P\frac{\partial\zeta}{\partial t}$, i.e. we may interchange time derivative and projection onto splines. This resembles the form of (3.10), leading to the rearranged equation

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}|\eta_t - P\zeta_t|^2_{\bar{A}^{-1}} &= \left\langle -\nabla_{\bar{A}^{-1}}\bar{H}(\eta_t) + \nabla_{\bar{A}^{-1}}\bar{H}(P\zeta_t), \eta_t - P\zeta_t \right\rangle_{\bar{A}^{-1}} \\
&\quad + \langle \bar{A}\varphi'(\zeta_t) - \nabla_{\bar{A}^{-1}}\bar{H}(P\zeta_t), \eta_t - P\zeta_t\rangle_{\bar{A}^{-1}} \\
&\quad + \left\langle -P\frac{\partial^2}{\partial\theta^2}\varphi'(\zeta_t) - \bar{A}\varphi'(\zeta_t), \eta_t - P\zeta_t \right\rangle_{\bar{A}^{-1}} \\
&= -\langle \nabla\bar{H}(\eta_t) - \nabla\bar{H}(P\zeta_t), \eta_t - P\zeta_t\rangle_{L^2} \\
&\quad + \langle \varphi'(\zeta_t) - \nabla\bar{H}(P\zeta_t), \eta_t - P\zeta_t\rangle_{L^2} \\
&\quad + \langle \varphi'(\zeta_t), (-\partial_\theta^2\bar{A}^{-1} - \mathrm{id})(\eta_t - P\zeta_t)\rangle_{L^2}.
\end{aligned} \tag{5.1}$$

The first term in (5.1) is now in the same form of (3.10) and can be estimated by the uniform strong convexity of $\bar{H}$ (see Theorem 3.9):

$$\langle \nabla\bar{H}(P\zeta_t) - \nabla\bar{H}(\eta_t), \eta_t - P\zeta_t\rangle_{L^2} \leq -\lambda|\eta_t - P\zeta_t|^2_{L^2}. \tag{5.2}$$

The second term in (5.1) is small because the gradient of the coarse-grained Hamiltonian $\bar{H}$ is a good approximation of the gradient of the macroscopic free energy $\mathcal{H}$.

Theorem 5.3 (Closeness of $\nabla\bar{H}$ and $\nabla\mathcal{H}$)**.** *There is an integer K^* such that if $K \geq K^*$ then it holds for all $\zeta \in L^2_0(\mathbb{T})$*

$$\left|\nabla\bar{H}(P\zeta) - \nabla\mathcal{H}(\zeta)\right|_{L^2} \lesssim \left(\frac{1}{K} + \frac{1}{M}\right)|\zeta|_{H^1} + \frac{1}{K^{\frac{1}{2}}}, \tag{5.3}$$

where the gradient $\nabla\mathcal{H}$ is taken in $L^2_0(\mathbb{T})$.

We prove Theorem 5.3 in Section 9. The last term in (5.1) is controlled by the following error estimates for exchanging the coarse-grained second-order difference operator $-\bar{A}$ and the second derivative ∂_θ^2.

Lemma 5.4 (Discrepancy between $-\bar{A}$ and ∂_θ^2)**.** *There exists an integer K^* such that for all $K \geq K^*, M$ and all $y, \tilde{y} \in Y_M$,*

$$|-\partial_\theta^2\bar{A}^{-1}y|_{L^2} \lesssim |y|_{L^2}, \tag{5.4}$$

$$|\langle -\partial_\theta^2\bar{A}^{-1}y, \tilde{y}\rangle_{L^2} - \langle y, \tilde{y}\rangle_{L^2}| \lesssim \frac{1}{K}|y|_{H^{-1}}|\tilde{y}|_{H^1}. \tag{5.5}$$

The proof of Lemma 5.4 is given in Section 6, where we gather and prove facts about splines. The error estimate (5.4) is closely related to the error estimate (4.15). We are now ready to prove Theorem 3.18.

Proof of Theorem 3.18. We first bound $P\zeta_t - \zeta_t$. By Lemma 3.3 and Lemma 5.1,

$$\begin{aligned}
\sup_{0\leq t\leq T}|P\zeta_t - \zeta_t|^2_{H^{-1}} &\lesssim \sup_{0\leq t\leq T}\frac{1}{M^2}|\zeta_t|^2_{L^2} \lesssim \frac{1}{M^2}\frac{\Lambda}{\lambda}|\zeta_0|^2_{L^2}, \\
\int_0^T|P\zeta_t - \zeta_t|^2_{L^2}dt &\lesssim \int_0^T\frac{1}{M^2}|\zeta_t|^2_{H_1}dt \lesssim \frac{1}{M^2}\frac{\Lambda}{\lambda^2}|\zeta_0|^2_{L^2}.
\end{aligned}$$

We now bound $\eta_t - P\zeta_t$. By Theorem 5.3, the second term in (5.1) is estimated as

$$\langle \varphi'(\zeta_t) - \nabla\bar{H}(P\zeta_t), \eta_t - P\zeta_t\rangle_{L^2} \lesssim \left(\left(\frac{1}{K} + \frac{1}{M}\right)|\zeta_t|_{H^1} + \frac{1}{K^{\frac{1}{2}}}\right)|\eta_t - P\zeta_t|_{L^2}. \tag{5.6}$$

By Lemma 5.4 and Lemma 3.3, the last term in (5.1) is estimated as

$$
\begin{aligned}
&\langle \varphi'(\zeta_t), (-\partial_\theta^2 \bar{A}^{-1} - \mathrm{id})(\eta_t - P\zeta_t)\rangle_{L^2} \\
&= \langle \varphi'(\zeta_t) - P\varphi'(\zeta_t), -\partial_\theta^2 \bar{A}^{-1}(\eta_t - P\zeta_t)\rangle_{L^2} + \langle P\varphi'(\zeta_t), (-\partial_\theta^2 \bar{A}^{-1} - \mathrm{id})(\eta_t - P\zeta_t)\rangle_{L^2} \\
&\lesssim |\varphi'(\zeta_t) - P\varphi'(\zeta_t)|_{L^2} |\eta_t - P\zeta_t|_{L^2} + \frac{1}{K} |P\varphi'(\zeta_t)|_{H^1} |\eta_t - P\zeta_t|_{H^{-1}} \\
&\lesssim \frac{1}{M} |\varphi'(\zeta_t)|_{H^1} |\eta_t - P\zeta_t|_{L^2} + \frac{1}{K} |\varphi'(\zeta_t)|_{H^1} |\eta_t - P\zeta_t|_{L^2}.
\end{aligned} \tag{5.7}
$$

Combining the estimates (5.2), (5.6), and (5.7) for equation (5.1) and applying Young's inequality yields that

$$
\frac{d}{dt}\frac{1}{2}|\eta_t - P\zeta_t|^2_{\bar{A}^{-1}} + \frac{\lambda}{2}|\eta_t - P\zeta_t|^2_{L^2} \lesssim \frac{1}{K} + \left(\frac{1}{K^2} + \frac{1}{M^2}\right)\left(|\zeta_t|^2_{H^1} + |\varphi'(\zeta_t)|^2_{H^1}\right).
$$

Integrating in time from 0 to T, applying the energy estimates in Lemma 5.1, and exchanging $\bar{A}^{-1}$ norm with H^{-1} norm (see Lemma 3.13), we get

$$
\sup_{0 \le t \le T} \frac{1}{2}|\eta_t - P\zeta_t|^2_{H^{-1}} + \frac{\lambda}{2}\int_0^T |\eta_t - P\zeta_t|^2_{L^2} dt \lesssim \frac{T}{K} + \left(\frac{1}{K^2} + \frac{1}{M^2}\right)|\zeta_0|^2_{L^2}.
$$

Combining the estimates for $\eta_t - P\zeta_t$ and $P\zeta_t - \zeta_t$ yields Theorem 3.18. □

6 Properties of spline approximations

In this section we prove the facts about splines $y \in Y_M$ used in the two-scale approach in this work. We begin with the observation that since Y_M is a finite-dimensional space, different norms on Y_M are equivalent for each M. More quantitatively:

Lemma 6.1 (Inverse Sobolev inequality on Y_M). *For all* $y \in Y_M$,

$$
|y|_{H^2} \lesssim M|y|_{H^1} \lesssim M^2|y|_{L^2}. \tag{6.1}
$$

The factors M and M^2 comes from a scaling argument, i.e. $\frac{1}{M}$ is the only internal length scale. We omit the proof of this fact, which consists of a simple dimensional analysis. This leads to a quick proof of Lemma 3.6.

Proof of Lemma 3.6. Recall that the adjoint operator $NP^t : Y_M \to X_N$ is an L^2-orthogonal projection onto piecewise constant functions on the intervals $[\frac{n-1}{N}, \frac{n}{N}]$. Therefore, for any $y \in Y_M$,

$$
\begin{aligned}
\langle (\mathrm{id} - PNP^t)y, y\rangle_{L^2} &= |y|^2_{L^2} - |NP^t y|^2_{L^2} \\
&= |y - NP^t y|^2_{L^2} \\
&\lesssim \frac{1}{N^2}|y|^2_{H^1} \overset{(6.1)}{\lesssim} \frac{M^2}{N^2}|y|^2_{L^2},
\end{aligned}
$$

where the equality in the second line is by Pythagorean theorem, and we then used a Poincaré inequality for each interval $[\frac{n-1}{N}, \frac{n}{N}]$ followed by the inverse Sobolev inequality. This verifies estimate (3.3) through the variational characterization

$$
\|T\| = \sup_{v \neq 0} \frac{\langle Tv, v\rangle}{|v|^2},
$$

for the self-adjoint operator $T = \mathrm{id} - PNP^t$. □

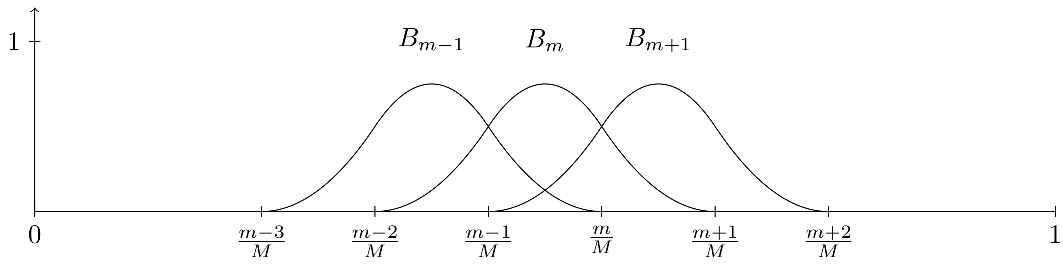

Figure 2: B-spline functions B_m.

6.1 Penalization of fluctuations around spline profiles

In this subsection we prove those auxiliary results that show fluctuations around spline profiles are penalized when measured in a weak norm, namely Lemma 3.3 and Lemma 4.10. For this purpose, we introduce a nice basis for the spline functions that forms a "partition of unity" on the torus $\mathbb{T}$.

Definition 6.2 (B-spline functions)**.** *The B-spline functions are given by*

$$B_m(\theta) = \begin{cases} \frac{M^2}{2}(\theta - \frac{m-2}{M})^2 & \text{for } \theta \in [\frac{m-2}{M}, \frac{m-1}{M}), \\ \frac{3}{4} - M^2(\theta - \frac{m-1/2}{M})^2 & \text{for } \theta \in [\frac{m-1}{M}, \frac{m}{M}), \\ \frac{M^2}{2}(\theta - \frac{m+1}{M})^2 & \text{for } \theta \in [\frac{m}{M}, \frac{m+1}{M}), \\ 0 & \text{else}. \end{cases} \tag{6.2}$$

Remark 6.3. The B-spline functions B_m have the following nice properties:

- $0 \le B_m < 1$,
- B_m is supported on $[\frac{m-2}{M}, \frac{m+1}{M}]$, and
- $\sum_{m=1}^{M} B_m = 1$.

This means the functions B_m have small overlap. More precisely,

$$\langle B_j, B_k \rangle_{L^2} = \frac{1}{M} B_{jk}, \tag{6.3}$$

where B is the symmetric matrix

$$B_{jk} = \frac{11}{20}\delta_{j=k} + \frac{13}{60}\delta_{|j-k|=1} + \frac{1}{120}\delta_{|j-k|=2}.$$

Proof of Lemma 3.3. The proof of (3.1) is based on the following spline interpolation: for $\zeta \in H^1(\mathbb{T})$, we define $I\zeta \in Y_M$ as

$$I\zeta(\theta) = \sum_{m=1}^{M} \zeta\left(\frac{m-1/2}{M}\right) B_m(\theta),$$

where $B_m \in Y_M$ is the B-spline basis defined in (6.2). We claim that

$$|\zeta - P\zeta|_{L^2} \le |\zeta - I\zeta|_{L^2} \lesssim \frac{1}{M} |\zeta|_{H^1}, \tag{6.4}$$

which establishes the second estimate of (3.1), from which the rest follows by duality. To verify (6.4), note the first inequality is simply due to the fact that $P\zeta$ is the best L^2 approximation of ζ in Y_M. The second estimate of (6.4) is well known in the literature on B-splines (see for example [10]). For the convenience of the reader we give a short proof of this fact. Using the fact that the B_m are supported on the intervals $[\frac{m-2}{M},\frac{m+1}{M}]$ and sum to 1, we obtain for $\theta \in \left(\frac{m-1}{M},\frac{m}{M}\right)$:

$$\zeta(\theta) - I\zeta(\theta) \overset{(6.2)}{=} \sum_{j=-1}^{1} \left(\zeta(\theta) - \zeta\left(\frac{m+j-1/2}{M}\right)\right) B_{m+j}(\theta). \tag{6.5}$$

Using the Fundamental Theorem of Calculus and the Cauchy-Schwarz inequality, we get

$$\begin{aligned} \int_{\frac{m-1}{M}}^{\frac{m}{M}} |\zeta(\theta) - I\zeta(\theta)|^2 \, d\theta &\le \int_{\frac{m-1}{M}}^{\frac{m}{M}} 3 \sum_{j=-1}^{1} \left(\zeta(\theta) - \zeta\left(\frac{m+j-1/2}{M}\right)\right)^2 B_{m+j}(\theta)^2 \, d\theta \\ &\le \int_{\frac{m-1}{M}}^{\frac{m}{M}} 3 \sum_{j=-1}^{1} \frac{2}{M} \left(\int_{\frac{m-3/2}{M}}^{\frac{m+1/2}{M}} |\zeta'(\widetilde{\theta})|^2 d\widetilde{\theta}\right) B_{m+j}(\theta)^2 \, d\theta \\ &\le \frac{6}{M^2} \left|\sum_{j=1}^{M} B_j^2\right|_{L^\infty} \left(\int_{\frac{m-3/2}{M}}^{\frac{m+1/2}{M}} |\zeta'(\widetilde{\theta})|^2 d\widetilde{\theta}\right). \end{aligned} \tag{6.6}$$

Summing over $m = 1, ..., M$ yields the second estimate of (6.4).

To verify (3.2), again by duality it suffices to show the first inequality only, which follows at once from the next two estimates

$$|\zeta - I\zeta|_{H^1} \lesssim |\zeta|_{H^1} \qquad \text{and} \qquad |P\zeta - I\zeta|_{H^1} \lesssim |\zeta|_{H^1}. \tag{6.7}$$

The first estimate of (6.7) can be deduced by differentiating the equation (6.5) and performing similar estimates as in (6.6). We omit the details of these calculations. The second estimate of (6.7) follows from the inverse Sobolev inequality 6.1 and combining the two estimates of (6.4). □

Proof of Lemma 4.10. The argument will be a discrete analogue of the previous proof. First, notice that for any $y \in Y_M$,

$$|x - x_\perp|^2 \le |x - NP^t y|^2,$$

since $\operatorname{im} NP^t = \ker P^\perp$. Consider the L^2-orthogonal projection onto X_N of the B-spline function B_m given by (6.2),

$$\beta^m := NP^t B_m.$$

Denote $\beta^m = (\beta_1^m, \beta_2^m, \cdots, \beta_N^m)$. The properties of B_m in Remark 6.3 then imply the similar for β^m:

- $0 \le \beta_n^m < 1$,
- $\beta_n^m \ne 0$ only for $n \in ((m-2)K, (m+1)K]$, and
- $\sum_{m=1}^{M} \beta_n^m = 1$.

For each $x \in X_N$, define a projected spline interpolation

$$I(x) := \sum_{j=1}^{M} \underbrace{\left(\frac{1}{3K} \sum_{i=(m-2)K+1}^{(m+1)K} x_i\right)}_{\hat{x}_m} \beta^m,$$

where $\hat{x}_m$ is the *mean spin* over the $3K$-site block $((m-2)K,(m+1)K]$. Using the fact that the β^m are supported on these blocks and sum to 1, we obtain for $n \in ((m-1)K, mK]$,

$$(x_n - I(x)_n)^2 = \left(\sum_{j=-1}^{1} (x_n - \hat{x}_{m+j}) \beta_n^{m+j} \right)^2 \leq \sum_{j=-1}^{1} (x_n - \hat{x}_{m+j})^2,$$

where we used the Cauchy-Schwarz inequality. Summing over n and rearranging leads to

$$\begin{aligned} \sum_{n=1}^{N} (x_n - I(x)_n)^2 &\leq \sum_{m=1}^{M} \sum_{n=(m-2)K+1}^{(m+1)K} (x_n - \hat{x}_m)^2 \\ &\overset{(2.7)}{\leq} \sum_{m=1}^{M} C(3K)^2 \sum_{n=(m-2)K+1}^{(m+1)K} (x_n - x_{n-1})^2 \\ &\leq 3^3 C K^2 \sum_{n=1}^{N} (x_n - x_{n-1})^2 \overset{(2.5)}{=} 3^3 C \frac{K^2}{N^2} x \cdot A x, \end{aligned}$$

where, in the second line, we used the discrete Poincaré inequality from Lemma 2.7 on the $3K$-site block $((m-2)K,(m+1)K]$. Thus (4.13) holds with $\gamma = 3^3 C$, where C is the constant in (2.7). □

6.2 Spline approximations involving the operator $\bar{A}$

In this section we prove those auxiliary results which make precise the idea that the coarse-grained operator $-\bar{A}$ is like a discrete version of the second derivative adapted to the spline space Y_M, namely Lemma 3.13, Lemma 4.13, and Lemma 5.4. We begin with showing that the H^1 inner product on Y_M is close to the inner product induced by the positive definite operator $\bar{A}$.

Lemma 6.4. *There exists an integer K^* such that for all $K \geq K^*, M$ and all $y, \tilde{y} \in Y_M$*

$$\begin{aligned} |\langle \tilde{y}, \bar{A} y \rangle_{L^2} - \langle \tilde{y}, y \rangle_{H^1}| &\lesssim \frac{1}{N} \left(|\tilde{y}|_{H^1} |y|_{H^2} + |\tilde{y}|_{H^2} |y|_{H^1} \right) &(6.8) \\ &\lesssim \frac{M}{N} |\tilde{y}|_{H^1} |y|_{H^1}. &(6.9) \end{aligned}$$

This result leads to a quick proof of Lemma 3.13: the equivalence of $\bar{A}$ and H^1 norms is a direct consequence of the estimate (6.9), from which the equivalence of $\bar{A}^{-1}$ and H^{-1} norms follows by a duality argument with the help of estimate (3.2) that bounds the projection P in H^1 norm.

To prove Lemma 6.4, we need to do some computations involving finite differences. Let

- Q be the L^2-orthogonal projection onto X_N (cf. (2.14)),
- D be the rescaled $N \times N$ forward difference matrix, satisfying that

$$(Dx)_i = N(x_{i+1} - x_i),$$

- ∂_θ^h be the difference quotient

$$\partial_\theta^h y(\theta) = \frac{y(\theta + h) - y(\theta)}{h} = \frac{1}{h} \int_\theta^{\theta+h} y'(s) ds, \qquad (6.10)$$

  which is also a *moving average* of the derivative.

Lemma 6.5. *It holds that $Q = NP^t$ on Y_M, $D^tD = A$, and*

$$DQ = Q\partial_\theta^{\frac{1}{N}}, \quad D^tQ = -Q\partial_\theta^{-\frac{1}{N}}. \tag{6.11}$$

We omit the proof of Lemma 6.5, which can be checked by a straightforward calculation.

Proof of Lemma 6.4. We will prove the estimate (6.8), and then the estimate (6.9) follows from (6.8) by an inverse Sobolev inequality (6.1). For $\tilde{y}, y \in Y_M$,

$$\begin{aligned}\langle \tilde{y}, \bar{A}y\rangle_{L^2} &= \langle NP^t\tilde{y}, A(NP^ty)\rangle_{L^2} \\ &= \langle Q\tilde{y}, D^tDQy\rangle_{L^2} \overset{(6.11)}{=} \langle \partial_\theta^{\frac{1}{N}}\tilde{y}, Q\partial_\theta^{\frac{1}{N}}y\rangle_{L^2}.\end{aligned}$$

Thus, we can decompose

$$\begin{aligned}\langle \tilde{y}, \bar{A}y\rangle_{L^2} - \langle \tilde{y}, y\rangle_{H^1} &= \langle \partial_\theta^{\frac{1}{N}}\tilde{y}, Q\partial_\theta^{\frac{1}{N}}y\rangle_{L^2} - \langle \tilde{y}', y'\rangle_{L^2} \\ &= \langle \partial_\theta^{\frac{1}{N}}\tilde{y} - \tilde{y}', Q\partial_\theta^{\frac{1}{N}}y\rangle_{L^2} + \langle \tilde{y}', Q(\partial_\theta^{\frac{1}{N}}y - y')\rangle_{L^2} + \langle \tilde{y}', Qy' - y'\rangle_{L^2}.\end{aligned}$$

This yields estimate (6.8) once we establish the following estimates:

$$|\partial_\theta^{\frac{1}{N}}y|_{L^2} \le |y'|_{L^2}, \quad |\partial_\theta^{\frac{1}{N}}y - y'|_{L^2} \le \frac{1}{N}|y''|_{L^2}, \quad |Qy' - y'|_{L^2} \lesssim \frac{1}{N}|y''|_{L^2}.$$

The last estimate follows from Poincaré inequality on the sub-intervals $(\frac{i-1}{N}, \frac{i}{N})$, since the projection Q takes *average* over these sub-intervals. The second estimate is similar, as $\partial_\theta^{\frac{1}{N}}y$ is a *moving average* of $\partial_\theta y$ over length $1/N$. The estimate then says taking moving average reduces the L^2 norm, which is a consequence of the Cauchy-Schwarz inequality. □

Now, we turn to the verification of Lemma 4.13.

Proof of Lemma 4.13. The statement is equivalent to

$$|PANP^ty|_{L^2} \ge \sigma|ANP^ty|_{L^2}. \tag{6.12}$$

Given $y \in Y_M$, let $z = ANP^ty$. Assume $z \ne 0$. We want to show $|Pz|_{L^2} \ge \sigma|z|_{L^2}$. For this purpose, it suffices to construct a spline approximation $I(z) \ne 0$ in Y_M satisfying

$$\langle z, I(z)\rangle_{L^2} \ge \sigma|z|_{L^2}|I(z)|_{L^2}, \tag{6.13}$$

since we always have

$$|Pz|_{L^2}|I(z)|_{L^2} \ge \langle Pz, I(z)\rangle_{L^2} = \langle z, I(z)\rangle_{L^2},$$

and the desired conclusion follows. Let us begin by computing $z = ANP^ty$. Applying Lemma 6.5,

$$z = D^tDQy = D^tQ\partial_\theta^{\frac{1}{N}}y = -Q\partial_\theta^{-\frac{1}{N}}\partial_\theta^{\frac{1}{N}}y.$$

Denote $z = (z_1, z_2, \cdots, z_N)$. Using (6.10) and the definition of Q, we compute that

$$\begin{aligned}\partial_\theta^{\frac{1}{N}}y(\theta) &= \fint_0^{\frac{1}{N}} y'(\theta + s)\,ds, \\ \partial_\theta^{-\frac{1}{N}}\partial_\theta^{\frac{1}{N}}y(\theta) &= \fint_{-\frac{1}{N}}^0 \fint_0^{\frac{1}{N}} y''(\theta + t + s)\,ds\,dt, \\ z_n &= -\fint_{\frac{n-1}{N}}^{\frac{n}{N}} \fint_{-\frac{1}{N}}^0 \fint_0^{\frac{1}{N}} y''(\theta + t + s)\,ds\,dt\,d\theta,\end{aligned} \tag{6.14}$$

where $\fint$ is the symbol for average of an integral. Since Y_M consists of quadratic splines,

$$y''(\theta) \equiv \alpha_m \tag{6.15}$$

for some $\alpha_m \in \mathbb{R}$ on the interval $(\frac{m-1}{M}, \frac{m}{M})$. Evaluating the integral (6.14) then gives

$$z_n = -\alpha_m, \text{ for } n = (m-1)K + 2, ..., mK - 1, \tag{6.16}$$

$$z_{(m-1)K+1} = -\alpha_m - \frac{1}{6}(\alpha_{m-1} - \alpha_m), \ \ z_{mK} = -\alpha_m - \frac{1}{6}(\alpha_{m+1} - \alpha_m). \tag{6.17}$$

In particular, $z = ANP^t y$ is almost piecewise constant on the K-site blocks $((m-1)K, mK]$. This motivates us to choose $I(z)$ as the B-spline interpolation

$$I(z) := \sum_{m=1}^{M} -\alpha_m B_m. \tag{6.18}$$

It remains to verify (6.13) for this choice of $I(z)$. Denote $\alpha := (\alpha_1, \alpha_2, \cdots, \alpha_m) \in \mathbb{R}^m$. It follows from (6.16) and (6.17) that

$$|z|_{L^2}^2 = \frac{1}{N}|z|^2 \leq \frac{K}{N}\sum_{m=1}^{M} \alpha_m^2 = \frac{1}{M}|\alpha|^2.$$

Using the property of B-spline functions in (6.3), we compute

$$|I(z)|_{L^2}^2 \overset{(6.18)}{=} \sum_{j,k=1}^{M} \alpha_j \alpha_k \langle B_j, B_k\rangle_{L^2} = \frac{1}{M}\, \alpha \cdot B\alpha \leq \frac{1}{M}|\alpha|^2,$$

where the last inequality is because the operator norm $\|B\| \leq 1$. Finally, we compute

$$\begin{aligned}
\langle z, I(z)\rangle_{L^2} &= \sum_{m=1}^{M} -\alpha_m \,\langle z, B_m\rangle_{L^2} \\
&= \sum_{m,k=1}^{M} \alpha_m \alpha_k \int_{\frac{k-1}{M}}^{\frac{k}{M}} B_m(\theta)d\theta + \sum_{m,k=1}^{M} \frac{1}{6}\alpha_m(\alpha_{k-1} - \alpha_k) \int_{\frac{k-1}{M}}^{\frac{k-1}{M}+\frac{1}{N}} B_m(\theta)d\theta \\
&\quad + \sum_{m,k=1}^{M} \frac{1}{6}\alpha_m(\alpha_{k+1} - \alpha_k) \int_{\frac{k}{M}-\frac{1}{N}}^{\frac{k}{M}} B_m(\theta)d\theta.
\end{aligned}$$

Evaluating this expression yields

$$\langle z, I(z)\rangle_{L^2} = \frac{1}{M}\, \alpha \cdot E\alpha + O\left(\frac{1}{N}\right)|\alpha|^2 \geq \frac{c}{M}|\alpha|^2,$$

where E is the symmetric matrix

$$E_{mk} = \frac{2}{3}\delta_{m=k} + \frac{1}{6}\delta_{|m-k|=1}$$

and the last inequality follows from the strict diagonal dominance of E, once K is large enough, for some universal constant $c > 0$. Putting everything together, we arrive at (6.13):

$$\frac{\langle z, I(z)\rangle_{L^2}^2}{|z|_{L^2}^2\, |I(z)|_{L^2}^2} \geq \frac{c^2}{M^2} M \frac{N}{CK} = \frac{c^2}{C}.$$

Using the results developed in this section, we can now quickly verify Lemma 5.4.

Proof of Lemma 5.4. Argument for (5.4): It follows from (6.15) - (6.17) that

$$|(-\partial_\theta^2 y) - ANP^t y|_{L^2}^2 = \sum_{m=1}^{M} \frac{1}{N}\frac{1}{6^2}(\alpha_{m-1}-\alpha_m)^2 + \frac{1}{N}\frac{1}{6^2}(\alpha_{m+1}-\alpha_m)^2$$
$$\lesssim \sum_{m=1}^{M} \frac{1}{N}|\alpha_m|^2 = \frac{1}{K}|-\partial_\theta^2 y|_{L^2}^2.$$

Consequently,

$$|-\partial_\theta^2 y|_{L^2} \le \left(1 + O\left(K^{-\frac{1}{2}}\right)\right)|ANP^t y|_{L^2} \overset{(6.12)}{\lesssim} |PANP^t y|_{L^2}.$$

Argument for (5.5): applying Lemma 6.4 to $\bar{A}^{-1}y, \tilde{y}$ yields

$$|\langle \tilde{y}, y\rangle_{L^2} - \langle \tilde{y}, -\partial_\theta^2 \bar{A}^{-1} y\rangle_{L^2}| = |\langle \tilde{y}, \bar{A}(\bar{A}^{-1}y)\rangle_{L^2} - \langle \tilde{y}, \bar{A}^{-1}y\rangle_{H^1}|$$
$$\overset{(6.8)}{\lesssim} \frac{1}{K}|\bar{A}^{-1}y|_{H^1}|\tilde{y}|_{H^1} \overset{(3.7)}{\lesssim} \frac{1}{K}|y|_{H^{-1}}|\tilde{y}|_{H^1}.$$ □

Part II

Thermodynamical ingredients for the two-scale approach

The remainder of this paper is devoted to the derivation of the main technical ingredients used in Part I, namely Theorem 3.9, Theorem 4.11, and Theorem 5.3. Towards this end, we will from now on *drop the zero-mean constraint* imposed in Part I:

- We replace the hyperplane X_N in (2.6) by the full microscopic space $\mathbb{R}^N$.
- We replace the zero-mean spline space Y_M in Definition 3.2 by the full spline space

$$Y_M := \left\{ y \in C^{L-1}(\mathbb{T}) \,\middle|\, \forall\, m \in [M] : y|_{\left(\frac{m-1}{M}, \frac{m}{M}\right)} \text{ polynomial of degree} \le \mathrm{L} \right\}.$$

  and the coarse-graining operator $P : L^2(\mathbb{T}) \to Y_M$ is now the L^2-orthogonal projection onto the full spline space Y_M.
- We replace the subspace $L_0^2(\mathbb{T})$ of functions of mean zero by the full space $L^2(\mathbb{T})$.

With these changes in place, we now define the coarse-grained Hamiltonian $\bar{H}(y)$ on the full spline space Y_M as

$$\bar{H}(y) := -\frac{1}{N}\log \int_{x\in\mathbb{R}^N : Px = y} \exp(-H_N(x))\,\mathcal{L}^{N-M}(dx)$$

and the conditional Gibbs measure for arbitrary spline profile y as

$$\mu(dx|y) := \frac{1}{Z}\,\mathbb{1}_{\{x\in\mathbb{R}^N : Px = y\}}(x)\ \exp(-H(x))\,\mathcal{L}^{N-M}(dx).$$

These definitions *extend* the definitions of $\bar{H}$ in (3.5) and $\mu(dx|y)$ in (3.4), i.e. they agree for spline profiles y of mean zero. In this case, $Px = y$ implies x has mean zero, i.e. $x \in X_N$, because P is now defined to be the projection onto the *full* spline space Y_M that includes all constant functions. Therefore $\{x \in R^N : Px = y\}$ appearing in the integrals above coincides with $\{x \in X_N : Px = y\}$, and the Lebesgue measure is taken over the same Euclidean space as in (3.5) an (3.4).

As a result, Theorem 3.9, Theorem 4.11, and Theorem 5.3 readily follow from their respective versions in the present setting of *unrestricted mean*:

- In Theorem 3.9, an upper/lower bound for $\operatorname{Hess}\bar{H}$ as a quadratic form on the full spline space clearly implies the same upper/lower bound on the subspace of zero-mean splines.
- In Theorem 4.11, uniform LSI for $\mu(dx|y)$ over all spline profiles y clearly implies the same over spline profiles y of mean zero.
- In Theorem 5.3, an upper bound for $|\nabla\bar{H} - \nabla\mathcal{H}|$ with gradient taken along the full spline space and the full space $L^2(\mathbb{T})$ clearly implies the same upper bound when gradients are taken along the subspace of zero-mean splines and the subspace $L^2_0(\mathbb{T})$, because the gradient on a subspace is an orthogonal projection of the gradient on the full space.

7 An auxiliary mesoscopic space

As mentioned in Remark 3.10, because the spline functions in Y_M are non-local, we will instead work with a larger mesoscopic space Y_M^{DG} consisting of *discontinuous Galerkin functions* and transfer results back to the spline space Y_M afterwards.

Definition 7.1 (The coarse-graining operator Q_M)**.** *For $M \in \mathbb{N}$, let Y_M^{DG} be the space of discontinuous Galerkin functions of degree $L \in \mathbb{N}$ on the torus $\mathbb{T} = [0,1]$ corresponding to the mesh $\{\frac{m}{M}\}_{m\in[M]}$. That is*

$$Y_M^{DG} := \left\{ y \in L^2(\mathbb{T}) \,\middle|\, \forall\, m \in [M] : y|_{\left(\frac{m-1}{M},\frac{m}{M}\right)} \text{ polynomial of degree} \leq L \right\}. \tag{7.1}$$

We endow Y_M^{DG} with the inner product inherited from $L^2(\mathbb{T})$. We define the coarse-graining operator $Q_M : L^2(\mathbb{T}) \to Y_M^{DG}$ as the L^2-orthogonal projection onto Y_M^{DG}.

Remark 7.2. Since splines are piecewise polynomials satisfying additional constraint, Y_M is a subspace of Y_M^{DG}.

Now, we define an adjoint operator $Q_M^t : Y_M^{DG} \to \mathbb{R}^N$ by

$$\langle Q_M x, y\rangle_{L^2} = x \cdot Q_M^t y \quad \forall\, x \in \mathbb{R}^N,\ y \in Y_M^{DG}.$$

It follows that $NQ_M^t : Y_M^{DG} \to \mathbb{R}^N$ is the L^2-orthogonal projection of Y_M^{DG} onto $\mathbb{R}^N$, where we identify $\mathbb{R}^N$ as a space of step functions on $\mathbb{T}$ in the same fashion as in (2.14) (but without the mean zero constraint).

Lemma 7.3. *It holds that*

$$\|Q_M N Q_M^t - \mathrm{id}_{Y_M^{DG}}\| \lesssim \frac{M^2}{N^2}. \tag{7.2}$$

We omit the proof of Lemma 7.3, which is analogous to the proof of Lemma 3.6 in Section 6. Let us record an easy corollary of this estimate for later use.

Corollary 7.4. *For all $y \in Y_M^{DG}$,*

$$|NQ_M^t y|_{L^2}^2 = \left(1 + O\left(\frac{M^2}{N^2}\right)\right) |y|_{L^2}^2, \tag{7.3}$$

and for all $x \in \operatorname{im} NQ_M^t$,

$$|Q_M x|_{L^2}^2 = \left(1 + O\left(\frac{M^2}{N^2}\right)\right) |x|_{L^2}^2. \tag{7.4}$$

Proof of Corollary 7.4. Estimate (7.3) follows from

$$|NQ_M^t y|_{L^2}^2 = \langle Q_M N Q_M^t y, y\rangle_{L^2} \overset{(7.2)}{=} \left(1 + O\left(\frac{M^2}{N^2}\right)\right) |y|_{L^2}^2.$$

Estimate (7.4) then similarly follows after making the substitution $x = NQ_M^t y$. □

From now on, we assume $N = KM$ for $K \in \mathbb{N}$ large enough so that $Q_M N Q_M^t : Y_M^{DG} \to Y_M^{DG}$ is invertible. In particular, this means the coarse-graining operator $Q_M : \mathbb{R}^N \to Y_M^{DG}$ has full range and the orthogonal projection $NQ_M^t : Y_M^{DG} \to \mathbb{R}^N$ is an embedding.

Definition 7.5 (The coarse-grained Hamiltonian $\bar{H}_{Y_M^{DG}}$)**.** *The coarse-grained Hamiltonian* $\bar{H}_{Y_M^{DG}} : Y_M^{DG} \to \mathbb{R}$ *associated to* Q_M *is given by*

$$\bar{H}_{Y_M^{DG}}(y) := -\frac{1}{N} \log \int_{\{x \in \mathbb{R}^N : Q_M x = y\}} \exp\left(-H_N(x)\right) dx. \tag{7.5}$$

where $H_N(x)$ *is the microscopic Hamiltonian defined in* (2.1)*.*

Let us now relate $\bar{H}_{Y_M^{DG}}$ back to $\bar{H}$. Since $Y_M \subset Y_M^{DG}$, the (restricted) orthogonal projection $P : Y_M^{DG} \to Y_M$ yields the orthogonal decomposition

$$Y_M^{DG} = Y_M \oplus Y_M^{\perp}, \quad Y_M^{\perp} = \{y \in Y_M^{DG} : Py = 0\}.$$

In particular, for any $x \in \mathbb{R}^N$, the decomposition above gives

$$Q_M x = Px \oplus z$$

for some (unique) $z \in Y_M^{\perp}$, which means

$$Px = y \iff Q_M x - y \in Y_M^{\perp}.$$

This relation allows us to view $\bar{H}$ as a coarse-grained version of $\bar{H}_{Y_M^{DG}}$ with the help of the *coarea formula* applied to the affine transformation $x \mapsto Q_M x - y$:

$$\begin{aligned} N\bar{H}(y) &= -\log \int_{x \in \mathbb{R}^N : Px = y} \exp(-H_N(x))\, dx \\ &= -\log \int_{z \in Y_M^{\perp}} \left(\int_{x \in \mathbb{R}^N : Q_M x - y = z} \exp(-H_N(x))\, dx \right) \mathcal{J}_{Q_M}^{-1}\, \mathcal{L}^{LM}(dz) \\ &= -\log \int_{z \in Y_M^{\perp}} \exp(-N\bar{H}_{Y_M^{DG}}(y+z)) \mathcal{L}^{LM}(dz) + \log \mathcal{J}_{Q_M}, \end{aligned} \tag{7.6}$$

where $\mathcal{J}_{Q_M} := \sqrt{\det Q_M Q_M^t}$ is the *constant* factor accounting for the volume change, and LM is the dimension of $Y_M^{\perp}$.

7.1 Reduction to one block

The advantage of working with discontinuous Galerkin functions is that everything can now be decomposed in a block-by-block manner. Let us first set up some notation.

- Given $\alpha \in Y_M^{DG}$, denote by $\alpha^{(m)} \in Y_1^{DG}$ the function obtained from restricting α to the subinterval $[\frac{m-1}{M}, \frac{m}{M})$, i.e.

$$\alpha^{(m)}(\theta) := \alpha\left(\frac{m-1}{M} + \frac{1}{M}\theta\right), \quad \theta \in [0,1). \tag{7.7}$$

- Given $x \in \mathbb{R}^N$, for $1 \le m \le M$, denote

$$x^{(m)} := (x_{(m-1)K+1}, \dots, x_{mK}) \in \mathbb{R}^K.$$

- Denote by $\psi_K := \bar{H}_1^{DG}$ the coarse-grained Hamiltonian for *one block*, i.e. for $\beta \in Y_1^{DG}$,

$$\psi_K(\beta) = -\frac{1}{K} \log \int_{\{x \in \mathbb{R}^K : Q_1 x = \beta\}} \exp(-H_K(x))\, dx.$$

- Let μ^N denote the Gibbs measure on $\mathbb{R}^N$ associated to the Hamiltonian H,

$$\frac{d\mu^N}{d\mathcal{L}^N}(x) = \frac{1}{Z}\exp\left(-H_N(x)\right).$$

Lemma 7.6 (Block-by-block decomposition)**.** *The space Y_M^{DG} decomposes as*

$$Y_M^{DG} = \bigoplus_{m=1}^{M} Y_1^{DG}$$

via the identification

$$\alpha = \bigoplus_{m=1}^{M} \alpha^{(m)}, \quad \langle \alpha, \beta\rangle_{Y_M^{DG}} = \frac{1}{M}\sum_{m=1}^{M} \langle \alpha^{(m)}, \beta^{(m)}\rangle_{Y_1^{DG}}.$$

Consequently, we have the decompositions

$$\begin{aligned} Q_M x &= \bigoplus_{m=1}^{M} Q_1 x^{(m)}, \\ \mu^N(dx|Q_M x = \alpha) &= \bigotimes_{m=1}^{M} \mu^K(dx^{(m)}|Q_1 x^{(m)} = \alpha^{(m)}), \qquad (7.8) \\ \bar{H}_{Y_M^{DG}}(\alpha) &= \frac{1}{M}\sum_{m=1}^{M} \psi_K(\alpha^{(m)}). \end{aligned}$$

8 Convexification of the coarse-grained Hamiltonian

This section is devoted to the proof of Theorem 3.9 that says $\bar{H}$ is uniformly strongly convex. The starting point of the proof is the coarse-graining relation (7.6) that says $\bar{H}$ is a coarse-grained version of $\bar{H}_{Y_M^{DG}}$. This allows us to transfer convexity of $\bar{H}_{Y_M^{DG}}$ to $\bar{H}$, because log-concavity of measures is preserved by marginalization. More precisely, we will invoke the following quantitative version of this well-known fact.

Lemma 8.1. *Let $W \oplus Z$ be an orthogonal decomposition of a finite dimensional Euclidean space. Suppose $F : W \oplus Z \to \mathbb{R}$ is a C^2 function such that $\int_{W\oplus Z}\exp(-F) < \infty$. Let $\bar{F}(z) := -\log\int_W \exp(-F(w,z))\,dw$. For any $c \geq 0$, it holds that*

$$\begin{aligned} \operatorname{Hess}_{W\oplus Z} F \geq c\,\mathrm{id}_{W\oplus Z} \;&\Rightarrow\; \operatorname{Hess}_Z \bar{F} \geq c\,\mathrm{id}_Z, \qquad (8.1) \\ \operatorname{Hess}_{W\oplus Z} F \leq c\,\mathrm{id}_{W\oplus Z} \;&\Rightarrow\; \operatorname{Hess}_Z \bar{F} \leq c\,\mathrm{id}_Z\,. \end{aligned}$$

Remark 8.2. In [7], it was shown in a very neat way that statement (8.1) is simple consequence of the well-known Brascamp-Lieb inequality. For completeness, we provide a short proof of this fact in Appendix A using the Prékopa-Leindler inequality from convex geometry.

Applying Lemma 8.1 with $Z = Y_M$, $W = Y_M^{\perp}$, and $F = N\bar{H}_{Y_M^{DG}}$, proving Theorem 3.9 reduces to showing the uniform strong convexity of $\bar{H}_{Y_M^{DG}}$ stated below.

Theorem 8.3 (Uniform strong convexity of $\bar{H}_{Y_M^{DG}}$)**.** *There are constants $0 < \lambda, \Lambda, K^* < \infty$ such that for all $K \geq K^*$, M and all $y \in Y_M^{DG}$ it holds*

$$2\lambda\,\mathrm{id}_{Y_M^{DG}} \leq \operatorname{Hess}_{Y_M^{DG}} \bar{H}_{Y_M^{DG}}(y) \leq 2\Lambda\,\mathrm{id}_{Y_M^{DG}}$$

in the sense of quadratic forms.

In turn, by the block-by-block decomposition from Lemma 7.6, proving Theorem 8.3 reduces to proving the case of just one block.

Theorem 8.4. *There are constants $0 < \lambda, \Lambda, J^* < \infty$ such that for all $J \geq J^*$, and all $\beta \in Y_1^{DG}$ it holds*

$$2\lambda \operatorname{id}_{Y_1^{DG}} \leq \operatorname{Hess}_{Y_1^{DG}} \psi_J(\beta) \leq 2\Lambda \operatorname{id}_{Y_1^{DG}},$$

in the sense of quadratic forms.

The proof of Theorem 8.4 closely follows the proof of [19, Proposition 31]: a local Cramér theorem through a local central limit theorem (CLT). The main difference here is that the local Cramér theorem has to be extended to canonical ensembles with multiple constraints, which means that we will have to use a multivariate CLT.

Before entering into the details, we sketch an outline of the argument. The strategy is to prove ψ_J converges to a strongly convex function $\bar{\psi}_J$ in the uniform C^2-topology as $J \to \infty$. Namely:

- Using Cramér's trick of exponential shift of measure, we construct for each $\beta \in Y_1^{DG}$ a product measure $\nu_{J,\beta}$ on $\mathbb{R}^J$ such that
  - the law of each spin is an "exponential shift" of the single-site measure,
  - the expectation of $Q_1 x$ under $\nu_{J,\beta}$ is equal to β, i.e. the conditioning $Q_1 x = \beta$ is a "typical" event.

  We refer to the product measure $\nu_{J,\beta}$ as the *modified grand canonical ensemble* for β. The required shifts of spins can be parameterized by a variable $\hat{\beta}$ that is dual to β.
- Due to the form of the single-site potential ψ, it follows that once J is large enough, the *specific free energy* $\bar{\psi}_J^*$ of $\nu_{J,\beta}$ is a strongly convex function of $\hat{\beta}$ with bounded Hessian. Consequently, the same is true for its Legendre transform $\bar{\psi}_J$.
- Moreover, the difference $\bar{\psi}_J(\beta) - \psi_J(\beta)$ can be interpreted as the difference between the specific free energies of $\nu_{J,\beta}$ and its restriction to the hyperplane determined by $Q_1 x = \beta$ (which is the "typical event"). Hence, we expect that this difference goes to zero as J grows large.

To verify that this difference indeed converges to zero in the uniform C^2-topology:

- Using a Cramér-type representation formula (equation (8.15) below), we relate this difference to the density at 0 of the law of the random variable $J^{\frac{1}{2}} (Q_1 x - \beta)$ under $\nu_{J,\beta}(dx)$.
- By establishing a uniform C^2 local central limit theorem, we ensure this density at 0 is bounded from above and from below uniformly in β and that moreover, it is bounded in the uniform C^2 norm as a function of β. These estimates are stated in Proposition 8.10 below and constitute the core of our proof.

8.1 Construction of the modified grand canonical ensembles $\nu_{J,\beta}$

We begin by introducing a family of "exponential shifts" of the law of a single spin. Let ψ^* be the *log partition function* of the single site potential ψ that was defined in (2.16). For $m \in \mathbb{R}$, let μ_m be the probability measure on $\mathbb{R}$ with Lebesgue density

$$\frac{d\mu_m}{dz} = \exp(-\psi^*(\sigma) + \sigma z - \psi(z)), \tag{8.2}$$

where σ is the *exponential shift required* for μ_m to have mean m, i.e. $(\psi^*)'(\sigma) = m$. We will need the following estimates of the 2nd and 3rd derivatives of ψ^*.

Lemma 8.5. *[19, Lemma 41] There are constants $0 < c < C < \infty$ such that it holds that:*

$$0 < c < \inf_{m\in\mathbb{R}} \mathrm{Var}(\mu_m) \le (\psi^*)''(\sigma) \le \sup_{m\in\mathbb{R}} \mathrm{Var}(\mu_m) < C < \infty, \tag{8.3}$$

$$|(\psi^*)'''(\sigma)| \le \sup_{m\in\mathbb{R}} \left| \int (z-m)^3 \mu_m(dz) \right| < C < \infty. \tag{8.4}$$

where $\mathrm{Var}(\mu_m)$ *denotes the variance of* μ_m*.*

To construct the product measure $\nu_{J,\beta}$ on $\mathbb{R}^J$, we will find a suitable dual variable $\hat\beta \in Y_1^{DG}$ and exponentially shift the law of the J spins according to the J entries of the vector $\hat x = JQ_1^t\hat\beta$. The right choice will be found from the following functions that are to be interpreted as the (specific) free energies of the ensemble $\nu_{J,\beta}$:

- We associate each $\hat x \in \mathbb{R}^J$ with a *free energy*

  $$\psi_J^*(\hat x) := \sum_{j=1}^J \psi^*(\hat x_j).$$

  Here the variable $\hat x$ should be viewed as "dual" to the spin configuration $x \in \mathbb{R}^J$.
- Using the projection operator $JQ_1^t : Y_1^{DG} \to \mathbb{R}^J$, we then associate each $\hat\beta \in Y_1^{DG}$ with a *specific free energy*

  $$\bar\psi_J^*(\hat\beta) := \frac{1}{J}\psi_J^*(JQ_1^t\hat\beta) = \frac{1}{J}\sum_{j=1}^J \psi^*\left((JQ_1^t\hat\beta)_j\right). \tag{8.5}$$

  Here the variable $\hat\beta$ should be viewed as "dual" to a variable $\beta \in Y_1^{DG}$. The specific free energy as a function of the variable β is then given by the Legendre transform,

  $$\bar\psi_J(\beta) := \sup_{\hat\beta\in Y_1^{DG}} \left(\langle\beta,\hat\beta\rangle \;-\; \bar\psi_J^*(\hat\beta)\right). \tag{8.6}$$

After dealing with the approximation error, the uniform convexity and C^2 bounds (8.3) transfer to the function $\bar\psi_J^*(\hat\beta)$ and, consequently, its Legendre transform $\bar\psi_J(\beta)$:

Lemma 8.6. *There is $J_1 \in \mathbb{N}$ such that for all $J \ge J_1$ and $\hat\beta, \beta \in Y_1^{DG}$:*

$$\|\,\mathrm{Hess}\ \bar\psi_J^*(\hat\beta)\| \simeq 1, \tag{8.7}$$

$$\|\,\mathrm{Hess}\ \bar\psi_J(\beta)\| \simeq 1. \tag{8.8}$$

Proof. For any $\hat x, \hat\xi \in \mathbb{R}^J$,

$$\hat\xi \cdot \mathrm{Hess}\ \psi_J^*(\hat x)\,\hat\xi = \sum_{j=1}^J (\psi^*)''(\hat x_j)\,|\hat\xi_j|^2 \overset{(8.3)}{\simeq} |\hat\xi|^2$$

where the derivatives are taken w.r.t. the Euclidean structure of $\mathbb{R}^J$. After accounting for the different Euclidean structure on $Y_1^{DG} \subseteq L^2(\mathbb{T})$, we obtain that for $\hat\eta \in Y_1^{DG}$,

$$\langle\hat\eta, \mathrm{Hess}\ \bar\psi_J^*(\hat\beta)\,\hat\eta\rangle = \langle JQ_1^t\hat\eta, \mathrm{Hess}\ \psi_J^*(JQ_1^t\hat\beta)\,JQ_1^t\hat\eta\rangle_{L^2} \simeq |JQ_1^t\hat\eta|_{L^2}^2 \overset{(7.3)}{=} \left(1+\frac{1}{J^2}\right)|\hat\eta|_{L^2}^2.$$

This establishes (8.7), which in turn yields (8.8) by a standard calculation. □

Let $\hat{\beta}^{max}(\beta)$ be the unique maximizer of (8.6), which exists by the convexity of $\bar{\psi}_J^*$. It satisfies

$$\beta \;=\; \nabla\,\bar{\psi}_J^*(\hat{\beta}^{max}) \overset{(8.5)}{=} Q_1 \nabla \psi_J^*(J Q_1^t \hat{\beta}^{max}). \tag{8.9}$$

The vector $\hat{\beta}^{max}$ serves to construct $\nu_{J,\beta}$. Set

$$\hat{m}(\beta) := J Q_1^t \hat{\beta}^{max}, \quad \hat{m}_j(\beta) := (\hat{m}(\beta))_j, \tag{8.10}$$

$$m(\beta) := \nabla \psi_J^*(\hat{m}(\beta)), \quad m_j(\beta) := (m(\beta))_j, \tag{8.11}$$

and define a product measure on $\mathbb{R}^J$ with Lebesgue density

$$\begin{aligned}\frac{d\nu_{J,\beta}}{dx}(x) \;:=\; \prod_{j=1}^{J} \frac{d\mu_{m_j(\beta)}}{dx_j}(x_j) &\overset{(8.2)}{=} \prod_{j=1}^{J} \exp(-\psi^*(\hat{m}_j(\beta)) + \hat{m}_j(\beta) x_j - \psi(x_j)) \\ &=\; \exp(-\psi_J^*(\hat{m}(\beta)) + \hat{m}(\beta)\cdot x - H_J(x)).\end{aligned} \tag{8.12}$$

Then $\nu_{J,\beta}$ has mean $m(\beta) = \nabla\psi_J^*(\hat{m}(\beta))$ and the expected value of $Q_1 x$ under $\nu_{J,\beta}(dx)$ is

$$Q_1 \nabla \psi_J^*(\hat{m}(\beta)) \overset{(8.10)}{=} Q_1 \nabla \psi_J^*(J Q_1^t \hat{\beta}^{max}) \overset{(8.9)}{=} \beta.$$

This completes the construction of the modified grand canonical ensemble $\nu_{J,\beta}$.

8.2 Uniform C^2 convergence of $\bar{\psi}_J - \psi_J$ to zero

For a given β, the specific free energy of the modified grand canonical ensemble $\nu_{J,\beta}$ is just

$$\bar{\psi}_J^*(\hat{\beta}^{max}) \overset{(8.6)}{=} \langle \beta, \hat{\beta}^{max}\rangle \;-\; \bar{\psi}_J(\beta). \tag{8.13}$$

On the other hand, the specific free energy of the canonical ensemble associated with the restriction of $\nu_{J,\beta}$ to the hyperplane $\{x \,|\, Q_1 x \;=\; \beta\}$, where it is highly concentrated anyway for large J by the usual law of large numbers, is given by

$$\begin{aligned}&\frac{1}{J}\log \int_{Q_1 x = \beta} \exp(\hat{m}(\beta)\cdot x - H_J(x))\mathcal{H}(dx) \\ &\overset{(8.10)}{=} \frac{1}{J}\log \int_{Q_1 x = \beta} \exp(J\langle Q_1 x, \hat{\beta}^{max}\rangle - H_J(x))\mathcal{H}(dx) \\ &= \langle \beta, \hat{\beta}^{max}\rangle - \psi_J(\beta).\end{aligned} \tag{8.14}$$

Consequently, $\bar{\psi}_J(\beta) - \psi_J(\beta)$ measures the difference in free energies and hence we expect it to converge to zero in some sense as $J \to \infty$. By Lemma 8.6, the function $\bar{\psi}_J$ is strongly convex with bounded Hessian. Therefore, proving Theorem 8.4 reduces to showing the difference $\bar{\psi}_J(\beta) - \psi_J(\beta)$ converges to zero in the uniform C^2 topology:

Proposition 8.7. *There exists J_2 such that for $J \geq J_2$ and for all $\beta \in Y_1^{DG}$,*

$$\begin{aligned}\left|\psi_J(\beta) - \bar{\psi}_J(\beta)\right| &\lesssim \frac{1}{J}, \\ \left\|\nabla\psi_J(\beta) - \nabla\bar{\psi}_J(\beta)\right\| &\lesssim \frac{1}{J}, \\ \|\operatorname{Hess}\psi_J(\beta) - \operatorname{Hess}\bar{\psi}_J(\beta)\| &\lesssim \frac{1}{J}.\end{aligned}$$

As we indicated above, the proof of Proposition 8.7 begins with a Cramér-type representation formula for the density at 0 of the law of $J^{\frac{1}{2}}\,(Q_1 x - \beta)$ under $\nu_{J,\beta}(dx)$, which is a centered $(L{+}1)$-dimensional vector of “suitably weighted” sums of independent random variables (cf. [19, equation (125)]). From now on, we identify Y_1^{DG} with $\mathbb{R}^{L+1}$.

Lemma 8.8. *Denote by $g_{J,\beta}$ the law of the $\mathbb{R}^{L+1}$-valued random variable $J^{\frac{1}{2}}(Q_1 x - \beta)$ where x is the configuration of J spins under $\nu_{J,\beta}$. The density of $g_{J,\beta}$ at $0 \in \mathbb{R}^{L+1}$ with respect to Lebesgue measure can be expressed as*

$$g_{J,\beta}(0) := \frac{dg_{J,\beta}}{d\mathcal{L}^{L+1}}(0) \; = \; (\mathcal{J}Q)^{-1} \; \exp\left[J\left(\bar{\psi}_J(\beta) - \psi_J(\beta)\right)\right], \tag{8.15}$$

where $\mathcal{J}Q := (\det Q_1 J Q_1^t)^{\frac{1}{2}}$ is the Jacobian of the (linear) transformation $x \mapsto J^{\frac{1}{2}}(Q_1 x - \beta)$.

Proof. The level set for $J^{\frac{1}{2}}(Q_1 x - \beta) = u$ is the hyperplane $Q_1 x = \beta + J^{-\frac{1}{2}} u$. Disintegrating measure along these hyperplanes and accounting for volume change, we have

$$g_{J,\beta}(u) = (\mathcal{J}Q)^{-1} \int_{\left\{Q_1 x = \beta + J^{-\frac{1}{2}} u\right\}} \frac{d\nu_{J,\beta}}{d\mathcal{L}^J}(x)\, \mathcal{H}^{J-(L+1)}(dx).$$

For $u = 0$, we compute

$$\begin{aligned} g_{J,\beta}(0) &= (\mathcal{J}Q)^{-1} \int_{Q_1 x = \beta} \frac{d\nu_{J,\beta}}{d\mathcal{L}^J}(x)\, \mathcal{H}(dx) \\ &\overset{(8.12)}{=} (\mathcal{J}Q)^{-1} \exp(-\psi_J^*(\hat{m}(\beta))) \int_{Q_1 x = \beta} \exp(\hat{m}(\beta)\cdot x - H_J(x))\, \mathcal{H}(dx) \\ &\overset{(8.14)}{=} (\mathcal{J}Q)^{-1} \exp(-J\bar{\psi}_J^*(\hat{\beta}^{max})) \exp[J(\langle \beta, \hat{\beta}^{max}\rangle - \psi_J(\beta))] \\ &\overset{(8.13)}{=} (\mathcal{J}Q)^{-1} \exp[J(\bar{\psi}_J(\beta) - \psi_J(\beta))], \end{aligned}$$

where we used that

$$\psi_J^*(\hat{m}(\beta)) \overset{(8.10)}{=} \psi_J^*(JQ_1^t \hat{\beta}^{max}) \overset{(8.5)}{=} J\bar{\psi}_J^*(\hat{\beta}^{max}). \qquad \square$$

Using formula (8.15), Proposition 8.7 basically reduces to the following estimates on the Jacobian $\mathcal{J}Q = \det(Q_1 J Q_1^t)^{\frac{1}{2}}$ (appearing on the right hand side of (8.15)) and the density $g_{J,\beta}(0)$ (appearing on left hand side of (8.15)).

Lemma 8.9. *There exists a constant $C < \infty$ and a positive integer $J_3 \in \mathbb{N}$ such that for $J \geq J_3$:*

$$\frac{1}{C} \; \leq \; \det(Q_1 J Q_1^t)^{\frac{1}{2}} \; \leq \; C. \tag{8.16}$$

This follows from the fact that $Q_1 J Q_1^t$ is close to the identity operator for large J, which is the case $M = 1, N = J$ of Lemma 7.3.

Proposition 8.10. *There exist a constant $C < \infty$ and a positive integer J_4 such that for all $J \geq J_4$ and all $\beta \in \mathbb{R}^{L+1}$:*

$$\frac{1}{C} \; \leq \; g_{J,\beta}(0) \; \leq \; C, \tag{8.17}$$

$$\|\nabla_\beta \, g_{J,\beta}(0)\| \; \leq \; C, \tag{8.18}$$

$$\|\mathrm{Hess}_\beta \; g_{J,\beta}(0)\| \; \leq \; C. \tag{8.19}$$

This result was proven in [19] for the case $L = 0$ (cf. equation (126) in [19]). In the general case considered here, establishing the estimates becomes somewhat more subtle. In particular, a uniform control on the magnitude of the projected spins $(JQ_1^t)_j$ enters as a new ingredient (see Lemma B.3 below). The proof also shares some similarities to the proof of the local Cramér theorem in [27]. As the proof as a whole becomes quite long we postpone it to Appendix B at the end of this article. We conclude with a quick derivation of Proposition 8.7 from these results.

Proof of Proposition 8.7. Rewrite formula (8.15) as:

$$\bar{\psi}_J(\beta) - \psi_J(\beta) \,=\, \frac{1}{J}\,\left[\log\left(\mathcal{J}Q\right) \,+\, \log g_{J,\beta}(0)\right].$$

For $J \,\geq\, \max\{J_3, J_4\}$, the estimates (8.16) - (8.19) thus yield

$$\begin{aligned}
|\bar{\psi}_J(\beta) - \psi_J(\beta)| \,&\leq\, \frac{\log C + \log C}{J} \\
\|\nabla\bar{\psi}_J(\beta) - \nabla\psi_J(\beta)\| &= \frac{1}{J}\left(g_{J,\beta}(0)\right)^{-1}\,\|\nabla_\beta g_{J,\beta}(0)\| \leq \frac{C^2}{J} \\
\|\operatorname{Hess}\bar{\psi}_J(\beta) - \operatorname{Hess}\psi_J(\beta)\| &\leq \frac{1}{J}\left(g_{J,\beta}(0)\right)^{-1}\,\|\operatorname{Hess}_\beta g_{J,\beta}(0)\| \\
&\quad + \frac{1}{J}\left(g_{J,\beta}(0)\right)^{-2}\,\|\,\nabla_\beta g_{J,\beta}(0)\,\otimes\,\nabla_\beta g_{J,\beta}(0)\,\| \\
&\leq \frac{C^2}{J} \,+\, \frac{C^4}{J}.
\end{aligned}$$

□

9 Convergence of the coarse-grained Hamiltonian to the macroscopic free energy

This section is devoted to the proof of Theorem 5.3, i.e. the convergence of $\nabla\bar{H}$ to $\nabla\mathcal{H}$. As in the proof of Theorem 3.9, the argument will go through the intermediate space Y_M^{DG} defined in (7.1). The main ingredient is the local Cramér theorem already established in Proposition 8.7, which says $\psi_K \,:=\, \bar{H}_{Y_1^{DG}}$ converges in C^2 to the free energy $\bar{\psi}_K$ defined in (8.6). By the block-by-block decomposition in Lemma 7.6, this result immediately extends to say that $\bar{H}_{Y_M^{DG}}$ converges in C^2 to a free energy $\mathcal{H}_{Y_M^{DG}}$:

Definition 9.1 (Mesoscopic free energies on Y_M^{DG})**.** *Define $\mathcal{H}_{Y_M^{DG}} : Y_M^{DG} \to \mathbb{R}$ by*

$$\mathcal{H}_{Y_M^{DG}}(z) := \frac{1}{M}\sum_{m=1}^{M}\bar{\psi}_K(z^{(m)}),$$

where $z^{(m)} \in Y_1^{DG}$ are defined in (7.7)*. Equivalently, $\mathcal{H}_{Y_M^{DG}}$ is the Legendre transform of the function $\varphi_N^* : Y_M^{DG} \to \mathbb{R}$ defined by*

$$\varphi_N^*(\hat{z}) : = \frac{1}{M}\sum_{m=1}^{M}\bar{\psi}_K^*(\hat{z}^{(m)}), \tag{9.1}$$

where $\bar{\psi}_K^*$ is the function defined in (8.5)*.*

Corollary 9.2. *There exists K^* such that for $K \geq K^*$ and for all M and $z \in Y_M^{DG}$,*

$$\begin{aligned}
\left|\bar{H}_{Y_M^{DG}}(z) - \mathcal{H}_{Y_M^{DG}}(z)\right| &\lesssim \frac{1}{K}, \\
\left\|\nabla\bar{H}_{Y_M^{DG}}(z) - \nabla\mathcal{H}_{Y_M^{DG}}(z)\right\| &\lesssim \frac{1}{K}, \\
\|\operatorname{Hess}\bar{H}_{Y_M^{DG}}(z) - \operatorname{Hess}\mathcal{H}_{Y_M^{DG}}(z)\| &\lesssim \frac{1}{K}.
\end{aligned} \tag{9.2}$$

The rest of the proof essentially consists of passing from $\bar{H}$ to $\bar{H}_{Y_M^{DG}}$ and passing from to $\mathcal{H}_{Y_M^{DG}}$ to $\mathcal{H}$. Before going into full detail, we give a summary of the main steps:

- To go from $\bar{H}$ to $\bar{H}_{Y_M^{DG}}$, we use the coarse-graining relation (7.6),

$$\bar{H}(y) = -\frac{1}{N}\log\int_{z\in Y_M^{\perp}}\exp(-N\bar{H}_{Y_M^{DG}}(y+z))\mathcal{L}^{LM}(dz) + \frac{1}{N}\log\mathcal{J}_{Q_M}. \tag{9.3}$$

For large K, the strong convexity of $\bar{H}_{Y_M^{DG}}$ means the integral on the right hand side above would concentrate around the minimum of $\bar{H}_{Y_M^{DG}}$,

$$\bar{H}_{Y_M^{DG}}(y+\bar{z}^*) = \inf_{z\in Y_M^\perp} \bar{H}_{Y_M^{DG}}(y+z), \tag{9.4}$$

where $\bar{z}^* \in Y_M^\perp$ is the unique minimizer.

- In parallel to the above, we consider the minimization

$$\mathcal{H}_{Y_M^{DG}}(y+z^*) = \inf_{z\in Y_M^\perp} \mathcal{H}_{Y_M^{DG}}(y+z), \tag{9.5}$$

where $z^* \in Y_M^\perp$ is the unique minimizer, which exists by the strong convexity of $\mathcal{H}_{Y_M^{DG}}$ (see Corollary 9.3 below). The closeness of $\bar{H}_{Y_M^{DG}}$ and $\mathcal{H}_{Y_M^{DG}}$ in C^2 then imply that the minimizers $\bar{z}^*$ and z^* are also close.

- On the other hand, we observe that $\mathcal{H}_{Y_M^{DG}}$ is the Legendre transform of the function φ_N^*. It follows from basic properties of the Legendre transform that the variational problem (9.5) can be reformulated as

$$\mathcal{H}_{Y_M^{DG}}(y+z^*) = \mathcal{H}_{Y_M}(y), \tag{9.6}$$

where $\mathcal{H}_{Y_M} : Y_M \to \mathbb{R}$ is defined to be the Legendre transform of the *same function* φ_N^* *with domain restricted to* Y_M.

- To go from $\mathcal{H}_{Y_M}$ to $\mathcal{H}$, we observe that $\mathcal{H}$ is the Legendre transform of the function $\varphi^* : L^2(\mathbb{T}) \to \mathbb{R}$ defined by

$$\varphi^*(\hat{y}) := \int_0^1 \psi^*(\hat{y}(\theta))d\theta, \tag{9.7}$$

while the restricted function $\varphi_N^* : Y_M \to \mathbb{R}$, whose Legendre transform is $\mathcal{H}_{Y_M}$, can also be expressed as

$$\varphi_N^*(\hat{z}) \overset{(9.1)}{=} \frac{1}{N}\sum_{i=1}^N \psi^*\left(\fint_{\frac{i-1}{N}}^{\frac{i}{N}} \hat{z}(s)ds \right) = \varphi^*(NP^t\hat{z}),$$

where the adjoint operator $NP^t : Y_M \to \mathbb{R}^N$ acts as an L^2 orthogonal projection, i.e. it takes average over blocks of size $K = N/M$.

  - As the block size $K = N/M \to \infty$, the function φ_N^* should converge to the function φ^* *with domain restricted to* Y_M. To track this effect we define another free energy $\hat{\mathcal{H}}_{Y_M} : Y_M \to \mathbb{R}$ as the Legendre transform of the restricted function $\varphi^* : Y_M \to \mathbb{R}$.
  - As the number of blocks $M \to \infty$, the spline space Y_M approximates the full space $L^2(\mathbb{T})$, so $\hat{\mathcal{H}}_{Y_M}$ would approximate $\mathcal{H}$.

To facilitate these approximation arguments, we need the following convexity estimates of the various free energies/coarse-grained Hamiltonians involved. They are direct consequences of the convexity/convexification results of the previous sections.

Corollary 9.3. *Let F denote any one of the following functions:*

$$\bar{H}_{Y_M^{DG}}, \mathcal{H}_{Y_M^{DG}}, \varphi_N^* : Y_M^{DG} \to \mathbb{R},$$
$$\bar{H}, \mathcal{H}_{Y_M}, \hat{\mathcal{H}}_{Y_M} : Y_M \to \mathbb{R}$$
$$\mathcal{H}, \varphi^* : L^2(\mathbb{T}) \to \mathbb{R}.$$

Then there exists K_0 such that for all $K \geq K_0$, F is uniformly strongly convex with uniformly bounded Hessian. Consequently, for all z, w in the domain of F

$$\begin{aligned} |z-w|_{L^2}^2 &\simeq \langle \nabla F(z) - \nabla F(w), z-w\rangle_{L^2} \\ &\simeq |\nabla F(z) - \nabla F(w)|_{L^2}^2. \end{aligned} \tag{9.8}$$

9.1 Closeness of $\nabla \bar{H}$ and $\nabla \bar{H}_{Y_M^{DG}}$

Let $\bar{z}^* \in Y_M^\perp$ be the minimizer of (9.4). From this variational characterization, it follows that $\nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) \in Y_M$. Moreover:

Lemma 9.4. *There exists K^* such that for all $K \geq K^*$ and for all $y \in Y_M$,*

$$|\nabla \bar{H}(y) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*)|_{L^2} \lesssim \frac{1}{K^{\frac{1}{2}}}. \tag{9.9}$$

Proof of Lemma 9.4. Taking gradient in (9.3) yields that

$$\nabla \bar{H}(y) = \int_{Y_M^\perp} P \nabla \bar{H}_{Y_M^{DG}}(y + z)\, \nu(dz|y),$$

where $\nu(dz|y)$ is the conditional Gibbs measure for the Hamiltonian $N\bar{H}_{Y_M^{DG}}$. Therefore

$$\nabla \bar{H}(y) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) = \int_{Y_M^\perp} \Big(P \nabla \bar{H}_{Y_M^{DG}}(y + z) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) \Big)\, \nu(dz|y). \tag{9.10}$$

Using the convexity estimates of Lemma 9.3,

$$\begin{aligned}
\Big| P \nabla \bar{H}_{Y_M^{DG}}(y + z) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) \Big|_{L^2}^2 &\overset{(9.8)}{\lesssim} \langle z - \bar{z}^*, \nabla \bar{H}_{Y_M^{DG}}(y + z) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) \rangle_{L^2} \\
&= \langle z - \bar{z}^*, \nabla \bar{H}_{Y_M^{DG}}(y + z) \rangle_{L^2},
\end{aligned}$$

where we also used the fact that $\nabla H_{Y_M^{DG}}(y + \bar{z}^*) \in Y_M$. After applying Cauchy-Schwarz, using this estimate on (9.10) and integration by parts for the Gibbs measure $\nu(dz|y)$,

$$\begin{aligned}
|\nabla \bar{H}(y) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*)|_{L^2}^2 &\leq \int_{Y_M^\perp} \Big| P \nabla \bar{H}_{Y_M^{DG}}(y + z) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) \Big|_{L^2}^2 \nu(dz|y), \\
&\lesssim \int_{Y_M^\perp} \langle z - \bar{z}^*, \nabla \bar{H}_{Y_M^{DG}}(y + z) \rangle_{L^2}\ \nu(dz|y) \\
&= \frac{1}{N} \int_{Y_M^\perp} \nabla \cdot (z - \bar{z}^*)\, \nu(dz|y) \\
&= \frac{\dim Y_M^\perp}{N} = \frac{LM}{N},
\end{aligned}$$

which implies the desired estimate (9.9). □

9.2 Closeness of $\nabla \bar{H}_{Y_M^{DG}}$ and $\nabla \mathcal{H}_{Y_M^{DG}}$

Let $z^* \in Y_M^\perp$ be the minimizer of (9.5). From this variational characterization, it follows that $\nabla \mathcal{H}_{Y_M^{DG}}(y + z^*) \in Y_M$. Moreover:

Lemma 9.5. *There exists K^* such that for all $K \geq K^*$ and for all $y \in Y_M$,*

$$|\bar{z}^* - z^*|_{L^2} \lesssim \frac{1}{K}, \tag{9.11}$$

$$|\nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) - \nabla \mathcal{H}_{Y_M^{DG}}(y + z^*)|_{L^2} \lesssim \frac{1}{K}. \tag{9.12}$$

Proof. The first estimate follows from

$$\begin{aligned}
|\bar{z}^* - z^*|_{L^2}^2 &\overset{(9.8)}{\lesssim} \langle \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) - \nabla \bar{H}_{Y_M^{DG}}(y + z^*), \bar{z}^* - z^* \rangle_{L^2} \\
&= \langle \nabla \mathcal{H}_{Y_M^{DG}}(y + z^*) - \nabla \bar{H}_{Y_M^{DG}}(y + z^*), \bar{z}^* - z^* \rangle_{L^2} \overset{(9.2)}{\lesssim} \frac{1}{K} |\bar{z}^* - z^*|_{L^2},
\end{aligned}$$

where we used the fact that $\nabla \bar{H}_{Y_M^{DG}}(y+\bar{z}^*), \nabla \mathcal{H}_{Y_M^{DG}}(y+z^*) \in Y_M$ in the second step. The second estimate then follows from combining the estimates

$$|\nabla \bar{H}_{Y_M^{DG}}(y+\bar{z}^*) - \nabla \bar{H}_{Y_M^{DG}}(y+z^*)|_{L^2} \overset{(9.8)}{\lesssim} |\bar{z}^* - z^*|_{L^2} \overset{(9.11)}{\lesssim} \frac{1}{K},$$
$$|\nabla \bar{H}_{Y_M^{DG}}(y+z^*) - \nabla \mathcal{H}_{Y_M^{DG}}(y+z^*)|_{L^2} \overset{(9.2)}{\lesssim} \frac{1}{K}. \qquad \square$$

9.3 Closeness of $\nabla \mathcal{H}_{Y_M^{DG}}$ and $\nabla \mathcal{H}_{Y_M}$

We now verify (9.6) as well as the relation

$$\nabla \mathcal{H}_{Y_M}(y) = \nabla \mathcal{H}_{Y_M^{DG}}(y+z^*). \tag{9.13}$$

As observed before, $\mathcal{H}_{Y_M^{DG}}$ and $\mathcal{H}_{Y_M}$ are Legendre transforms of the same function φ_N^* with different domains:

$$\begin{aligned}\mathcal{H}_{Y_M^{DG}}(y+z^*) &= \sup_{\hat{z} \in Y_M^{DG}} \langle y+z^*, \hat{z}\rangle_{L^2} - \varphi_N^*(\hat{z}) \\ &\geq \sup_{\hat{z} \in Y_M} \langle y, \hat{z}\rangle_{L^2} - \varphi_N^*(\hat{z}) = \mathcal{H}_{Y_M}(y).\end{aligned}$$

By basic properties of Legendre transform, the maximization problem in the first line has the unique solution $\hat{z} = \nabla \mathcal{H}_{Y_M^{DG}}(y+z^*)$ and the maximization problem in the second line has the unique solution $\hat{z} = \nabla \mathcal{H}_{Y_M}(y)$. But since $\nabla \mathcal{H}_{Y_M^{DG}}(y+z^*) \in Y_M$, it must also solve the maximization problem in the second line. This establishes both (9.6) and (9.13).

9.4 Closeness of $\nabla \mathcal{H}_{Y_M}$ and $\nabla \mathcal{H}$

Recall that we define $\hat{\mathcal{H}}_{Y_M} : Y_M \to \mathbb{R}$ to be the Legendre transform of the function φ^* defined in (9.7) *with domain restricted to* Y_M, i.e.

$$\hat{\mathcal{H}}_{Y_M}(y) = \sup_{\hat{y} \in Y_M} \left(\langle y, \hat{y}\rangle_{L^2} - \varphi^*(\hat{y})\right).$$

We need to show the gradients of $\mathcal{H}_{Y_M}$ and $\hat{\mathcal{H}}_{Y_M}$ are close for large K. This will follow from an analogous estimate for the functions φ_N^* and φ^*.

Lemma 9.6. *There exists K^* such that for all $K \geq K^*$ and all $\hat{y} \in Y_M$*

$$|\nabla_{Y_M} \varphi_N^*(\hat{y}) - \nabla_{Y_M} \varphi^*(\hat{y})|_{L^2} \lesssim \frac{1}{K} |\hat{y}|_{L^2}. \tag{9.14}$$

Here ∇_{Y_M} indicates the gradient is restricted to Y_M.

Proof. Using the formulation (9.1) of φ_N^*, we find

$$\begin{aligned}|\nabla_{Y_M} \varphi^*(\hat{y}) - \nabla_{Y_M} \varphi_N^*(\hat{y})|_{L^2}^2 &= |P\nabla\varphi^*(\hat{y}) - P\nabla\varphi^*(NP^t\hat{y})|_{L^2}^2 \\ &\overset{(9.8)}{\lesssim} |\hat{y} - NP^t\hat{y}|_{L^2}^2 \\ &\lesssim \frac{1}{N^2}|\hat{y}|_{H^1}^2 \overset{(6.1)}{\lesssim} \frac{M^2}{N^2}\, |\hat{y}|_{L^2}^2\,,\end{aligned}$$

where we used a Poincaré inequality on an interval of length $\frac{1}{N}$ and then the inverse Sobolev inequality (6.1) on Y_M from Section 6 (cf. proof of Lemma 3.6 given in that section). $\square$

Corollary 9.7. *There exists K^* such that for all $K \geq K^*$ and all $y \in Y_M$*

$$|\nabla \mathcal{H}_{Y_M}(y) - \nabla \hat{\mathcal{H}}_{Y_M}(y)|_{L^2} \lesssim \frac{1}{K}|y|_{L^2}. \tag{9.15}$$

Proof. Denote $\hat{y}_N = \nabla \mathcal{H}_{Y_M}(y)$ and $\hat{y} = \nabla \hat{\mathcal{H}}_{Y_M}(y)$. We have the duality relations

$$y = \nabla_{Y_M} \varphi_N^*(\hat{y}_N) = \nabla_{Y_M} \varphi^*(\hat{y}).$$

The preceding lemma then leads to the estimate

$$\begin{aligned} |\hat{y}_N - \hat{y}|_{L^2} &\overset{(9.8)}{\simeq} |\nabla_{Y_M} \varphi^*(\hat{y}_N) - \nabla_{Y_M} \varphi^*(\hat{y})|_{L^2} \\ &= |\nabla_{Y_M} \varphi^*(\hat{y}_N) - \nabla_{Y_M} \varphi_N^*(\hat{y}_N)|_{L^2} \overset{(9.14)}{\lesssim} \frac{1}{K} |\hat{y}_N|_{L^2}. \end{aligned} \tag{9.16}$$

Using the fact that $\nabla \varphi_N^*(0) = (\psi^*)'(0) = 0$ (cf. Assumption 2.13), we find

$$|\hat{y}_N|_{L^2} \overset{(9.8)}{\simeq} |\nabla_{Y_M} \varphi_N^*(\hat{y}_N)|_{L^2} = |y|_{L^2}.$$

Combining this with (9.16) yields the desired estimate. □

The last auxiliary result we need is that $\nabla \hat{\mathcal{H}}_{Y_M}$ and $\nabla \mathcal{H}$ are close for large M.

Lemma 9.8. *It holds that for any $z \in L^2(\mathbb{T})$*

$$|\nabla \hat{\mathcal{H}}_{Y_M}(Pz) - \nabla \mathcal{H}(z)|_{L^2} \lesssim \frac{1}{M} |z|_{H^1}. \tag{9.17}$$

Proof of Lemma 9.8. Denote $\hat{y} = \nabla \hat{\mathcal{H}}_{Y_M}(Pz)$ and $\hat{z} = \nabla \mathcal{H}(z)$. We have the duality relations

$$Pz = P \nabla \varphi^*(\hat{y}) \ \text{ and } \ z = \nabla \varphi^*(\hat{z}).$$

In particular, $\nabla \varphi^*(\hat{y}) - \nabla \varphi^*(\hat{z}) \in Y_M^{\perp}$. Since $\hat{y}, P\hat{z} \in Y_M$, we have the estimate

$$\begin{aligned} |\hat{y} - \hat{z}|_{L^2}^2 &\overset{(9.8)}{\lesssim} \langle \nabla \varphi^*(\hat{y}) - \nabla \varphi^*(\hat{z}), \hat{y} - \hat{z} \rangle_{L^2} \\ &= \langle \nabla \varphi^*(\hat{y}) - \nabla \varphi^*(\hat{z}), P\hat{z} - \hat{z} \rangle_{L^2} \overset{(9.8)}{\lesssim} |\hat{y} - \hat{z}|_{L^2} |P\hat{z} - \hat{z}|_{L^2}. \end{aligned} \tag{9.18}$$

The spline estimate from Lemma 3.3 (proven in Section 6) then yields

$$|\hat{y} - \hat{z}|_{L^2} \lesssim |P\hat{z} - \hat{z}|_{L^2} \overset{(3.1)}{\lesssim} \frac{1}{M} |\hat{z}|_{H^1}.$$

Finally, using the uniform bound on φ'' from Lemma 2.12,

$$|\hat{z}|_{H^1} = |\nabla \mathcal{H}(z)|_{H^1} = |\varphi'(z)|_{H^1} = |\varphi''(z) \partial_\theta z|_{L^2} \overset{(2.17)}{\lesssim} |\partial_\theta z|_{L^2} = |z|_{H^1}.$$

Combined with (9.18), this gives the desired estimate. □

9.5 Proof of Theorem 5.3

Using the auxiliary results that were provided in Sections 9.1 - 9.4, Theorem 5.3 is straightforward to prove.

Proof of Theorem 5.3. For any $\zeta \in L^2(\mathbb{T})$ and $y = P\zeta$,

$$\begin{aligned} |\nabla \bar{H}(P\zeta) - \nabla \mathcal{H}(\zeta)|_{L^2} \le\ & |\nabla \bar{H}(y) - \nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*)|_{L^2} \\ &+ |\nabla \bar{H}_{Y_M^{DG}}(y + \bar{z}^*) - \nabla \mathcal{H}_{Y_M^{DG}}(y + z^*)|_{L^2} \\ &+ |\nabla \mathcal{H}_{Y_M^{DG}}(y + z^*) - \nabla \mathcal{H}_{Y_M}(y)|_{L^2} \\ &+ |\nabla \mathcal{H}_{Y_M}(y) - \nabla \hat{\mathcal{H}}_{Y_M}(y)|_{L^2} \\ &+ |\nabla \hat{\mathcal{H}}_{Y_M}(y) - \nabla \mathcal{H}(\zeta)|_{L^2}. \end{aligned}$$

The first term on the right hand side is estimated by (9.9). The second term is estimated by (9.12). The third term vanishes by (9.13). The fourth term is estimated by (9.15). The fifth term is estimated by (9.17). Summing up yields the desired estimate (5.3). □

10 A uniform log-Sobolev inequality for conditional measures

This section is devoted to the proof of Theorem 4.11, which states that the conditional measures $\mu(dx|Px = y)$ satisfies a uniform *logarithmic Sobolev inequality* (LSI). In Section 10.1, we recall four principles for establishing LSI:

1. the tensorization principle;
2. the Holley-Stroock perturbation principle;
3. the Bakry-Émery criterion; and
4. the two-scale criterion of Otto and Reznikoff.

The first three of these are standard results that have been long known and proven useful for deducing LSI in many cases. The fourth principle is a more specialized criterion that has been successfully applied for deducing LSI for spin systems. It will guide our main strategy of proof, while the rest are needed to verify its assumptions. After presenting these principles, we explain in Section 10.2 how they are applied to deduce a uniform LSI for the conditional measures $\mu(dx|Px = y)$ needed in this work.

10.1 Basic principles for the LSI

We begin with formulating the LSI in the generic setting of an Euclidean space X. We write T_xX for the tangent space of X at $x \in X$ and denote by $|\cdot|$, $\langle\cdot,\cdot\rangle$, and ∇ the norm, inner product, and gradient derived from the Euclidean structure of X. We also write $\mathcal{P}(X)$ for the space of Borel probability measures on X.

Definition 10.1 (LSI)**.** *Let $\Phi(z) := z \log z$. We say that a measure $\nu \in \mathcal{P}(X)$ satisfies a logarithmic Sobolev inequality* (LSI) *with constant $\rho > 0$ if for all test functions $h : X \to \mathbb{R}_+$,*

$$\mathrm{Ent}(h\nu|\nu) := \int \Phi(h)\,\nu(dx) - \Phi\left(\int h\,\nu(dx)\right) \le \frac{1}{2\rho}\int \frac{|\nabla h|^2}{h}\,\nu(dx).$$

In this case, we also use the notation $\mathrm{LSI}(\nu) \ge \rho$.

Remark 10.2. The logarithmic Sobolev inequality was first discovered by Gross [18]. It characterizes the speed of convergence to equilibrium of the natural associated drift diffusion process. For more facts about the LSI, we refer to the books [34, 5] and the survey article [25].

The following tensorization principle has been known ever since the notion of LSI came up (see [18]). It is the basic reason for why LSI is well-suited for high-dimensional systems.

Lemma 10.3 (Tensorization principle)**.** *If measures $\nu_1, \nu_2, \cdots, \nu_N \in \mathcal{P}(X)$ satisfy $\mathrm{LSI}(\nu_n) \ge \rho_n$ for $n = 1, 2, \cdots, N$, then*

$$\mathrm{LSI}\left(\bigotimes_{n=1}^{N} \nu_n\right) \ge \min_n \rho_n.$$

We next recall two fundamental criteria for proving logarithmic Sobolev inequalities. The first one is a simple perturbation result due to Holley and Stroock [22].

Lemma 10.4 (Holley-Stroock)**.** *We assume that $\nu \in \mathcal{P}(X)$ satisfies $\mathrm{LSI}(\nu) \ge \rho$. For a bounded function $\delta\psi : X \to \mathbb{R}$, define a measure $\widetilde{\nu} \in \mathcal{P}(X)$ that is absolutely continuous with respect to ν via*

$$\frac{d\widetilde{\nu}}{d\nu}(x) = \frac{1}{Z}\exp[-\delta\psi(x)].$$

Then $\mathrm{LSI}(\widetilde{\nu}) \ge \rho\exp\left[-2\,\mathrm{osc}(\delta\psi)\right]$. Here $\mathrm{osc}(\delta\psi) = \sup_X \delta\psi - \inf_X \delta\psi$ stands for the total oscillation of the perturbation $\delta\psi$.

The second criterion is due to Bakry and Émery [4]. It says that strong convexity of the Hamiltonian yields LSI.

Lemma 10.5 (Bakry-Émery)**.** *Let $\nu \in \mathcal{P}(X)$ be absolutely continuous with respect to the Hausdorff measure $\mathcal{H}$ on X and given by a Hamiltonian H, i.e.*

$$\frac{d\nu}{d\mathcal{H}}(x) = \frac{1}{Z} e^{-H(x)}.$$

If H is C^2 and satisfies

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

then $\operatorname{LSI}(\nu) \,\geq\, \lambda$.

Proofs of the facts mentioned so far can be found e.g. in [20] or in [25]. As pointed out above, we will also need the two-scale criterion that was presented in [31] and which is also contained, in a slightly different formulation, in [19]. We first define a disintegration of measure analogous to Definition 3.7 in the setting of a product space.

Definition 10.6. *Let $\nu \in \mathcal{P}(X_1 \times X_2)$ be a measure with smooth positive density w.r.t. to the Hausdorff measure. We decompose ν into a family of conditional measures $\{\nu(dx_1|x_2)\}_{x_2 \in X_2} \subset \mathcal{P}(X_1)$ and a marginal measure $\bar{\nu} \in \mathcal{P}(X_2)$. This decomposition is such that for all measurable $h : X_1 \times X_2 \to \mathbb{R}$:*

$$\int_{X_1 \times X_2} h\, d\nu \,=\, \int_{X_2} \int_{X_1} h(x_1, x_2)\, \nu(dx_1|x_2)\, \bar{\nu}(dx_2).$$

The two-scale criterion reads as follows.

Lemma 10.7 (Two-scale criterion for LSI)**.** *Let $\nu \in \mathcal{P}(X_1 \times X_2)$ be a measure whose Hamiltonian H is C^2. Assume that there exist constants $\rho_1, \rho_2 > 0$ such that*

$$\operatorname{LSI}(\nu(dx_1|x_2)) \geq \rho_1 \quad \text{for all } x_2 \in X_2, \tag{10.1}$$

$$\operatorname{LSI}(\bar{\nu}) \geq \rho_2. \tag{10.2}$$

Moreover, assume that

$$\frac{1}{\rho_1} \frac{1}{\rho_2} \sup_{X_1 \times X_2} |\nabla_{X_1} \nabla_{X_2} H(x)|^2 \,=\, \kappa < \infty. \tag{10.3}$$

Here

$$|\nabla_{X_1} \nabla_{X_2} H(x)| \,=\, \sup \left\{ \langle \operatorname{Hess} H(x)\, u, v\rangle \,\middle|\, u \in T_x X_1, v \in T_x X_2, |u| = |v| = 1 \right\}$$

is finite if $\operatorname{Hess} H$ is bounded. Then

$$\operatorname{LSI}(\nu) \,\geq\, \frac{1}{2} \left(\rho_1 + (1+\kappa)\rho_2 \,-\, \sqrt{(\rho_1 + (1+\kappa)\rho_2)^2 - 4\rho_1\rho_2} \right).$$

Lemma 10.7 says that LSI for the conditional and marginal measures may be combined to yield a LSI for the full measure, under the coupling assumption (10.3). A proof of the two-scale criterion can be found in [31] where it is stated as Theorem 2.

10.2 Uniform LSI for conditional measures

In this section we explain how the basic principles of Section 10.1 are used to deduce Theorem 4.11. The proof adapts the strategy in [19], which covered the case for $L = 0$, when Y_M is the space of piecewise constant functions. However, for $L \geq 1$, due to the non-locality of the spline functions in Y_M, we need to modify the strategy in [19] by introducing an additional step. Namely, we first deduce a uniform LSI on the level of Y_M^{DG}, the space of piecewise polynomials of degree $\leq L$ introduced in Definition 7.1, and then apply two-scale criterion to get back a uniform LSI on the level of Y_M.

Theorem 10.8. *The measures* $\mu^N(dx|Q_M x = y)$ *satisfy* $LSI(\varrho_{DG})$ *with a constant* $\varrho_{DG} > 0$ *uniform in* N, M *and* $y \in Y_M^{DG}$.

Proof of Theorem 10.8. Since the measures $\mu^N(dx|Q_M x = y)$ factor (cf. equation (7.8)), by the tensorization principle (cf. Lemma 10.3) it suffices to show that the measures

$$\mu_\alpha^K(dx) := \mu^K(dx|Q_1 x = \alpha)$$

satisfy a uniform LSI for K large enough. The strategy is to apply the same block decomposition, from which we obtained M blocks of size K from $N = MK$ sites, to the K sites in each block. Namely, we assume $K = RJ$ and divide the K sites into R many *sub-blocks* of size J, where $J \in \mathbb{N}$ is large but *fixed*. By the adjoint relation between $Q_R : \mathbb{R}^K \to Y_R^{DG}$ and $KQ_R^t : Y_R^{DG} \to \mathbb{R}^K$, we have the orthogonal decomposition

$$\mathbb{R}^K = \ker Q_R \oplus \operatorname{im} KQ_R^t. \tag{10.4}$$

The key observation here is that $\ker Q_R \subset \ker Q_1$ because a polynomial in Y_1^{DG} is also a piecewise polynomial in Y_R^{DG}. Thus, the decomposition (10.4) induces a decomposition on the state space of $\mu^K(dx|Q_1 x = \alpha)$:

$$\{x \in \mathbb{R}^K : Q_1 x = \alpha\} = \underbrace{\{x_\parallel \in \ker Q_R\}}_{=:V} \oplus \underbrace{\{x_\perp \in \operatorname{im} KQ_R^t : Q_1 x = \alpha\}}_{=:W_\alpha}.$$

The strategy now is to apply the two-scale criterion (cf. Lemma 10.7) with $X_1 = V, X_2 = W_\alpha$.

- Uniform LSI (10.1) for conditional measures $\mu_\alpha^K(dx_\parallel|x_\perp)$: let $\beta = Q_R x_\perp$, then

$$\mu_\alpha^K(dx_\parallel|x_\perp) = \mu^K(dx|Q_R x = \beta) \overset{(7.8)}{=} \bigotimes_{r=1}^{R} \mu^J(dx^{(r)}|Q_1 x^{(r)} = \beta^{(r)}).$$

  By the tensorization principle, it suffices to show uniform LSI for the conditional measure $\mu^J(dx^{(r)}|Q_1 x^{(r)} = \beta^{(r)})$, whose Hamiltonian is just a restriction of the Hamiltonian H_J of the full measure μ^J. From the explicit form of H_J in (2.1), a combination of Holley-Stroock (cf. Lemma 10.4) and Bakry-Émery (cf. Lemma 10.5) yields

$$\operatorname{LSI}\left(\mu^J(dx^{(r)}|Q_1 x^{(r)} = \beta^{(r)})\right) \geq \exp\left(-2J \operatorname{osc}(\delta\psi)\right) =: \rho_1 > 0.$$

- LSI (10.2) for the marginal measure $\bar{\mu}_\alpha^K(dx_\perp)$: observe that the Hamiltonian of the marginal measure $\bar{\mu}_\alpha^K(dx_\perp)$, which we denote $\hat{H}_{W_\alpha}$, is closely related to the coarse-grained Hamiltonian $\bar{H}_{Y_R^{DG}}$ defined in equation (7.5):

$$\hat{H}_{W_\alpha}(x_\perp) := -\log \int_V \exp(-H_K(x_\parallel + x_\perp))\, dx_\parallel = K\bar{H}_{Y_R^{DG}}(Q_R x_\perp). \tag{10.5}$$

  By Theorem 8.3, $\bar{H}_{Y_R^{DG}}$ is uniformly strongly convex for large enough J. This property is transferred to $\hat{H}_{W_\alpha}$ via the relation (10.5): for $u \in T_{x_\perp} W_\alpha \simeq \operatorname{im} KQ_R^t \subset \mathbb{R}^K$,

$$\begin{aligned} u \cdot \operatorname{Hess} \hat{H}_{W_\alpha}(x_\perp)\, u &= K\langle Q_R u, \operatorname{Hess} \bar{H}_{Y_R^{DG}}(Q_R x_\perp)\, Q_R u\rangle_{L^2} \\ &\geq 2K\lambda |Q_R u|_{L^2}^2 \\ &\overset{(7.4)}{=} 2K\lambda \left(1 + O\left(J^{-2}\right)\right) |u|_{L^2}^2 = 2\lambda \left(1 + O\left(J^{-2}\right)\right) |u|^2, \end{aligned}$$

  where we accounted for different Euclidean structures on $\mathbb{R}^K$. This yields the uniform strong convexity of $\hat{H}_{W_\alpha}$, and the Bakry-Émery criterion (cf. Lemma 10.5) implies (10.2) with constant $\rho_2 = \lambda$.

- Coupling condition (10.3): this follows from the uniform C^2 bound of H_K.

Overall, we may hence apply Lemma 10.7 which yields that for large enough J:

$$\mathrm{LSI}(\mu_\alpha^K) \geq \rho_{DG} := \frac{1}{2}\left(\rho_1 + (1+\kappa)\lambda - \sqrt{(\rho_1 + (1+\kappa)\lambda)^2 - 4\rho_1\lambda}\right),$$

which is bounded from below uniformly in K. □

With Theorem 10.8 at hand, the proof of Theorem 4.11 consists of another application of the two-scale criterion (see Lemma 10.7), which is very similar to the one in proof of Theorem 10.8: there we introduced an intermediate level by prescribing a global polynomial as the constraint, here we will introduce an intermediate level by prescribing a spline on the same mesh as the constraint.

Proof of Theorem 4.11. Since $\ker Q_M \subset \ker P$ (as a spline in Y_M is also a piecewise polynomial in Y_M^{DG}), the orthogonal decomposition

$$\mathbb{R}^N = \ker Q_M \oplus \operatorname{im} NQ_M^t$$

induces a decomposition on the state space of $\mu(dx|y) = \mu^N(dx|Px = y)$:

$$\{x \in \mathbb{R}^N : Px = y\} = \underbrace{\{x_\parallel \in \ker Q_M\}}_{=:V} \oplus \underbrace{\{x_\perp \in \operatorname{im} NQ_M^t : Px_\perp = y\}}_{=:W_y}.$$

We now apply the two-scale criterion (cf. Lemma 10.7) with $X_1 = V, X_2 = W_y$.

- Uniform LSI (10.1) for conditional measures $\mu^N(dx_\parallel|x_\perp)$: this is directly provided by Theorem 10.8 since $\mu^N(dx_\parallel|x_\perp) = \mu^N(dx|Q_M x = Q_M x_\perp)$.
- LSI (10.2) for the marginal measure $\bar{\mu}^N(dx_\perp|Px_\perp = y)$: the same argument used in the proof of Theorem 10.8 applies to show that the uniform strong convexity of $\bar{H}_{Y_M^{DG}}$ transfers to the Hamiltonian $\hat{H}_{W_y}$ of the measure $\bar{\mu}^N(dx_\perp|Px_\perp = y)$, and Bakry-Émery criterion then yields an LSI constant of λ.
- Coupling condition (10.3): this follows from the uniform C^2 bound of H_N.

Overall, Lemma 10.7 then yields that for large enough K:

$$\mathrm{LSI}(\mu(dx|y)) \geq \rho := \frac{1}{2}\left(\rho_{DG} + (1+\kappa)\lambda - \sqrt{(\rho_{DG} + (1+\kappa)\lambda)^2 - 4\rho_{DG}\lambda}\right)$$

which is uniformly bounded from below in N. □

A Preservation of convexity under coarse-graining

In this section, we verify Lemma 8.1 which gives a quantitative statement on the preservation of convexity under coarse-graining. For $c \geq 0$, consider the function $F_c : W \oplus Z \to \mathbb{R}$

$$F_c(w,z) := F(w,z) - \frac{c}{2}|z|^2,$$

and the coarse-grained function $\overline{F_c} : Z \to \mathbb{R}$

$$\overline{F_c}(z) := -\log \int_W \exp(-F_c(w,z))\, dw = \bar{F}(z) - \frac{c}{2}|z|^2.$$

It follows immediately that

$$\operatorname{Hess}_{W\oplus Z} F \geq c\,\mathrm{id}_{W\oplus Z} \;\Rightarrow F_c \text{ is convex on } W \oplus Z$$
$$\operatorname{Hess}_{W\oplus Z} F \leq c\,\mathrm{id}_{W\oplus Z} \;\Rightarrow F_c \text{ is concave on } Z.$$

and

$$\overline{F_c} \text{ is convex} \Rightarrow \ \operatorname{Hess}_Z \bar{F} \geq c\,\mathrm{id}_Z,$$
$$\overline{F_c} \text{ is concave} \Rightarrow \ \operatorname{Hess}_Z \bar{F} \leq c\,\mathrm{id}_Z\,.$$

Thus, we need to show that

$$F_c \text{ is convex on } W \oplus Z \Rightarrow \ \overline{F_c} \text{ is convex}, \tag{A.1}$$
$$F_c \text{ is concave on } Z \Rightarrow \ \overline{F_c} \text{ is concave}. \tag{A.2}$$

Argument for (A.1): this is a simple consequence of the Prékopa-Leindler inequality (see e.g. [6, Theorem 4.24]). Given $\lambda \in (0,1)$, $z_1, z_2 \in Z$, consider functions $h, f, g : W \to \mathbb{R}$

$$h(w) := \exp(-F_c(w, (1-\lambda)z_1 + \lambda z_2)), \tag{A.3}$$
$$f(w) := \exp(-F_c(w, z_1)),$$
$$g(w) := \exp(-F_c(w, z_2)).$$

Convexity of F_c on $W \oplus Z$ means that for any $w_1, w_2 \in W$

$$h((1-\lambda)w_1 + \lambda w_2) \geq f(w_1)^{1-\lambda} g(w_2)^{\lambda}.$$

By the Prékopa-Leindler inequality, this implies

$$\int_W h(w)\,dw \geq \left(\int_W f(w)\,dw\right)^{1-\lambda} \left(\int_W g(w)\,dw\right)^{\lambda},$$

which yields

$$\overline{F_c}((1-\lambda)z_1 + \lambda z_2)) \leq (1-\lambda)\overline{F_c}(z_1) + \lambda \overline{F_c}(z_2).$$

Argument for (A.2): this is a simple consequence of the Hölder inequality. Given $\lambda \in (0,1)$, $z_1, z_2 \in Z$, let h, f, g be defined as in (A.3). Concavity of F_c on Z means that for any $w \in W$,

$$h(w) \leq f(w)^{1-\lambda} g(w)^{\lambda}.$$

By the Hölder inequality, this implies

$$\int_W h(w)\,dw \leq \left(\int_W f(w)\,dw\right)^{1-\lambda} \left(\int_W g(w)\,dw\right)^{\lambda},$$

which yields

$$\overline{F_c}((1-\lambda)z_1 + \lambda z_2)) \geq (1-\lambda)\overline{F_c}(z_1) + \lambda \overline{F_c}(z_2).$$

B A multivariate local central limit theorem

We now begin with the long and technical proof of Proposition 8.10. We recommend the interested reader to compare with the proof of Proposition 31 in [19]. As in the usual proof of the (local) central limit theorem, we start by using independence and the Fourier transform to obtain an explicit formula for $g_{J,\beta}(0)$. To clear up notation, we denote

$$[\xi] := JQ_1^t \xi, \quad [\xi]_j := (JQ_1^t \xi)_j$$

for the J projected spins of $\xi \in Y_1^{DG}$.

Lemma B.1. *For each $m \in \mathbb{R}$, denote by*

$$h(m,z) := \int_{\mathbb{R}} e^{iz(x-m)}\, \mu_m(dx)$$

for the characteristic function of the centered *version of the single-site measure μ_m defined in* (8.2). *We have*

$$g_{J,\beta}(0) = \left(\frac{1}{2\pi}\right)^{L+1} \int_{\mathbb{R}^{L+1}} \prod_{j=1}^{J} h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)\, d\xi, \tag{B.1}$$

Proof of Lemma B.1. By Fourier inversion,

$$g_{J,\beta}(0) = \left(\frac{1}{2\pi}\right)^{L+1} \int_{\mathbb{R}^{L+1}} \left(\int_{\mathbb{R}^{L+1}} \exp\left(i\,\xi\cdot u\right) g_{J,\beta}(du) \right) d\xi.$$

Since $g_{J,\beta}$ is the law of the random variable

$$J^{\frac{1}{2}}(Q_1 x - \beta) = J^{\frac{1}{2}} Q_1 (x - m(\beta))$$

under $\nu_{J,\beta}(dx) = \prod_{j=1}^{J} \mu_{m_j(\beta)}(dx_j)$, the inner integral equals to

$$\begin{aligned}\int_{\mathbb{R}^J} \exp(i\,\xi\cdot J^{\frac{1}{2}} Q_1(x-m(\beta)))\,\nu_{J,\beta}(dx) &= \int_{\mathbb{R}^J} \exp(i\,J^{-\frac{1}{2}}(JQ_1^t\xi)\cdot(x-m(\beta)))\,\nu_{J,\beta}(dx)\\ &= \prod_{j=1}^{J} \int_{\mathbb{R}} \exp(i\,J^{-\frac{1}{2}}[\xi]_j\,(x_j - m_j(\beta)))\,\mu_{m_j(\beta)}(dx_j),\end{aligned}$$

which yields the right hand side of (B.1). □

The strategy for the rest of the proof is to split the integral on the right hand side of (B.1) into an inner and an outer part. We will show that for small enough δ and for large enough J (depending on δ)

$$\lim_{J\to\infty} \int_{\left\{|\xi|>J^{\frac{1}{2}}\delta\right\}} \prod_{j=1}^{J} |h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)|\, d\xi = 0, \tag{B.2}$$

$$\int_{\left\{|\xi|\le J^{\frac{1}{2}}\delta\right\}} \prod_{j=1}^{J} |h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)|\, d\xi \;\le\; C, \tag{B.3}$$

$$\left| \int_{\left\{|\xi|\le J^{\frac{1}{2}}\delta\right\}} \prod_{j=1}^{J} h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)\, d\xi \right| \;\ge\; \frac{1}{C}, \tag{B.4}$$

$$\lim_{J\to\infty} \left\| \operatorname{Hess}_\beta \int_{\left\{|\xi|>J^{\frac{1}{2}}\delta\right\}} \prod_{j=1}^{J} h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)\, d\xi \right\| = 0, \tag{B.5}$$

$$\left\| \operatorname{Hess}_\beta \int_{\left\{|\xi|\le J^{\frac{1}{2}}\delta\right\}} \prod_{j=1}^{J} h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j)\, d\xi \right\| \;\le\; C. \tag{B.6}$$

The bounds for $g_{J,\beta}(0)$ in (8.17) follow from (B.2) - (B.4). The bounds for the Hessian in (8.19) follow from (B.5) and (B.6). The bounds for the gradient in (8.18) are then immediate from interpolation. To establish these estimates, we will need some auxiliary results stated and verified in Section B.1 below. We then establish bounds (B.2) - (B.4)) in Section B.2 and bounds (B.5) - (B.6) in Section B.3, which completes the proof of Proposition 8.10.

B.1 Auxiliary results

The first ingredient we need is a collection of elementary properties of the function h.

Lemma B.2. *We have the following bounds and decay properties for h and its derivatives:*

- *For all $m, z \in \mathbb{R}$*
$$|h(m,z)| \leq 1. \tag{B.7}$$
- *Given $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that for $m \in \mathbb{R}, |z| \geq \varepsilon$*
$$|h(m,z)| \leq \frac{1}{1 + c_\varepsilon |z|}. \tag{B.8}$$
- *For all $m, z \in \mathbb{R}$*
$$\left|\frac{\partial h}{\partial m}(m,z)\right| \lesssim 1 + |z|, \quad \left|\frac{\partial^2 h}{\partial m^2}(m,z)\right| \lesssim 1 + |z|^2. \tag{B.9}$$

Moreover, there exists $\delta_0 > 0$ such that for $z \in [-\delta_0, \delta_0]$, $m \in \mathbb{R}$, we can express h as
$$h(m,z) = \exp(-z^2\, h_2(m,z)), \tag{B.10}$$
for some function $h_2 : \mathbb{R} \times [-\delta_0, \delta_0] \to \mathbb{C}$. The function h_2 satisfies that

- *For all $m \in \mathbb{R}$*
$$h_2(m,0) = \frac{1}{2}\operatorname{Var}(\mu_m). \tag{B.11}$$
- *There exists constants $0 < c_{h_2}, C_{h_2} < \infty$ such that for all $m \in \mathbb{R}, z \in [-\delta_0, \delta_0]$*
$$\operatorname{Re} h_2(m,z) \geq c_{h_2}, \quad \left|\frac{\partial h_2}{\partial z}(m,z)\right| \leq C_{h_2}. \tag{B.12}$$
- *For all $m \in \mathbb{R}, z \in [-\delta_0, \delta_0]$*
$$\left|\frac{\partial h_2}{\partial m}(m,z)\right| \lesssim 1, \quad \left|\frac{\partial^2 h_2}{\partial m^2}(m,z)\right| \lesssim 1. \tag{B.13}$$

The estimates of Lemma (B.2) should not be surprising as $h(m,\cdot)$ is just the Fourier transform of μ_m which belongs to the exponential family of a perturbed standard Gaussian measure. For the proofs, we refer the reader to [19, Lemma 39, 40].

The second ingredient for the proof of Proposition 8.10 is a uniform control on the magnitude of the projected spins $[\xi]_j := (JQ_1^t\xi)_j$ which enters into the second argument of h (after rescaling). This is new compared to [19].

Lemma B.3. *There exists constant $1 < C_{max} < \infty$ such that for $J \geq 1$,*
$$\max_{1 \leq j \leq J} |[\xi]_j| \leq C_{max}\, |\xi|_{L^2}, \quad \forall\, \xi \in Y_1^{DG}. \tag{B.14}$$
Given $0 < \alpha < 1$, define
$$I_\alpha(\xi) := \{1 \leq j \leq J : |[\xi]_j| > \alpha|\xi|_{L^2}\}$$
for $\xi \in Y_1^{DG}$. Then there exists constants J_α and $0 < \kappa(\alpha) < 1$ such that for $J \geq J_\alpha$,
$$|I_\alpha(\xi)| > \kappa(\alpha)\, J, \quad \forall\, \xi \in Y_1^{DG}. \tag{B.15}$$

Proof of Lemma B.3. Since the projected spin $[\xi]_j = (JQ_1^t)_j$ is just the average of $\xi(t)$ over the subinterval $\left[\frac{j-1}{J}, \frac{j}{J}\right]$, we have

$$\max_{1\le j\le J} |[\xi]_j| \le \max_{t\in[0,1]} |\xi(t)| \le C_{max}|\xi|_{L^2} \tag{B.16}$$

for some constant $1 < C_{max} < \infty$ because norms on the finite-dimensional space Y_1^{DG} are equivalent. This establishes (B.14). Moreover, the uniform upper bound (B.16) on $\xi(t)$ also forces it to be bounded below on a significant portion of the interval $[0,1]$, e.g.

$$\text{Leb}\left(\{t \in [0,1] : |\xi(t)| > \alpha|\xi|_{L^2}\}\right) > \frac{1-\alpha^2}{C_{max}^2}, \tag{B.17}$$

where Leb denotes the Lebesgue measure on $\mathbb{R}$. Estimate (B.17) may be seen as a continuous version of (B.15) and implies that for all $J \ge 1$,

$$\left|\left\{1 \le j \le J : \max_{t\in\left[\frac{j-1}{J},\frac{j}{J}\right]} |\xi(t)| > \alpha|\xi|_{L^2}\right\}\right| > \frac{1-\alpha^2}{C_{max}^2} J.$$

It only remains to verify that $[\xi]_j$ approximates $\xi(t)$ uniformly well for J large enough. By scaling, it suffices to show that as $J \to \infty$,

$$\max_{1\le j\le J} \max_{t\in\left[\frac{j-1}{J},\frac{j}{J}\right]} |[\xi]_j - \xi(t)| \to 0$$

uniformly over the unit sphere

$$\mathbb{S} := \{\xi \in Y_1^{DG} : |\xi|_{L^2} = 1\}.$$

But this is an immediate consequence of the *equicontinuity* of $\mathbb{S}$ as a family of continuous functions on $[0,1]$, which holds true by Arzelá-Ascoli because $\mathbb{S}$ is compact (in the uniform topology), being a closed and bounded set in a finite-dimensional space. $\square$

The last ingredient we need is the following uniform C^2 estimates for the mean spins $m_j(\beta)$.

Lemma B.4. *For all $J \ge 1$ and $\beta \in Y_1^{DG}$, it holds uniformly in $1 \le j \le J$ that*

$$\|\nabla m_j(\beta)\| \lesssim 1, \tag{B.18}$$

$$\|\operatorname{Hess} m_j(\beta)\| \lesssim 1. \tag{B.19}$$

In order to prove Lemma B.4, we need a 3rd derivative bound on $\bar{\psi}_J$:

Lemma B.5. *For all $J \ge 1$ and $\hat{\beta}, \beta \in Y_1^{DG}$:*

$$\|\, D^3\bar{\psi}_J^*(\hat{\beta})\,\| \lesssim 1, \tag{B.20}$$

$$\|\, D^3\bar{\psi}_J(\beta)\,\| \lesssim 1. \tag{B.21}$$

Proof of Lemma B.5. Using (8.4) from Lemma 8.5 with similar calculations in the proof of Proposition 8.6, we have that for any $\hat{x}, \hat{\xi} \in \mathbb{R}^J$

$$|\, D^3\psi_J^*(\hat{x})(\hat{\xi},\hat{\xi},\hat{\xi})\,| \le \sum_{j=1}^J |(\psi^*)'''(\hat{x}_j)|\,|\hat{\xi}_j|^3 \lesssim \max_{1\le j\le J}(\hat{\xi}_j)\,|\hat{\xi}|^2,$$

and consequently, for any $\hat{\eta} \in Y_1^{DG}$

$$\begin{aligned} |\, D^3\bar{\psi}_J^*(\hat{\beta})(\hat{\eta},\hat{\eta},\hat{\eta})\,| &= \frac{1}{J}|\, D^3\psi_J^*([\hat{\beta}])([\hat{\eta}],[\hat{\eta}],[\hat{\eta}])\,| \\ &\lesssim \max_{1\le j\le J} |[\hat{\eta}]_j|\,|[\hat{\eta}]|_{L^2}^2 \le C_{max}|\hat{\eta}|\left(1 + O\left(J^{-2}\right)\right)|\hat{\eta}|^2, \end{aligned}$$

where we used both the uniform upper bound (B.14) and the estimate (7.3) in the last step. This establishes (B.20). Together with the 2nd derivative estimates of Proposition 8.6, this in turn yields (B.21) by a standard calculation. □

Proof of Lemma B.4. We recall that

$$m_j(\beta) = \int_{\mathbb{R}} z \, \exp(-\psi^*(\hat{m}_j(\beta)) + \hat{m}_j(\beta)z - \psi(z)) \, dz.$$

Standard calculation yields

$$\begin{aligned} \nabla m_j &\overset{(8.11)}{=} \mathrm{Var}(\mu_{m_j}) \, \nabla \hat{m}_j, \\ \mathrm{Hess}\, m_j &= \mathrm{Var}(\mu_{m_j}) \; \mathrm{Hess}\, \hat{m}_j + \left(\int (z - m_j)^3 \; \mu_{m_j}(dz) \right) \, \nabla \hat{m}_j \, \otimes \, \nabla \hat{m}_j. \end{aligned}$$

By the uniform estimates on $\mathrm{Var}(\mu_m)$ and $\int (z - m)^3 \, \mu_m(dz)$ in (8.3) and (8.4), it remains to bound $\nabla \hat{m}_j$ and $\mathrm{Hess}\, \hat{m}_j$. By equation (8.10) and (8.9),

$$\hat{m}_j = (J Q_1^t \nabla \bar{\psi}_J)_j = [\nabla \bar{\psi}_J]_j.$$

For $\eta \in Y_1^{DG}$, let ∂_η denote the η-directional derivative. Then by the uniform bound (B.14),

$$\begin{aligned} \langle \nabla \hat{m}_j, \eta \rangle &= \partial_\eta \, \hat{m}_j = [\partial_\eta \nabla \bar{\psi}_J]_j \leq C_{max} |\partial_\eta \nabla \bar{\psi}_J| \leq C_{max} \| \, \mathrm{Hess}\, \bar{\psi}_J \| |\eta|, \\ \langle \mathrm{Hess}\, \hat{m}_j \cdot \eta, \eta \rangle &= \partial_\eta^2 \, \hat{m}_j = [\partial_\eta^2 \nabla \bar{\psi}_J]_j \leq C_{max} |\partial_\eta^2 \nabla \bar{\psi}_J| \leq C_{max} \| D^3 \bar{\psi}_J \| |\eta|^2. \end{aligned}$$

Since the Hessian and 3rd derivative of $\bar{\psi}_J$ are uniformly bounded (cf. Lemma 8.6 and Lemma B.5), this concludes the proof of (B.18) and (B.19). □

B.2 Proof of bounds (B.2) - (B.4)

From now on, we fix some $\alpha \in (0, 1)$ and assume from now on $J \geq J_\alpha$, so that $|I_\alpha(\xi)| \geq \kappa(\alpha) J$ (cf. Lemma B.3). We establish the bounds (B.2) - (B.4) by estimating the integrand

$$\prod_{j=1}^{J} h(m_j, J^{-\frac{1}{2}} [\xi]_j)$$

in the two regions $|\xi| > J^{\frac{1}{2}} \delta$ and $|\xi| \leq J^{\frac{1}{2}} \delta$.

B.2.1 The region $|\xi| > J^{\frac{1}{2}} \delta$

In this region, observe that for $j \in I_\alpha(\xi)$,

$$|J^{-\frac{1}{2}} [\xi]_j| \geq \alpha J^{-\frac{1}{2}} |\xi| \geq \alpha \delta.$$

Hence choosing $\varepsilon = \alpha\delta$ in the decay property (B.8) gives

$$|h(m_j, J^{-\frac{1}{2}} [\xi]_j)| \leq \frac{1}{1 + c_\varepsilon \alpha J^{-\frac{1}{2}} |\xi|} \leq \frac{1}{1 + c_\varepsilon \varepsilon}.$$

Consequently, using the global bound (B.7) on h to deal with $j \notin I_\alpha(\xi)$, we have

$$\begin{aligned}\prod_{j=1}^{J} |h(m_j, J^{-\frac{1}{2}}[\xi]_j)| &\le \left(\frac{1}{1 + c_\varepsilon \alpha J^{-\frac{1}{2}}|\xi|}\right)^{|I_\alpha(\xi)|} \\ &\le \left(\frac{1}{1 + c_\varepsilon \varepsilon}\right)^{|I_\alpha(\xi)|-(L+2)} \left(\frac{1}{1 + c_\varepsilon \alpha J^{-\frac{1}{2}}|\xi|}\right)^{L+2} \\ &= \exp\left[-\Theta(|I_\alpha(\xi)|)\right] O\left(J^{-\frac{1}{2}}|\xi|\right)^{-(L+2)} \\ &= \exp\left[-\Theta(J)\right] O\left(|\xi|^{-(L+2)}\right). \end{aligned} \tag{B.22}$$

Integrating over the region $\{|\xi| > J^{\frac{1}{2}}\delta\}$, this yields (B.2) because the function $|\xi|^{-(L+2)}$ is integrable away from zero.

B.2.2 The region $|\xi| \le J^{\frac{1}{2}}\delta$ (upper bound)

In this region, if we choose $\delta \le \frac{\delta_0}{C_{max}}$, then for all J and $1 \le j \le J$,

$$|J^{-\frac{1}{2}}[\xi]_j| \le \delta_0.$$

Thus, we may use the representation of h via h_2 in Lemma B.2 to compute that

$$\begin{aligned}\prod_{j=1}^{J} h(m_j, J^{-\frac{1}{2}}[\xi]_j) &= \exp\left(-\sum_{j=1}^{J} (J^{-\frac{1}{2}}\,[\xi]_j)^2\; h_2(m_j, J^{-\frac{1}{2}}\,[\xi]_j)\right) \\ &= \exp\left(-\frac{1}{J}\sum_{j=1}^{J} [\xi]_j^2\; h_2(m_j, J^{-\frac{1}{2}}\,[\xi]_j)\right) = \exp\left(-S(\xi, J^{-\frac{1}{2}}\,[\xi])\right), \end{aligned} \tag{B.23}$$

where we denote

$$S(\xi, z) := \frac{1}{J}\sum_{j=1}^{J} [\xi]_j^2\; h_2(m_j, z_j)$$

for $\xi \in Y_1^{DG}, z \in \mathbb{R}^J$. Observe that

$$\frac{1}{J}\sum_{j=1}^{J} [\xi]_j^2 = |[\xi]|_{L^2}^2 \overset{(7.3)}{=} \left(1 + O(J^{-2})\right)|\xi|^2. \tag{B.24}$$

Thus, applying the lower bound $\operatorname{Re} h_2 \ge c_{h_2}$ in (B.12) to equation (B.23) gives

$$\exp\left(-S(\xi, J^{-\frac{1}{2}}\,[\xi])\right) \le \exp\left(-c_{h_2}\frac{1}{J}\sum_{j=1}^{J} [\xi]_j^2\right) = \exp\left(-c_{h_2}\left(1 + O(J^{-2})\right)|\xi|^2\right). \tag{B.25}$$

Integrating over the region $\{|\xi| \le J^{\frac{1}{2}}\delta\}$, this yields the upper bound (B.3).

B.2.3 The region $|\xi| \le J^{\frac{1}{2}}\delta$ (lower bound)

Let us now deduce the lower bound (B.4) from equation (B.23). Set $z = J^{-\frac{1}{2}}\,[\xi]$. Our strategy is to split the integrand as

$$\begin{aligned}\exp\left(-S(\xi, z)\right) &= \exp\left(-S(\xi, 0)\right) + \left[\exp\left(-S(\xi, z)\right) - \exp\left(-S(\xi, 0)\right)\right] \\ &= \exp\left(-S(\xi, 0)\right) + \exp\left(-S(\xi, 0)\right)\left(\exp\left[S(\xi, z) - S(\xi, 0)\right] - 1\right)\end{aligned} \tag{B.26}$$

and deal with the two terms separately:

- Using the fact (B.11) that $h_2(m_j, 0) = \frac{1}{2}\operatorname{Var}(\mu_{m_j})$ and the uniform bound (8.3) on $\operatorname{Var}(\mu_{m_j})$, we have
$$\frac{c}{2} \le h_2(m_j, 0) \le \frac{C}{2}.$$
where c, C are constants in (8.3). Together with observation (B.24), it follows that
$$\frac{c}{2}\left(1 + O(J^{-2})\right)|\xi|^2 \le S(\xi, 0) \le \frac{C}{2}\left(1 + O(J^{-2})\right)|\xi|^2.$$
Thus, the first term in (B.26) is positive real-valued and
$$\exp\left(-C_1|\xi|^2\right) \le \exp\left(-S(\xi, 0)\right) \le \exp\left(-C_2|\xi|^2\right).$$
for some universal constants $C_1, C_2 > 0$. In particular, its integral over the region $\{|\xi| \le J^{\frac{1}{2}}\delta\}$ is uniformly bounded from below.
- Thus, it remains to show the second term in (B.26) has a smaller contribution in the integral. Using first order Taylor approximation, we have
$$\begin{aligned}|h_2(m_j, z_j) - h_2(m_j, 0)| &\le \sup_{z\in[-\delta_0,\delta_0]}\left|\frac{\partial h_2}{\partial z}\right||z_j| \\ &\overset{(B.12)}{\le} C_{h_2}|J^{-\frac{1}{2}}[\xi]_j| \le C_{h_2}C_{max}J^{-\frac{1}{2}}|\xi| \le C_{h_2}C_{max}\,\delta.\end{aligned}$$
Together with observation (B.24), it follows that
$$|S(\xi, z) - S(\xi, 0)| \le C_{h_2}C_{max}\,\delta\left(1 + O(J^{-2})\right)|\xi|^2.$$
Thus, by the elementary inequality $|e^w - 1| \le e^{|w|} - 1$, we have
$$|\left(\exp\left[S(\xi, z) - S(\xi, 0)\right] - 1\right)| \le \exp\left(C_3\,\delta|\xi|^2\right) - 1$$
for some universal constant $C_3 > 0$.

In conclusion, we have shown that the integral in (B.4) is bounded below by
$$\int_{\{|\xi|\le J^{\frac{1}{2}}\delta\}} \exp\left(-C_1|\xi|^2\right)d\xi - \int_{\mathbb{R}^{L+1}} \exp\left(-C_2|\xi|^2\right)\left(\exp\left(C_3\,\delta|\xi|^2\right) - 1\right)d\xi.$$
As $\delta \to 0$, the second integral vanishes (independently of J). Thus, by choosing δ small enough (and then J large enough), the first integral dominates the second integral and yields the uniform lower bound (B.4) for all large enough J.

B.3 Proof of bounds (B.5) - (B.6)

Let us now turn to the terms that involve derivatives with respect to β. Interchanging differentiation and integration, we need to estimate
$$\operatorname{Hess}_\beta \prod_{j=1}^{J} h(m_j(\beta), J^{-\frac{1}{2}}[\xi]_j) \tag{B.27}$$
in the two regions of integration $|\xi| > J^{\frac{1}{2}}\delta$ and $|\xi| \le J^{\frac{1}{2}}\delta$. To clear up notation, let us denote
$$\begin{aligned}z_j(\xi) &:= J^{-\frac{1}{2}}[\xi]_j, \\ \mathbf{u}_j(\beta, \xi) &:= (m_j(\beta), z_j(\xi)).\end{aligned}$$

We compute the integrand in (B.27) as

$$\nabla_\beta \prod_{j=1}^{J} h(\mathbf{u}_j) = \sum_{j=1}^{J} \frac{\partial h}{\partial m}(\mathbf{u}_j)\, \nabla m_j \prod_{n\neq j} h(\mathbf{u}_n),$$

$$\operatorname{Hess}_\beta \prod_{j=1}^{J} h(\mathbf{u}_j) = \sum_{j=1}^{J} \left(\frac{\partial^2 h}{\partial m^2}(\mathbf{u}_j)\, \nabla m_j \otimes \nabla m_j + \frac{\partial h}{\partial m}(\mathbf{u}_j)\, \operatorname{Hess} m_j \right) \prod_{n\neq j} h(\mathbf{u}_n)$$
$$+ \sum_{j=1}^{J} \sum_{k\neq j} \frac{\partial h}{\partial m}(\mathbf{u}_j)\, \frac{\partial h}{\partial m}(\mathbf{u}_k)\, \nabla m_j \otimes \nabla m_k \prod_{n\neq j,k} h(\mathbf{u}_n).$$

Using the uniform C^2 bounds (B.18) and (B.19) on $m_j(\beta)$, we arrive at the pointwise estimate

$$\left|\operatorname{Hess}_\beta \prod_{j=1}^{J} h(\mathbf{u}_j)\right| \lesssim \sum_{j=1}^{J} \left(\left|\frac{\partial^2 h}{\partial m^2}(\mathbf{u}_j)\right| + \left|\frac{\partial h}{\partial m}(\mathbf{u}_j)\right| \right) \prod_{n\neq j} |h(\mathbf{u}_n)|$$
$$+ \sum_{j=1}^{J} \sum_{k\neq j} \left|\frac{\partial h}{\partial m}(\mathbf{u}_j)\right| \left|\frac{\partial h}{\partial m}(\mathbf{u}_k)\right| \prod_{n\neq j,k} |h(\mathbf{u}_n)|\,.$$

B.3.1 The regime $|\xi| > J^{\frac{1}{2}}\delta$

By essentially the same calculation as in (B.22), we can estimate the product of h as:

$$\prod_{n\neq j} |h(\mathbf{u}_n)| \leq \prod_{n\neq j,k} |h(\mathbf{u}_n)| = e^{-\Theta(J)} O\left(|\xi|^{-(L+4)}\right).$$

Then, applying the bounds (B.9) on the derivatives of h gives that

$$\left|\operatorname{Hess}_\beta \prod_{j=1}^{J} h(\mathbf{u}_j)\right| \lesssim \left(\sum_{j=1}^{J} (1+|z_j|^2) + \sum_{j=1}^{J}\sum_{k\neq j} (1+|z_j|)(1+|z_k|) \right) e^{-\Theta(J)} O\left(|\xi|^{-(L+4)}\right)$$
$$\lesssim J^2(1+|\xi|^2)\, e^{-\Theta(J)} O\left(|\xi|^{-(L+4)}\right) = e^{-\Theta(J)} O\left(|\xi|^{-(L+2)}\right)$$

where, in the second step, we collected like terms after using the uniform upper bound (B.14) on the projected spins $[\xi]_j$. Integrating over the region $\{|\xi| > J^{\frac{1}{2}}\delta\}$ yields (B.5).

B.3.2 The regime $|\xi| \leq J^{\frac{1}{2}}\delta$

For the inner integral of (B.6), we again use the representation via h_2 from (B.10). In this case, the derivatives of h can be represented as

$$\frac{\partial h}{\partial m}(m,z) = -z^2 \frac{\partial h_2}{\partial m}(m,z)\; h(m,z),$$
$$\frac{\partial^2 h}{\partial m^2}(m,z) = \left(-z^2 \frac{\partial^2 h_2}{\partial m^2}(m,z) \,+\, z^4 \left(\frac{\partial h_2}{\partial m}(m,z)\right)^2 \right) h(m,z).$$

Together with the uniform bounds (B.13) on the derivatives of h_2,

$$\left|\operatorname{Hess}_\beta \prod_{j=1}^{J} h(\mathbf{u}_j)\right| \lesssim \left(\sum_{j=1}^{J} \left(|z_j|^2 + |z_j|^4\right) + \sum_{j=1}^{J}\sum_{k\neq j} |z_j|^2 |z_k|^2 \right) \prod_{n=1}^{J} |h(\mathbf{u}_n)|$$
$$\lesssim (|\xi|^2 + |\xi|^4)\, \exp\left(-c_{h_2}\left(1+O(J^{-2})\right)|\xi|^2\right)$$

where we applied the estimate (B.23) for the product of h (cf. (B.25)) and collected like terms after using the uniform upper bound (B.14) on the projected spins $[\xi]_j$. Integrating over the region $\{|\xi| \leq J^{\frac{1}{2}}\delta\}$ yields (B.6). This completes the proof of Proposition 8.10.

References

[1] S. Adams, N. Dirr, M. Peletier, and J. Zimmer, *From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage*, Communications in Mathematical Physics **307** (2010). MR2842966

[2] S. Adams, N. Dirr, M. Peletier, and J. Zimmer, *Large deviations and gradient flows*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **371** (2013), no. 2005, 20120341, 17. MR3150638

[3] L. Ambrosio, N. Gigli, and G. Savare, *Gradient flows in metric spaces and in the space of probability measures*, (2005). MR2129498

[4] D. Bakry and M. Émery, *Diffusions hypercontractives*, Sem. Probab. XIX, Lecture Notes in Math., Springer **1123** (1985), 177–206. MRMR889476

[5] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR3155209

[6] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, Oxford University Press, 02 2013. MR3185193

[7] H. J. Brascamp and E. H. Lieb, *On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation*, J. Funct. Anal. **22** (1976), no. 4, 366–389. MRMR0450480

[8] C. C. Chang and H.-T. Yau, *Fluctuations of one-dimensional Ginzburg-Landau models in nonequilibrium*, Commun. Math. Phys. **145** (1992), no. 2, 209–234. MR1162798

[9] D. A. Dawson and J. Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics **20** (1987), no. 4, 247–308. MR0885876

[10] R. A. DeVore and G. G. Lorentz, *Constructive approximation*, Grundlehren der mathematischen Wissenschaften, vol. 303, Springer, 1993. MR1261635

[11] J. Diehl, M. Gubinelli, and N. Perkowski, *The Kardar–Parisi–Zhang equation as scaling limit of weakly asymmetric interacting Brownian motions*, Commun. Math. Phys. **354** (2017), no. 2, 549–589. MR3663617

[12] D. Dizdar, *Schritte zu einer optimalen Konvergenzrate im hydrodynamischen Limes der Kawasaki Dynamik (Towards an optimal rate of convergence in the hydrodynamic limit for Kawasaki dynamics)*, Diplomarbeit, Universität Bonn (2007).

[13] H. Duong, V. Laschos, and M. Renger, *Wasserstein gradient flows from large deviations of many-particle limits*, European Series in Applied and Industrial Mathematics (ESAIM): Control, Optimization and Calculus of Variations **19** (2013). MR3182684

[14] M. Fathi, *A two-scale approach to the hydrodynamic limit part II: local Gibbs behavior*, ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 2, 625–651. MR3089703

[15] M. Fathi and M. Simon, *The gradient flow approach to hydrodynamic limits for the simple exclusion process*, (2015). MR3557721

[16] Max Fathi, *A gradient flow approach to large deviations for diffusion processes*, Journal de Mathématiques Pures et Appliquées **106** (2016), no. 5, 957–993. MR3550853

[17] J. Fritz, *On the hydrodynamic limit of a one-dimensional Ginzburg-Landau lattice model. The a priori bounds*, Journal of Statistical Physics **47** (1987), 551–572. MR0894407

[18] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR0420249

[19] N. Grunewald, F. Otto, C. Villani, and M. Westdickenberg, *A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit*, Ann. Inst. H. Poincaré Probab. Statist. **45** (2009), no. 2, 302–351. MRMR2521405

[20] A. Guionnet and B. Zegarlinski, *Lectures on logarithmic Sobolev inequalities*, Séminaire de Probabilités, XXXVI, Lecture Notes in Math., vol. 1801, Springer, Berlin, 2003, pp. 1–134. MRMR1971582

[21] 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. MRMR954674

[22] R. Holley and D. Stroock, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys. **46** (1987), 1159–1194. MRMR893137

[23] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM journal on mathematical analysis **29** (1998), no. 1, 1–17. MR1617171

[24] E. Kosygina, *The behavior of the specific entropy in the hydrodynamic scaling limit*, Ann. Probab. **29** (2001), no. 3, 1086–110. MRMR1872737

[25] M. Ledoux, *Logarithmic Sobolev inequalities for unbounded spin systems revisted*, Sem. Probab. XXXV, Lecture Notes in Math., Springer **1755** (2001), 167–194. MRMR1837286

[26] S. L. Lu and H. T. Yau, *Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics*, Commun. Math. Phys. **156** (1993), no. 2, 399–433. MRMR1233852

[27] G. Menz, *LSI for Kawasaki dynamics with weak interaction*, Commun. Math. Phys. **307** (2011), no. 3, 817–860. MR2842967

[28] A. Mielke, *On evolutionary γ-convergence for gradient systems*, pp. 187–249, Springer International Publishing, Cham, 2016. MR3468299

[29] A. Montefusco, C. Schutte, and S. Winkelmann, *A route to the hydrodynamic limit of a reaction-diffusion master equation using gradient structures*, arXiv:2201.02613 (2022). MR4580769

[30] F. Otto, *The geometry of dissipative evolution equations: The porous medium equation*, Communications in Partial Differential Equations **26** (2001), 101–174. MR1842429

[31] F. Otto and M. G. Reznikoff, *A new criterion for the logarithmic Sobolev inequality and two applications*, J. Funct. Anal. **243** (2007), no. 1, 121–157. MR2291434

[32] 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. MRMR1760620

[33] G. A. Pavliotis, *Stochastic processes and applications*, Texts in Applied Mathematics, vol. 60, Springer, New York, 2014, Diffusion processes, the Fokker-Planck and Langevin equations. MR3288096

[34] G. Royer, *Une initiation aux inégalités de Sobolev logarithmiques*, Cours Spécialisés, Soc. Math. de France (1999). MRMR1704288 MR1704288

[35] E. Sandier and S. Serfaty, *Gamma-convergence of gradient flows and application to Ginzburg–Landau*, Communications on Pure and Applied Mathematics **57** (2004), 1627–1672. MR2082242

[36] F. Santambrogio, *Euclidean, Metric, and Wasserstein gradient flows: an overview*, Bulletin of Mathematical Sciences **7** (2016). MR3625852

[37] S. Serfaty, *Gamma-convergence of gradient flows on hilbert and metric spaces and applications*, 2011, p. 1427. MR2836361

[38] H. Spohn, *Equilibrium fluctuations for interacting Brownian particles*, Commun. Math. Phys. **103** (1986), no. 1, 1–33. MR826856

[39] S. R. S. Varadhan, *Scaling limits for interacting diffusions*, Comm. Math. Phys. **135** (1991), no. 2, 313–353. MR1087387

[40] M. Zhu, *Equilibrium fluctuations for one-dimensional Ginzburg-Landau lattice model*, Nagoya Math. J. **117** (1990), 63–92. MR1044937

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: https://imstat.org/shop/donation/