Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

LSI for Kawasaki dynamics with weak interaction

by

Georg Menz

Preprint no.: 31 2010

LSI for Kawasaki dynamics with weak interaction

Georg Menz *

May 27, 2010

We consider a large lattice system of unbounded continuous spins that are governed by a Ginzburg-Landau type potential and a weak quadratic interaction. We derive the logarithmic Sobolev inequality (LSI) for Kawasaki dynamics uniform in the boundary data. The scaling of the LSI constant is optimal in the system size and our argument is independent of the geometric structure of the system. The proof consists of an application of the two-scale approach of Grunewald, Otto, Westdickenberg & Villani. Several ideas are needed to solve new technical difficulties due to the interaction. Let us mention the application of a new covariance estimate, a conditioning technique, and a generalization of the local Cramér theorem.

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

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

Contents

*Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany.

1 Introduction

The logarithmic Sobolev inequality (LSI) – introduced by Gross [Gro75] – is a powerful tool for studying spin systems. It implies exponential convergence to equilibrium of the naturally associated diffusion process and also characterizes the rate of convergence (cf. [SZ92a, SZ92b, SZ95, Yos00, Zeg96] and Remark 1.6). Therefore an appropriate scaling of the LSI constant in the system size indicates the absence of phase transitions. The LSI is also useful to deduce the hydrodynamic limit (see [Kos01, GOVW09]).

In this article we consider a large system of real-valued unbounded spins. The Hamiltonian of the system is given by a Ginzburg-Landau type potential and a two-body interaction (see (2)). The quadratic interaction is not restricted to finite range. Any two spins of the system are allowed to interact. Because Kawasaki dynamics conserves the mean spin of the system, we work with the canonical ensemble.

Even if there is no interaction term in the Hamiltonian, there is a long-range interaction due to the conservation of the mean spin. Therefore it was a challenge to establish the LSI for the canonical ensemble in the case of a non-interacting quadratic Hamiltonian (cf. [LY93] for discrete spin and [LPY02, Cha03, GOVW09] for continuous spin). The main difficulty was to attain the optimal scaling behavior of the LSI constant in the system size. In [Yau96] this result was generalized to weak interaction of finite range for bounded discrete spin values (see also [CMR02]). In this article we show that the LSI also holds for unbounded continuous spin values with weak two-body interaction. The LSI constant is uniform in the boundary data and scales optimal in the system size. Compared to the discrete case we have to deal with new technical difficulties due to the fact that the spin values and the range of interaction are unbounded. Because we apply the two-scale approach [GOVW09], deriving the hydrodynamic limit should also work. However, this is omitted in this article. Our approach is independent of the geometrical structure of the system. Hence in contrast to existing results on the hydrodynamic limit (cf. [GPV88, Yau91]) there is no restriction to lattices of certain dimensions or nearest neighbor interaction. For the proof of the main result we also establish a generalized version of the local Cramér theorem [GOVW09], where the single-site potentials have an additional linear term and depend on the site.

Recently, Felix Otto and the author of this article derived in [MO10a] the LSI for a class of non-interacting quartic Hamiltonians that also contains the classical Ginzburg-Landau potential $(x^2-1)^2$. Again, the scaling behavior of the LSI constant is optimal in the system size. It is a natural question if one could extend their result to weak interaction. Our approach allows only bounded perturbations of quadratic Hamiltonians because we use the two-scale criterion for LSI [GOVW09, Theorem 3], which is restricted to this class of Hamiltonians (cf. Remark 2.2).

The paper is organized in the following way. We present the main result in Section 1.1

and its proof in Section 2. The local Cramér theorem is discussed in Section 3. A short appendix contains basic facts about the LSI.

1.1 Basic setting and main result

We consider the following type of spin system. Let Λ be an arbitrary set of sites that are indexed by $\{1, \ldots, N\}$, $N \in \mathbb{N}$. Note that we do not make any assumption on the geometrical structure of Λ. To each site $i \in \{1, \ldots, N\}$ we assign a real value $x_i \in \mathbb{R}$ called the spin. A vector $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ represents a configuration of the spin system. The energy of a configuration is given by the Hamiltonian $H(x) \in \mathbb{R}$. In our case there are three contributions to the Hamiltonian:

- for each site $i \in \{1, \ldots, N\}$ a Ginzburg-Landau type single-site potential $\psi_i : \mathbb{R} \to \mathbb{R}$ that satisfies
$$\psi_i(x) = \frac{1}{2}\, x^2 + \delta\psi_i(x) \quad \text{and} \quad \|\delta\psi_i\|_{C^2(\mathbb{R})} \le c_1 < \infty, \tag{1}$$
uniformly in $i \in \{1, \ldots, N\}$;
- a two-body interaction given by a real-valued symmetric matrix $M = (m_{ij})_{N\times N} \in \mathbb{R}^{N\times N}$ with zero diagonal $m_{ii} = 0$;
- a linear term given by the real-valued vector $s \in \mathbb{R}^N$, which models the dependence on the boundary data (see also Remark 1.8).

Explicitly, the Hamiltonian of the system is given by
$$H(x) := \sum_{i=1}^{N} \psi_i(x_i) + \frac{1}{2} \sum_{i,j=1}^{N} m_{ij} x_i x_j + \sum_{i=1}^{N} s_i x_i. \tag{2}$$

Note that in contrast to [GOVW09] we do not consider homogeneous single-site potentials $\psi_i = \psi$, $i \in \{1, \ldots, N\}$ (cf. Remark 1.7). Note that $|m_{ij}|$ determines the strength of the interaction between the spin x_i and x_j. The sign of m_{ij} determines if the interaction is repulsive or attractive. To avoid phase transition, it is natural to assume that the interaction is small in a certain sense.

Definition 1.1 (Condition of smallness). *The interaction matrix M satisfies the smallness condition CS(ε) with $\varepsilon > 0$, if for all $x, y \in \mathbb{R}^N$*
$$\sum_{i,j=1}^{N} x_i\, |m_{ij}|\, x_j \le \varepsilon \sum_{i=1}^{N} x_i^2.$$

Remark (Alternative condition of smallness I)**.** *Note that CS(ε) is weaker than a condition Yoshida used in [Yos01] i.e.*

$$\max_{j=1\ldots N} \sum_{i=1}^{N} |m_{ij}| \leq \varepsilon \qquad \textit{and} \qquad m_{ij} = 0, \quad \textit{if} \quad |i-j| \geq R,$$

for some fixed $R \in \mathbb{N}$. There is an obvious difference between both conditions: the CS(ε) allows infinite range of interaction and Yoshida's condition not. Even if we allow infinite range of interaction in Yoshida's condition, we will give an example to distinguish both conditions. Let us consider the interaction matrix $M = (m_{ij})_{N\times N}$ given by

$$m_{ij} = \begin{cases} \frac{\varepsilon}{2\sqrt{N}}, & \textit{if } i = 1 \textit{ and } j \neq 1, \\ \frac{\varepsilon}{2\sqrt{N}}, & \textit{if } j = 1 \textit{ and } i \neq 1, \\ 0, & \textit{else}. \end{cases}$$

Let us consider the condition CS(ε). By Cauchy-Schwarz we have

$$\sum_{j=1}^{N} |x_j| \leq \left(\sum_{j=1}^{N} 1\right)^{\frac{1}{2}} \left(\sum_{j=1}^{N} |x_j|^2\right)^{\frac{1}{2}} = \sqrt{N} \left(\sum_{j=1}^{N} |x_j|^2\right)^{\frac{1}{2}}.$$

Then direct calculation reveals that

$$\begin{aligned} \sum_{i,j=1}^{N} x_i \, |m_{ij}| \, x_j &= \frac{\varepsilon}{\sqrt{N}} \, |x_1| \sum_{j=1}^{N} |x_j| \\ &\leq \varepsilon \, |x_1| \left(\sum_{j=1}^{N} |x_j|^2\right)^{\frac{1}{2}} \leq \varepsilon \sum_{j=1}^{N} |x_j|^2, \end{aligned}$$

which yields that the matrix M satisfies CS(ε).
Considering Yoshida's condition one sees directly that

$$\max_{j=1\ldots N} \sum_{i=1}^{N} |m_{ij}| = \frac{\varepsilon}{2} \left(\sqrt{N} - \frac{1}{\sqrt{N}}\right).$$

This bound is not uniform in the system size N.

Remark (Alternative condition of smallness II)**.** *Unfortunately, we cannot use the smallness condition*

$$\sum_{i,j=1}^{N} x_i m_{ij} x_j \leq \varepsilon \sum_{i=1}^{N} x_i^2, \qquad (\widetilde{\mathrm{CS}}(\varepsilon))$$

which would be sensitive to competing interaction (more precisely, to changes in the sign of m_{ij}). Obviously the condition CS(ε) implies $\widetilde{CS}(\varepsilon)$. The best $\tilde{\varepsilon}$ satisfying $\widetilde{CS}(\varepsilon)$ is given by

$$\tilde{\varepsilon} = \sup_{x\in\mathbb{R}^N} \frac{\langle Mx, x\rangle}{\langle x, x\rangle},$$

where $\langle\cdot,\cdot\rangle$ denotes the standard scalar product on $\mathbb{R}^N$. By Lagrange multipliers one sees that

$$\tilde{\varepsilon} = \max\{\lambda \geq 0\,, \lambda \in \sigma(M)\}\,,$$

where $\sigma(M)$ denotes the spectrum of M. By the same argument the best ε satisfying CS(ε) is given by

$$\varepsilon = \max\left\{\lambda \geq 0, \lambda \in \sigma(\tilde{M})\right\},$$

where $\tilde{M}$ denotes the matrix with entries $|m_{ij}|$, $i,j \in \{1,\ldots,N\}$. Let us illustrate the difference between both conditions with an example of three interacting sites. We consider two interaction matrices given by (cf. Figure 1)

$$M_1 = \begin{pmatrix} 0 & \varepsilon & -\varepsilon \\ \varepsilon & 0 & \varepsilon \\ -\varepsilon & \varepsilon & 0 \end{pmatrix} \qquad \textit{and} \qquad M_2 = \begin{pmatrix} 0 & \varepsilon & \varepsilon \\ \varepsilon & 0 & \varepsilon \\ \varepsilon & \varepsilon & 0 \end{pmatrix}.$$

Direct calculation reveals that

$$\sigma(M_1) = \{-2\varepsilon, \varepsilon, \varepsilon\} \qquad \textit{and} \qquad \sigma(M_2) = \{-\varepsilon, -\varepsilon, 2\varepsilon\}\,,$$

which implies that M_1 satisfies $\widetilde{CS}(\varepsilon)$ and M_2 satisfies $\widetilde{CS}(2\varepsilon)$. However, on the level of the condition CS the interaction matrices M_1 and M_2 are indistinguishable. We have to work with the smallness condition CS(ε), because of an application of the criterion of Otto & Reznikoff (see Section 2.2) and of a covariance estimate (see Section 2.3.2).

The grand canonical ensemble μ_{gc} is a probability measure on $\mathbb{R}^N$ given by

$$\mu_{gc}(dx) := \frac{1}{Z}\exp\left(-H(x)\right)dx.$$

Here and later on, Z denotes a generic normalization constant. The Kawasaki dynamics conserves the mean spin $m = \frac{1}{N}\sum_{i=1}^{N} x_i$ of an initial configuration $x \in \mathbb{R}^N$. Therefore we want to restrict the system to the $N-1$ dimensional hyper-plane $X_{N,m}$ defined by

$$X_{N,m} := \left\{x \in \mathbb{R}^N \mid \frac{1}{N}\sum_{i=1}^{N} x_i = m\right\}.$$

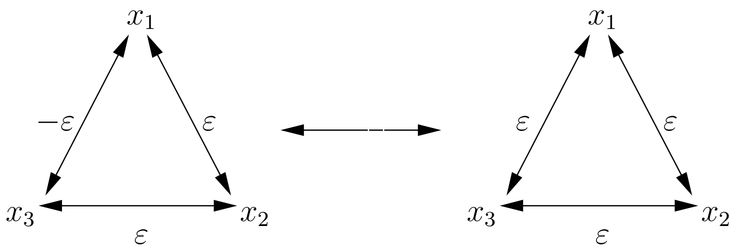

Figure 1: Competing vs. uniform interaction

We equip $X_{N,m}$ with the standard scalar product and norm induced by $\mathbb{R}^N$ i.e.

$$\langle x, \tilde{x} \rangle := \sum_{i=1}^{N} x_i \tilde{x}_i \qquad \text{and} \qquad |x| = \sum_{i=1}^{M} x_i^2.$$

The restriction of μ_{gc} to $X_{N,m}$ is called canonical ensemble μ i.e.

$$\mu(dx) := \frac{1}{Z} \, 1_{\left\{ \frac{1}{N} \sum_{i=1}^{N} x_i = m \right\}} \exp\left(-H(x)\right) \; \mathcal{H}(dx). \tag{3}$$

Here and later on, $\mathcal{H}$ denotes the Hausdorff measure in the appropriate dimension. Note that the canonical ensemble μ depends on the system size $N \in \mathbb{N}$ and the mean spin $m \in \mathbb{R}$.

Definition 1.2 (LSI)**.** *Let X be an Euclidean space. A Borel probability measure μ on X satisfies the LSI with constant $\varrho > 0$ (in short: LSI(ϱ)), if for all functions $f \geq 0$*

$$\operatorname{Ent}\left(f\mu|\mu\right) := \int f \log f \; d\mu - \int f d\mu \log\left(\int f d\mu\right) \leq \frac{1}{2\varrho} \int \frac{|\nabla f|^2}{f} d\mu. \tag{4}$$

Here ∇ denotes the gradient determined by the Euclidean structure of X.

Remark 1.3 (Gradient on $X_{N,m}$)**.** *If we choose in Definition 1.2 $X = X_{N,m}$, we can calculate $|\nabla f|^2$ in the following way: Extend the function $f : X_{N,m} \to \mathbb{R}$ to be constant on the direction normal to $X_{N,m}$, then*

$$|\nabla f|^2 = \sum_{i=1}^{N} |\partial_{x_i} f|^2.$$

Now, we are able to state the main result of this article.

Theorem 1.4. *Assume that the Hamiltonian H is given by* (2) *and that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Then there exist $\varepsilon > 0$ and $\varrho > 0$ depending only on c_1 such that:*
If the interaction matrix M satisfies CS(ε), then the canonical ensemble μ satisfies LSI(ϱ) independent of N, m, and s.

In the next remark we explain in which sense the scaling behavior of Theorem 1.4 is optimal in the system size.

Remark 1.5 (From Glauber to Kawasaki)**.** *The bound on the right hand side of* (4) *is given in terms of the Glauber dynamics in the sense that we have endowed $X_{N,m}$ with the standard Euclidean structure inherited from $\mathbb{R}^N$. By the discrete Poincaré inequality one can recover the bound for the Kawasaki dynamics (cf. [Cap03] or [GOVW09, Remark 15]) in the sense that one endows $X_{N,m}$ with the Euclidean structure coming from the discrete H^{-1}-norm. More precisely, let A denote the discrete second-order difference operator. For example for a one-dimensional lattice A is given by*

$$A := \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ & & & & \\ 0 & \cdots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

On the d-dimensional lattice we consider a cube Λ with side length W and total number of sites N. Caputo showed in [Cap03] that there exists a constant C depending only on the dimension d such that

$$|\nabla f|^2 \leq CW^2 |\sqrt{A}\nabla f|^2.$$

Thus if a measure μ satisfies LSI(ϱ), then for all functions $f \geq 0$

$$\mathrm{Ent}\,(f\mu|\mu) \leq \frac{CW^2}{2\varrho} \int \frac{|\sqrt{A}\nabla f|^2}{f} d\mu.$$

The diffusive scaling behavior in W is optimal (cf. [Yau96]).

In the next remark we explain what is meant by Kawasaki Dynamics and how the LSI is connected to exponential convergence to equilibrium.

Remark 1.6 (Convergence to equilibrium)**.** *The Kawasaki dynamics on $X_{N,m}$ is given by the stochastic differential equation*

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

where B_t denotes a standard Brownian motion on $\mathbb{R}^N$. If the process X_t is distributed as $f_t\mu$, then f_t satisfies the time-evolution

$$\frac{d}{dt}(f_t\mu) = \nabla \cdot (A\nabla f_t\mu)\,.$$

Using this equation one sees that

$$\frac{d}{dt}\,\mathrm{Ent}\,(f_t\mu|\mu) = -\frac{1}{2}\int \frac{|\sqrt{A}\nabla f|^2}{f}d\mu.$$

Hence it follows from Gronwall's Lemma and Remark 1.5 that if μ satisfies LSI(ϱ), then

$$\mathrm{Ent}\,(f_t\mu|\mu) \leq \exp\left(-CW^{-2}\varrho\right)\mathrm{Ent}\,(f_0\mu|\mu)$$

with a constant C, that depends only on the dimension of the lattice.

There are several criteria for the LSI (cf. Appendix A), but none of them applies to our situation:

- The **Tensorization Principle** for LSI does not apply because of the interaction $M \neq 0$.
- The criterion of **Bakry-Emery** does not apply because the single-site potentials ψ_i are allowed to be non-convex.
- The criterion of **Holley-Stroock** does not help because we want the LSI constant ϱ to be independent of the system size N.
- The criterion of **Otto & Reznikoff** does not help because of the restriction to the hyper-plane $X_{N,m}$.

Therefore new tools are needed. The most common approach to LSI for Kawasaki dynamics is the Lu-Yau martingale method [LY93, LPY02, Cha03]. Using this method Landim, Panizo & Yau [LPY02] proved Theorem 1.4 in the special case $M = 0$ for the Kawasaki bound. An adaptation of this approach by Chafaï [Cha03] led to the stronger bound for Glauber Dynamics. Providing a new technique – called the two-scale approach – Grunewald, Otto, Westdickenberg (former Reznikoff) and Villani [GOVW09] reproduced Theorem 1.4 for $M = 0$. We will follow their approach, but our setting differs in two aspects: On the one hand we consider inhomogeneous single-site potentials (i.e. ψ_i depends on the site i) and on the other hand –and more fundamentally– we allow for interaction $M \neq 0$. These differences lead to new technical difficulties, which are dealt with using the following ideas (see also Section 2.1):

- the interaction between blocks is controlled by the covariance estimate of [MO10b];

- the convexification of the coarse-grained Hamiltonian with interaction is achieved by a conditioning technique (that artificially reduces the system size) and a non-standard perturbation argument;
- the local Cramér theorem (cf. [GOVW09, Proposition 31]) is generalized to Hamiltonians with inhomogeneous single-site potentials and linear terms.

In our argument we want to point out the proof of Lemma 2.12 which contains crucial estimates to deal with unbounded spin values and unbounded range of interaction. If the spin values and the range of interaction are bounded, the proof of Lemma 2.12 would be a lot easier (cf. comments after (30)). As a consequence one would not have to generalize the local Cramér theorem. Therefore Lemma 2.12 is the interesting step compared to the discrete and bounded case (cf. [Yau96, CMR02]).

The proof of the main result (Theorem 1.4) is structured in the following way. In Section 2.1 we outline the two-scale approach. The proof of the main result is given directly after we formulated the two-scale criterion for LSI (see Theorem 2.1), which is the main tool of our argument. In the remaining part of Chapter 2 we verify the ingredients of the two-scale criterion. The microscopic LSI and the macroscopic LSI are deduced in Section 2.2 and Section 2.3 respectively. For the proof of the macroscopic LSI we need a generalized version of the local Cramér theorem, which we state and prove in Section 3.

Remark 1.7 (Homogeneous single-site potentials)**.** *Let us consider the situation of homogeneous single-site potentials. More precisely, assume that for $i \in \{1, \dots N\}$ the single site-potential $\psi_i(x)$, $x \in \mathbb{R}$ is given by*

$$\psi_i(x) := \psi(x) := \frac{1}{2}\, x^2 + \delta\psi(x), \qquad \text{and} \qquad \|\delta\psi\|_{C^2} \leq c_1.$$

Then the proof of Theorem 1.4 would be exactly the same as for inhomogeneous single-site potentials except of one detail: one has to generalize local Cramér theorem (cf. Theorem 3.1 or Corollary 3.2) only to Hamiltonians of the form

$$H_h(x) := \sum_{i=1}^{K} \frac{1}{2} x_i^2 + s_i x_i + \delta\psi(x_i).$$

and not to Hamiltonians of the form

$$H_{ih}(x) := \sum_{i=1}^{K} \frac{1}{2} x_i^2 + s_i x_i + \delta\psi_i(x_i).$$

However, the argument for the generalized local Cramér theorem is almost the same for Hamiltonians given by H_h or H_{ih}. Therefore we decided to state the main result in the more general case of inhomogeneous single-site potentials.

In the next remark we explain that the dependence on the boundary data can be expressed by adding a linear term to the Hamiltonian H.

Remark 1.8 (Introducing boundary data)**.** *We start with a system of N sites and a Hamiltonian that just consists of single-site potentials and interaction i.e.*

$$H(x_1,\ldots,x_N) := \sum_{i=1}^{N} \psi_i(x_i) + \frac{1}{2}\sum_{i,j=1}^{N} m_{ij} x_i x_j.$$

To model the dependence on the boundary data, we embed the original system into a larger system of $\tilde{N} \geq N$ spins, $\tilde{N} \in \mathbb{N}$. Because the indexing of the sites is arbitrary, we can assume that the new sites are indexed by the set $\{N+1,\ldots,\tilde{N}\}$. The new system has the Hamiltonian

$$\tilde{H}(x_1,\ldots x_{\tilde{N}}) := \sum_{i=1}^{\tilde{N}} \psi_i(x_i) + \frac{1}{2}\sum_{i,j=1}^{\tilde{N}} m_{ij} x_i x_j,$$

where the real numbers $m_{ji} = m_{ij}$, $i \in \{1,\ldots,\tilde{N}\}, j \in \{N+1,\ldots\tilde{N}\}$, express the interaction with the additional sites $\{N+1,\ldots,\tilde{N}\}$. We consider the new grand canonical ensemble $\tilde{\mu}_{gc}$ on $\mathbb{R}^{\tilde{N}}$ associated to $\tilde{H}$,

$$\tilde{\mu}_{gc}(dx) := \frac{1}{Z}\exp\left(-\tilde{H}(x)\right) dx.$$

We fix the values of the new spins x_i, $i \in \{N, N+1,\ldots,\tilde{N}\}$, that now play the role of the boundary data of the original system. Hence, we restrict $\tilde{\mu}_{gc}$ to the space

$$\left\{y \in \mathbb{R}^{\tilde{N}} \mid y_i = x_i,\ N+1 \leq i \leq \tilde{N}\right\} \approx \mathbb{R}^N$$

and denote the restriction as

$$\mu_b(dx) := \frac{1}{Z}\,1_{\left\{y_i=x_i,\ N+1\leq i\leq \tilde{N}\right\}} \exp\left(-\tilde{H}(x)\right)\ \mathcal{H}(dx).$$

Changing the normalization constant Z one can cancel all terms of the Hamiltonian $\tilde{H}$ that are independent of $(x_1,\ldots,x_N)$. Therefore μ_b can be considered as a measure on $\mathbb{R}^N$ with density

$$\mu_b(dx) = \frac{1}{Z}\ \exp\left(-H_b(x)\right)\ dx,$$

where $H_b(x)$, $x \in \mathbb{R}^N$, is given by

$$H_b(x) := \sum_{i=1}^{N} \psi_i(x_i) + \frac{1}{2}\sum_{i,j=1}^{N} m_{ij} x_i x_j + \sum_{i=1}^{N} \underbrace{\left(\sum_{j=N+1}^{\tilde{N}} m_{ij} x_j\right)}_{=:s_i} x_i.$$

This calculation shows that in our setup the dependence on the boundary data can be modelled by adding a linear term to the Hamiltonian.

2 Proof of the main result

2.1 The two-scale approach

In this section we will explain the two-scale approach, point out the new difficulties arising from the interaction, and explain how they are solved. Our presentation is based on Subsection 2.1 and 5.1 of [GOVW09], which we recommend for better understanding. We decompose the spin system into L blocks each containing K sites. Therefore $N = KL$. The index set of the l-th block, $l \in \{1, \ldots, L\}$ is given by (cf. Figure 2)

$$B(l) := \{(l-1)K + 1, \ldots, l\}.$$

Hence a configuration $x \in X_{N,m}$ of the spin system can be written as

$$x = (x^1, \ldots, x^L), \quad \text{with } x^l := (x_i)_{i \in B(l)}. \tag{5}$$

Note that the block decomposition is entirely arbitrary and has no geometric significance. The coarse-graining operator $P : X_{N,m} \to X_{L,m} =: Y$ assigns to each block its mean spin i.e.

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

We endow Y with the scalar product

$$\langle y, z \rangle_Y := \frac{1}{L} \sum_{i=1}^{L} y_i z_i, \qquad \text{for } y, z \in Y. \tag{6}$$

Let $P^* : Y \to X_{N,m}$ denote the adjoint operator of P. Note that due to the special Euclidean structure of Y we have

$$P^* = \frac{1}{L} P^t,$$

where P^t denotes the transpose of P. The orthogonal projection of $X_{N,m}$ on $\ker P$ is given by $\mathrm{Id} - NP^*P$, which can be seen using the identity

$$PNP^* = \mathrm{Id}_Y \, . \tag{7}$$

Hence we can decompose $x \in X_{N,m}$ into a macroscopic profile and a microscopic fluctuation i.e.

$$x = \underbrace{(NP^*P)\, x}_{\in (\ker P)^{\perp}} + \underbrace{(\mathrm{Id} - NP^*P) x}_{\in \, \ker P} .$$

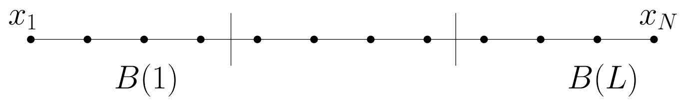

Figure 2: Block decomposition of the spin system

The coarse graining also induces a natural decomposition of measures. Recall that μ denotes the canonical ensemble given by (3) associated to the Hamiltonian H and the mean spin m. Let $\bar{\mu} := P_{\#}\mu$ be the push forward of μ under P and let $\mu(dx|y)$ denote the conditional measure of x given $Px = y$. Then by disintegration

$$\mu(dx) = \mu(dx|y)\bar{\mu}(dy). \tag{8}$$

This equation has to be understood in a weak sense i.e. for any test function ξ

$$\int \xi \, d\mu = \int_Y \left(\int_{\{Px=y\}} \xi \; \mu(dx|y) \right) \bar{\mu}(dy).$$

By the Coarea Formula one can determine the density of $\bar{\mu}(dy)$ as

$$\bar{\mu}(dy) = \exp(-N\bar{H}(y)) \, dy,$$

where the coarse-grained Hamiltonian $\bar{H}$ is given by

$$\bar{H}(y) := -\frac{1}{N} \log \int_{\{Px=y\}} \exp(-H(x)) \, \mathcal{H}(dx). \tag{9}$$

The coarse-grained Hamiltonian $\bar{H}(y)$ represents the energy of a macroscopic profile y. Overall, we observe the system at two different scales:

- the microscopic scale $\mu(dx|y)$ considers all fluctuations of the system around a macroscopic profile $y \in Y$, and
- the macroscopic scale $\bar{\mu}(dy)$ considers the macroscopic profiles and neglects all fluctuations.

We want to apply the two-scale criterion for LSI (see [GOVW09, Theorem 3]) to derive the LSI for μ. In our setting the two-scale criterion becomes

Theorem 2.1 (Two-scale criterion)**.** *Assume that the Hamiltonian H is given by* (2) *and that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Assume that:*

(i) *There is* $\varrho > 0$ *such that for all* N, m, s, *and* $y \in Y$ *the conditional measures* $\mu(dx|y)$ *satisfy LSI(*ϱ*).*

(ii) *There is* $\lambda > 0$ *such that for all* N, m, *and* s *the marginal* $\bar{\mu}$ *satisfies LSI(*λN*).*

Then μ *satisfies LSI(*$\hat{\varrho}$ *) with* $\hat{\varrho}$ *independent of* N, m, *and* s.

Remark. *The two-scale criterion in [GOVW09] also contains an explicit representation of the LSI constant* $\hat{\varrho}$ *in terms of* ϱ, λ, *and a constant* κ, *which represents the strength of the coupling between the microscopic and macroscopic scale. However, for our purpose it is just important that* $\hat{\varrho}$ *is independent of the system size* N, *the mean spin* m, *and the boundary data* s.

Remark 2.2. *In the introduction we mentioned the question of generalizing the main result (Theorem 1.4) to quartic Hamiltonians* H *i.e.*

$$H(x) := \sum_{i=1}^{N} \left(\frac{x_i^4}{4} + \delta\psi_i(x_i) \right) + \frac{1}{2} \sum_{i,j=1}^{N} m_{ij} x_i x_j + \sum_{i=1}^{N} s_i x_i,$$

where the functions $\delta\psi_i(x_i) : \mathbb{R} \to \mathbb{R}$ *satisfy the uniform bound*

$$\sup_{x \in \mathbb{R}} |\delta\psi_i''(x)| \leq c_1 < \infty.$$

Our approach does not cover that case because Theorem 2.1 (or [GOVW09, Theorem 3]) cannot be applied to quartic Hamiltonians. The reason is that $\kappa = \infty$ *in that case (see the last remark and [MO10a]). In [MO10a] this problem is avoided by an adaptation of the two-scale criterion [GOVW09, Theorem 3] to the quartic case. The adapted two-scale criterion is formulated only for coarsening of pairs and the main result of [MO10a] is established by an iteration of dyadic coarsening. This approach cannot be transferred to the interactive case directly, because the iteration is based on a product structure which is lost due to non-zero interaction.*

The ingredients *(i)* and *(ii)* of Theorem 2.1 are deduced in Section 2.2 and Section 2.3 respectively. Provided these ingredients are satisfied, the proof of the main result (see Theorem 1.4) is just an application of the two-scale criterion (see Theorem 2.1).

Proof of Theorem 1.4. We carry out the coarse-graining procedure with a large but fixed block size $K \geq K_0$, where K_0 is determined by Proposition 2.4 below. Note that K_0 is independent of the system size N, the mean spin m, and the boundary data s. The ingredients of the two-scale criterion of Theorem 2.1, namely the microscopic LSI and the macroscopic LSI, are verified by Proposition 2.3 and Corollary 2.5 respectively. Then Theorem 1.4 follows directly from an application of Theorem 2.1. □

Now, we will discuss how the ingredients of Theorem 2.1 are verified. The microscopic LSI follows directly from an application of the Otto & Reznikoff criterion for LSI (see Subsection 2.2). Difficulties arise deducing the macroscopic LSI (see Subsection 2.3). We follow the strategy of [GOVW09] and want to show that $\bar{H}$ is uniformly convex if the block size K is large enough and the interaction ε is small enough. The uniform convexity of $\bar{H}$ would yield the macroscopic LSI by the criterion of Bakry-Emery (see Theorem A.3). Due to the interaction between blocks we lose the product structure of $\bar{\mu}$ (cf. [GOVW09, (63)]), that was crucial for the argument of [GOVW09]. As a consequence, the off-diagonal entries $h_{ln}, l \neq n$, of the Hessian of $\bar{H}$ become non trivial (see (16)). However, applying a new covariance estimate [MO10b] yields sufficient control of h_{ln} in terms of ε (see Subsection 2.3.2).

The main difficulty of the proof is encountered when checking the positivity of the diagonal elements h_{ll} of the Hessian of $\bar{H}$. It is not possible to transfer the positivity of h_{ll} from the case of $\varepsilon = 0$ to the case of small ε by a simple perturbation argument. The reason is that due to the loss of the product structure, h_{ll} will depend on all spins of system. In the case $\varepsilon = 0$ the diagonal elements h_{ll} depend only on the spins of the l-th block, which has size K. Hence one could not choose ε independent from the system size N and the LSI constant would depend on N. We avoid this problem by conditioning on all spins except of a single block (see Subsection 2.3.3). This procedure artificially reduces the system size to the number K and introduces new boundary data, which is expressed by an additional linear term in the Hamiltonian (cf. Remark 1.8). Independently, we observe that for $\varepsilon = 0$ the positivity of h_{ll} for large K is untouched by a linear term (cf. Remark 3.3). Therefore we are able to apply a perturbation argument to transfer the positivity of h_{ll} to small ε depending only on K and c_1 and not on the system size N (see Lemma 2.11 and Lemma 2.12).

2.2 Microscopic LSI

In this subsection we will prove the following statement.

Proposition 2.3 (Microscopic LSI)**.** *Assume that the Hamiltonian H is given by* (2) *and that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Then there is $\varepsilon > 0$ independent of N, m, s, and $y \in Y$ (depending only on the block size K and c_1) such that:*
If M satisfies CS(ε), then the conditional measures $\mu(dx|y)$ given by (8) *satisfy LSI(ϱ) with $\varrho > 0$ independent of N, m, s and y (depending only on K, c_1, and ε).*

Proof of Proposition 2.3. The statement follows from a direct application of the criterion for LSI of Otto & Reznikoff (see Theorem A.4 in the appendix or Theorem 1 in

[OR07]). We check the first requirement of Theorem A.4. Recall that for a given configuration $x \in \mathbb{R}^N$ the spins inside the block $B(l)$ are denoted by $x^l := (x_i)_{i \in B(l)}$. For a fixed $y = (y_1, \ldots, y_L) \in Y$ we define

$$X_l := \left\{ x^l \in \mathbb{R}^{B(l)} \mid \frac{1}{K} \sum_{i \in B(l)} x_i = y_l \right\}, \quad l \in \{1, \ldots, L\}.$$

Because of the block decomposition (5) we have

$$\{x \in \mathbb{R}^N \mid Px = y\} = X_1 \times \ldots \times X_L.$$

Hence we can decompose a configuration $x \in X$ as

$$x = (x^1, \ldots, x^L) \quad \text{with} \quad x^i \in X_i.$$

For convenience, the spins outside the block $B(l)$ are denoted by $\bar{x}^l := (x_i)_{i \notin B(l)}$. Disintegration of the microscopic measure $\mu(dx|y)$ with respect to x^l yields

$$\mu(dx|y) = \mu(dx^l|\bar{x}^l, y)\ \bar{\mu}(d\bar{x}^l|y),$$

where $\mu(dx^l|\bar{x}^l, y)$ and $\bar{\mu}(d\bar{x}^l|y)$ denotes the conditional measure and the corresponding marginal respectively (cf. Figure 3 below). More precisely, we have for all test functions $\xi : \{Px = y\} \to \mathbb{R}$

$$\int \xi(x)\mu(dx|y) = \int\int \xi(x^l, \bar{x}^l)\mu(dx^l|\bar{x}^l, y)\bar{\mu}(d\bar{x}^l|y). \tag{10}$$

For the first requirement of Theorem A.4 we have to show that on X_l, $1 \le l \le L$, the conditional measures

$$\mu(dx^l|\bar{x}^l, y)$$

satisfy the LSI($\tilde{\varrho}$) with constant $\tilde{\varrho} > 0$ independent of N, m, s, y, l and $\bar{x}^l$. For this purpose let us have a closer look at the Hamiltonian of the conditional measure $\mu(dx^l|\bar{x}^l, y)$. For an arbitrary vector $s^* \in \mathbb{R}^{B(l)}$ we define the Hamiltonian $H(x^l|M, s^*)$ by

$$H(x^l|M, s^*) = \sum_{i \in B(l)} \psi_i(x_i) + \frac{1}{2} \sum_{i,j \in B(l)} m_{ij} x_i x_j + \sum_{i \in B(l)} s_i^* x_i.$$

The definition (2) of the Hamiltonian H yields

$$\begin{aligned}
H(x) &= \sum_{i=1}^{N} \psi_i(x_i) + \frac{1}{2}\sum_{i,j=1}^{N} m_{ij}x_ix_j + \sum_{i=1}^{N} s_ix_i \\
&= \sum_{i\in B(l)} \psi_i(x_i) + \frac{1}{2}\sum_{i,j\in B(l)} m_{ij}x_ix_j + \sum_{i\in B(l)} \left(s_i + \sum_{j\notin B(l)} m_{ij}x_j \right) x_i \\
&\quad + \sum_{i\notin B(l)} \psi_i(x_i) + \frac{1}{2}\sum_{i,j\notin B(l)} m_{ij}x_ix_j + \sum_{i\notin B(l)} s_ix_i \\
&= H\left(x^l|M, s_c\right) + \sum_{i\notin B(l)} \psi_i(x_i) + \frac{1}{2}\sum_{i,j\notin B(l)} m_{ij}x_ix_j + \sum_{i\notin B(l)} s_ix_i,
\end{aligned}$$

where the vector $s_c = s_c(s, M, \bar{x}^l) \in \mathbb{R}^{B(l)}$ is defined by

$$s_{c,i} := s_i + \sum_{j\notin B(l)} m_{ij}x_j \qquad \text{for } i \in B(l).$$

Because one can cancel all terms that are independent of $x^l = (x_i)_{i\in B(l)}$ with terms of the normalization constant Z, the effective Hamiltonian of the conditional measure $\mu(dx^l|\bar{x}^l, y)$ is given by $H(x^l|M, s_c)$. More precisely,

$$\mu(dx^l|\bar{x}^l, y) = \frac{1}{Z}\, 1_{\left\{\frac{1}{K}\sum_{i\in B(l)} x_i = y_l\right\}} \exp\left(-H(x^l|M, s_c)\right) \mathcal{H}(dx).$$

Note that the dependence of $H(x^l|M, s_c)$ on y_l is hidden in the condition

$$x^l \in X_l = \left\{ x^l \in \mathbb{R}^{B(l)} \mid \frac{1}{K}\sum_{i\in B(l)} x_i = y_l \right\}.$$

Using the assumption (1) on the single-site potentials ψ_i we can rewrite $H(x^l|M, s_c)$ as

$$H(x^l|M, s_c) = \underbrace{\sum_{i\in B(l)} \left[\frac{x_i^2}{2} + \left(s_i + \sum_{j\notin B(l)} m_{ij}x_j \right) x_i \right] + \frac{1}{2}\sum_{i,j\in B(l)} m_{ij}x_ix_j}_{=:H_1(x^l|M,s_c)} + \underbrace{\sum_{i\in B(l)} \delta\psi_i(x_i)}_{=:H_2(x^l|M,s_c)}.$$

Using CS(ε) it follows that

$$\sum_{i,j\in B(l)} m_{ij}x_ix_j \le \varepsilon |x^l|^2.$$

Hence if ε is small enough, then $H_1(x^l|M, s_c)$ is a uniformly strictly convex function with constant $\lambda \geq \frac{1}{4}$. By the assumption (1) on the functions $\delta\psi_i$ it follows that $H_2(x^l|M, s_c)$ is a bounded function satisfying

$$\left| \sup_{x^l \in X_l} H_2(x^l|M, s_c) - \inf_{x^l \in X_l} H_2(x^l|M, s_c) \right| \leq K \max_{i \in \{1,\ldots,N\}} |\sup \delta\psi_i - \inf \delta\psi_i| \leq 2Kc_1.$$

Therefore by a combination of the criterion of Bakry & Emery (see Theorem A.3) and of the criterion of Holley & Stroock (see Theorem A.2), the conditional measures $\mu(dx^l|\bar{x}^l, y)$ satisfy a uniform LSI with constant

$$\tilde{\varrho} = \exp\left(-2Kc_1\right) \frac{1}{4}.$$

Note that $\tilde{\varrho}$ is independent of N, m, s, y, l and $\bar{x}^l$ (depending only on the block size K and the constant c_1 given by (1)).
Now, we verify the remaining ingredients of the criterion of Otto & Reznikoff. For $n, m \in \{1, \ldots, L\}$ let M_{nm} denote the $K \times K$ matrix

$$M_{nm} = (m_{ij})_{i \in B(n),\ j \in B(m)}.$$

Let $\|M_{nm}\|$ be defined as the operator norm of M_{nm} as a bilinear form i.e.

$$\|M_{nm}\| = \max \left\{ \sum_{i \in B(n),\ j \in B(m)} \frac{x_i m_{ij} y_j}{|x|\ |y|} \ \middle|\ x \in \mathbb{R}^{B(n)}, \tilde{y} \in \mathbb{R}^{B(m)} \right\}.$$

Let the matrix $A = (a_{nm})_{K \times K}$ be defined by the elements

$$a_{nm} = \begin{cases} \tilde{\varrho}, & \text{if} \quad n = m, \\ -\|M_{nm}\|, & \text{if} \quad n \neq m, \end{cases} \qquad n, m \in \{1, \ldots, K\}. \tag{11}$$

We have to show that A satisfies for some $\varrho > 0$ independent of N, m,s, y, l and $\bar{x}^l$

$$A \geq \varrho \, \mathrm{Id}$$

in the sense of quadratic forms. For the rest of the proof let $C < \infty$ denote a generic constant that depends only on K. Firstly, we will show that

$$\left(\|M_{nm}\|\right)_{L \times L} \leq C\varepsilon \ \mathrm{Id}. \tag{12}$$

in the sense of quadratic forms. Because of the equivalence of norms in a finite dimensional vector space we have for $n, m \in \{1, \ldots, L\}$

$$\|M_{nm}\| \leq C \sum_{i \in B(n), j \in B(m)} |m_{ij}|.$$

For any vector $x \in \mathbb{R}^L$ we have

$$\begin{aligned}\sum_{n,m=1}^{L} x_n \, \|M_{nm}\| \, x_m \quad &\leq \quad C \sum_{n,m=1}^{L} \sum_{i \in B(n), j \in B(m)} |x_n| \, |m_{ij}| \, |x_m| \\ &\overset{\mathrm{CS}(\varepsilon)}{\leq} \quad C\varepsilon \sum_{n=1}^{L} x_n^2.\end{aligned}$$

This inequality already yields (12). Because $\tilde{\varrho}$ depends only on the block size K and c_1, we can choose ε small independent of N and y such that

$$\begin{aligned}A &= \tilde{\varrho}\,\mathrm{Id} - (\|M_{nm}\|)_{L\times L} + \mathrm{diag}\,(\|M_{11}\|, \ldots, \|M_{LL}\|) \\ &\geq \tilde{\varrho}\,\mathrm{Id} - (\|M_{nm}\|)_{L\times L} \\ &\geq (\tilde{\varrho} - C\varepsilon)\,\mathrm{Id} \\ &\geq \varrho\,\mathrm{Id}, \qquad (13)\end{aligned}$$

for some $\varrho > 0$ depending only on K, c_1, and ε. Hence we can apply the criterion of Otto & Reznikoff and get that the conditional measures $\mu(dx|y)$ satisfy LSI(ϱ) uniformly in N, m, s, and y, if the strength of interaction ε is small enough. □

2.3 Macroscopic LSI

In this section we will derive the macroscopic LSI. More precisely, we will prove that $\bar{H}$ becomes uniformly convex for large K and small ε.

Proposition 2.4. *Assume that the Hamiltonian H is given by* (2) *and that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Let $\bar{H}$ denote the coarse-grained Hamiltonian defined by* (9) *and let $\mathrm{Hess}_Y \bar{H}$ denote the Hessian of $\bar{H}$ w.r.t. the Euclidean structure $\langle \cdot, \cdot \rangle_Y$ on Y given by* (6). *Then there exists $K_0 \in \mathbb{N}$ depending only on c_1 such that:*
If the block size $K \geq K_0$ and the interaction matrix M satisfies $\mathrm{CS}(\varepsilon)$, then there are constants $\lambda > 0$ and $C < \infty$ independent of N, m, and s (depending only on K and c_1) such that for all $y \in Y$

$$\mathrm{Hess}_Y \bar{H}(y) \geq (\lambda - C\varepsilon)\,\mathrm{Id}$$

in the sense of quadratic forms.

By the definition (9) of $\bar{H}$ we have $\bar{\mu}(dy) = \exp(-N\bar{H}(y))\mathcal{H}(dy)$. Hence the macroscopic LSI is a direct consequence of Proposition 2.4 and the criterion of Bakry & Emery (see Theorem A.3), if we choose ε small enough. More precisely, we have

Corollary 2.5 (Macroscopic LSI). *Assume that the Hamiltonian H is given by* (2) *and that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Choose a fixed block size $K \geq K_0$ where K_0 is given by Proposition 2.4. Consider the marginal $\bar{\mu}$ defined by* (8). *Then there exist $\varepsilon > 0$ and $\lambda > 0$ independent of N, m, and s (depending only on K and c_1) such that:*
If the interaction matrix M satisfies $\mathrm{CS}(\varepsilon)$, *then $\bar{\mu}$ satisfies LSI(λN).*

The proof of Proposition 2.4 consists of three steps. In the next subsection we will deduce a formula for the elements of $\operatorname{Hess}_Y \bar{H}$. In Subsection 2.3.2 we will show that the off-diagonal elements of $\operatorname{Hess}_Y \bar{H}$ are small in a certain sense (cf. Lemma 2.8). In Subsection 2.3.3 we will show that the diagonal elements of $\operatorname{Hess}_Y \bar{H}$ are uniformly positive for large K and small ε (cf. Lemma 2.10).

Proof of Proposition 2.4. We decompose the $\operatorname{Hess}_Y \bar{H}(y)$ into its diagonal matrix and its remainder i.e.

$$\begin{aligned}\operatorname{Hess}_Y \bar{H}(y) = \operatorname{diag}\left(\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{11}, \ldots, \left(\operatorname{Hess}_Y \bar{H}(y)\right)_{LL}\right) \\ + \left[\operatorname{Hess}_Y \bar{H}(y) - \operatorname{diag}\left(\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{11}, \ldots, \left(\operatorname{Hess}_Y \bar{H}(y)\right)_{LL}\right)\right]\end{aligned}$$

Then a combination of Lemma 2.8 and Lemma 2.10 from below yields the statement. □

2.3.1 Formula for the elements of the Hessian of $\bar{H}$

Before we derive the formula for the elements of the Hessian of $\bar{H}$, we will deduce an alternative representation of the coarse-grained Hamiltonian $\bar{H}$.

Lemma 2.6. *Assume that the Hamiltonian H and the coarse-grained Hamiltonian $\bar{H}$ are given by* (2) *and* (9) *respectively. For $x \in \{Px = 0\}$ and $y \in Y$ let $H_M(x, y)$ be defined by*

$$H_M(x,y) := \frac{1}{2}\langle x, (\mathrm{Id} + M)x\rangle + \langle x, MNP^*y\rangle + \langle s, x\rangle + \sum_{i=1}^{N} \delta\psi_i(x_i + (NP^*)_i).$$

Then

$$\bar{H}(y) = \frac{1}{2}\langle y, (\mathrm{Id} + PMNP^*)y\rangle_Y + \langle Ps, y\rangle_Y - \frac{1}{N}\log\int_{\{Px=0\}} \exp\left(-H_M(x,y)\right)\,\mathcal{H}(dx), \tag{14}$$

where the scalar product $\langle\cdot,\cdot\rangle_Y$ is given by (6)

Proof of Lemma 2.6. Because Hamiltonian of our system is given by (2) it follows that

$$H(x + NP^*y) = \frac{1}{2}\langle NP^*y, (\mathrm{Id} + M)NP^*y\rangle + \langle x, (\mathrm{Id} + M)NP^*y\rangle + \frac{1}{2}\langle x, (\mathrm{Id} + M)x\rangle + \langle s, x\rangle + \langle s, NP^*y\rangle + \sum_{i=1}^{N} \delta\psi_i(x_i + (NP^*y)_i).$$

Because $x \in \ker P$, we have $\langle x, NP^*y\rangle = 0$. Additionally, note that

$$\langle s, NP^*y\rangle = N\langle Ps, y\rangle_Y .$$

Hence the equality from above yields

$$H(x + NP^*y) = \frac{1}{2}N\langle y, (\mathrm{Id} + PMNP^*)y\rangle_Y + N\langle Ps, y\rangle_Y + H_M(x, y), \tag{15}$$

where we used the definition of $H_M(x, y)$ from above. Note that by (7) we have

$$\begin{aligned}\left\{x \in \mathbb{R}^N \mid Px = y\right\} &= \left\{x \in \mathbb{R}^N \mid Px = PNP^*y\right\} \\ &= \left\{x \in \mathbb{R}^N \mid P(x - NP^*y) = 0\right\}.\end{aligned}$$

This identity yields by the translation $x \mapsto x - NP^*y$ that

$$\begin{aligned}\bar{H}(y) &\overset{(9)}{=} -\frac{1}{N}\log \int_{\{Px=y\}} \exp\left(-H(x)\right)\, \mathcal{H}(dx) \\ &= -\frac{1}{N}\log \int_{\{Px=0\}} \exp\left(-H(x + NP^*y)\right)\, \mathcal{H}(dx).\end{aligned}$$

Applying the representation of $H(x + NP^*y)$ from above implies (14). □

From now on we will use the standard notation that for a probability measure ν

$$\mathrm{cov}_\nu(f, g) := \int \left(f - \int f d\nu\right)\left(g - \int g d\nu\right) \quad \text{and} \quad \mathrm{var}_\nu(f) := \mathrm{cov}_\nu(f, f).$$

The following representation of the Hessian of $\bar{H}$ follows from differentiation of (14).

Lemma 2.7. *Assume that the Hamiltonian H and the coarse-grained Hamiltonian $\bar{H}$ are given by* (2) *and* (9) *respectively. Recall that the conditional measures $\mu(dx|y)$ are defined by* (8). *For $1 \leq l, n \leq L$ we have*

$$\begin{aligned}\left(\mathrm{Hess}_Y\, \bar{H}(y)\right)_{ln} = \;& \delta_{ln} + \delta_{ln}\frac{1}{K}\int \sum_{i\in B(l)} \delta\psi_i''(x_i)\, \mu(dx|y) + \frac{1}{K}\sum_{i\in B(l),\, j\in B(n)} m_{ij} \\ & -\frac{1}{K}\mathrm{cov}_{\mu(dx|y)}\left(\sum_{j\in B(l)}\left(\sum_{i=1}^{N} m_{ij}x_i\right) + \delta\psi_j'(x_j)\ ,\ \sum_{j\in B(n)}\left(\sum_{i=1}^{N} m_{ij}x_i\right) + \delta\psi_j'(x_j)\right).\end{aligned} \tag{16}$$

Proof of Lemma 2.7. Let $\bar{\nabla}$ denote the gradient w.r.t. to the scalar product $\langle \cdot, \cdot \rangle_Y$ given by (6) and let ∇ denote the gradient w.r.t. the standard scalar product $\langle \cdot, \cdot \rangle$. Then we have $\bar{\nabla} = L\nabla$. Differentiation of (14) yields

$$\bar{\nabla}\bar{H}(y) = L\nabla\bar{H}(y) = (\mathrm{Id} + PMNP^*)y + Ps + \frac{1}{K}\int \nabla H_M(x,y)\mu_M(dx),$$

where μ_M denotes the probability measure associated to H_M, namely

$$\mu_M(dx) := Z^{-1}\delta\left(Px = 0\right)\exp(-H_M(x,y))\mathcal{H}(dx).$$

A second differentiation yields that for $1 \leq l, n \leq L$

$$\begin{aligned}
\left(\mathrm{Hess}_Y\, \bar{H}(y)\right)_{ln} &= L\frac{d^2}{dy_l dy_n}\bar{H}(y) \\
&= \delta_{ln} + (PMNP^*)_{ln} + \frac{1}{K}\int \frac{d^2}{dy_l dy_n}H_M(x,y)\mu_M(dx) \\
&\quad - \frac{1}{K}\operatorname{cov}_{\mu_M}\left(\frac{d}{dy_l}H_M\, ,\, \frac{d}{dy_n}H_M\right).
\end{aligned} \tag{17}$$

Let us consider each term successively. Note that

$$(PMNP^*)_{ln} = \frac{1}{K}\sum_{i\in B(l),\; j\in B(n)} m_{ij}.$$

For the next term we get by differentiation that

$$\frac{d}{dy_l}H_M(x,y) = \sum_{j\in B(l)}\left(\sum_{i=1}^{N} m_{ij}x_i\right) + \delta\psi_j'\left(x_j + (NP^*y)_j\right).$$

Another differentiation combined with the fact

$$j \in B(n) \quad \Rightarrow \quad (NP^*y)_j = y_n$$

yields that

$$\frac{d^2}{dy_n dy_l}H_M(x,y) = \delta_{ln}\sum_{j\in B(l)}\delta\psi_j''\left(x_j + (NP^*y)_j\right).$$

Because adding constants does not change covariances, we can write the fourth term as

$$\begin{aligned}
&\operatorname{cov}_{\mu_M}\left(\frac{d}{dy_l}H_M\, ,\, \frac{d}{dy_n}H_M\right) \\
&= \operatorname{cov}_{\mu_M}\left(\frac{d}{dy_l}H_M + \sum_{j\in B(l)}\sum_{i=1}^{N} m_{ij}(NP^*y)_i\, ,\, \frac{d}{dy_n}H_M + \sum_{j\in B(n)}\sum_{i=1}^{N} m_{ij}(NP^*y)_i\right).
\end{aligned}$$

Applying these identities to (17) yields

$$
\begin{aligned}
\left(\mathrm{Hess}_Y \bar{H}(y)\right)_{ln} = \ & \delta_{ln} + \frac{1}{K} \sum_{i\in B(l),\, j\in B(n)} m_{ij} + \frac{1}{K}\,\delta_{ln} \int \sum_{i\in B(l)} \delta\psi_i''\left((x+NP^*y)_i\right)\mu_M(dx) \\
& - \frac{1}{K} \operatorname{cov}_{\mu_M(dx)}\left(\sum_{j\in B(l)} \left(\sum_{i=1}^{N} m_{ij}(x+NP^*y)_i \right) + \delta\psi_j'\left((x+NP^*y)_j\right), \right. \\
& \left. \sum_{j\in B(n)} \left(\sum_{i=1}^{N} m_{ij}(x+NP^*y)_i \right) + \delta\psi_j'\left((x+NP^*y)_j\right) \right).
\end{aligned}
$$

This identity already yields (16) by the translation $x \mapsto x + NP^*y$. More precisely, we apply the fact that for any measurable function ξ

$$
\begin{aligned}
& \int \xi(x+NP^*y)\mu_M(dx) \\
& = \frac{\int_{\{Px=0\}} \xi(x+NP^*y)\exp\left(-H_M(x,y)\right)\mathcal{H}(dx)}{\int_{\{Px=0\}} \exp\left(-H_M(x,y)\right)\mathcal{H}(dx)} \\
& \overset{(15)}{=} \frac{\int_{\{P(x+NP^*y)=PNP^*y\}} \xi(x+NP^*y)\exp\left(-H(x+NP^*y)\right)\mathcal{H}(dx)}{\int_{\{P(x+NP^*y)=PNP^*y\}} \exp\left(-H(x+NP^*y)\right)\mathcal{H}(dx)} \\
& \overset{(7)}{=} \frac{\int_{\{Px=y\}} \xi(x)\exp\left(-H(x)\right)\mathcal{H}(dx)}{\int_{\{Px=y\}} \exp\left(-H(x)\right)\mathcal{H}(dx)} \\
& = \int \xi(x)\mu(dx|y),
\end{aligned}
$$

where we canceled in the second step the term

$$
\exp\left(-\frac{1}{2}N\left\langle y, (\mathrm{Id} + PMNP^*)y\right\rangle_Y - N\left\langle Ps, y\right\rangle_Y\right)
$$

with the denominator. □

2.3.2 Estimation of the off-diagonal elements of the Hessian of $\bar{H}$

In this section we will show, that the off-diagonal elements of the Hessian of $\bar{H}$ are controlled by ε. Explicitly, we will prove the following statement.

Lemma 2.8. *Assume that the Hamiltonian H and the coarse-grained Hamiltonian $\bar{H}$ are given by* (2) *and* (9) *respectively. Additionally, assume that the single-site potentials*

ψ_i *satisfy* (1) *with a constant* $c_1 < \infty$ *independent of the system size* $N \in \mathbb{N}$*, the mean spin* $m \in \mathbb{R}$*, and the boundary data* $s \in \mathbb{R}^N$*.*
If the interaction matrix M *satisfies CS(*ε*), then there is a constant* $0 \leq C < \infty$ *independent of* N*,* m*, and* s *(depending only on the block size* K *and* c_1*) such that*

$$\operatorname{Hess}_Y \bar{H}(y) - \operatorname{diag}\left(\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{11}, \ldots, \left(\operatorname{Hess}_Y \bar{H}(y)\right)_{LL}\right) \geq -C\varepsilon \operatorname{Id}.$$

This lemma is not obvious. Considering (16) one has to estimate for example the covariance

$$\operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \delta\psi_j'(x_j), \sum_{j \in B(n)} \delta\psi_j'(x_j)\right) \qquad \text{for } 1 \leq l \neq n \leq L.$$

It is not clear how to exploit the control CS(ε) on this term. The key observation is that the first function depends only on spins of the block $B(l)$, whereas the second function depends only on spins of block $B(n)$. One hopes that the covariance is decaying in the distance of the blocks, if ε is small enough. For this purpose we apply the covariance estimate [MO10b, Theorem 1], which connects the smallness condition CS(ε) to decay of covariances. For convenience, we state [MO10b, Theorem 1] as the next theorem.

Theorem 2.9 (Covariance estimate of [MO10b])**.** *Let* $d\mu := \frac{1}{Z}\exp(-H(x))dx$ *be a probability measure on a direct product of Euclidean spaces* $X = X_1 \times \cdots \times X_L$*. We assume that*

- *the conditional measures* $\mu(dx^l | x^n \in X_n \; n \neq l)$ *on* X_l*,* $l \in \{1, \ldots, L\}$*, satisfy a uniform SG(*$\tilde{\varrho}$*) (see Definition A.5).*
- *the numbers* κ_{ln} *satisfy*

  $$\kappa_{ln} := \max_x |\nabla_l \nabla_n H(x)| \leq C < \infty$$

  uniformly in $1 \leq l \neq n \leq L$*; here* $|\cdot|$ *denotes the operator norm of a bilinear form.*
- *the matrix* $A = (a_{ij})_{L \times L}$ *defined by*

  $$a_{ij} = \begin{cases} \varrho_i & \text{if } i = j, \\ -\kappa_{ij} & \text{else} \end{cases}$$

  is strictly positive definite.

Then for any function f *and* g

$$\operatorname{cov}_\mu(f, g) \leq \sum_{l,n=1}^{L} A_{ln}^{-1} \left(\int |\nabla_l f|^2 \, d\mu\right)^{\frac{1}{2}} \left(\int |\nabla_n g|^2 \, d\mu\right)^{\frac{1}{2}}.$$

Now, we can proceed to the proof of Lemma 2.8.

Proof of Lemma 2.8. Because of (16) we can write

$$\operatorname{Hess}_Y \bar{H}(y) - \operatorname{diag}\left(\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{11}, \ldots, \left(\operatorname{Hess}_Y \bar{H}(y)\right)_{LL}\right) = W_1 + W_2,$$

where the matrix W_1 is given by

$$(W_1)_{ln} = \begin{cases} \frac{1}{K} \sum_{i \in B(l), j \in B(n)} m_{ij}, & \text{if } 1 \leq n \neq l \leq L, \\ 0, & \text{if } l = n, \end{cases}$$

and the elements of the matrix W_2 are defined for $1 \leq n \neq l \leq L$ by

$$(W_2)_{ln} = -\frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \left(\sum_{i=1}^{N} m_{ij} x_i \right) + \delta\psi_j'(x_j), \sum_{j \in B(n)} \left(\sum_{i=1}^{N} m_{ij} x_i \right) + \delta\psi_j'(x_j) \right)$$

and for $l = n$ by

$$(W_2)_{ln} = 0.$$

Firstly, we will estimate W_1. Using CS(ε) we have for any vector $x \in \mathbb{R}^L$

$$\begin{aligned} \sum_{l,n=1}^{L} x_l (W_1)_{ln} x_n &\geq -\frac{1}{K} \sum_{l,n=1}^{L} |x_l|\,|x_n| \sum_{i \in B(l), j \in B(n)} |m_{ij}| \\ &\geq -\varepsilon \sum_{l=1}^{L} x_l^2. \end{aligned}$$

The last estimate shows that

$$W_1 \geq -\varepsilon \ \operatorname{Id}$$

in the sense of quadratic forms. The estimation of W_2 is a little bit more subtle. By bilinearity of the covariance the matrix W_2 can be rewritten as

$$W_2 = W_3 + W_4 + W_5 + W_6,$$

where the elements of the matrices $W_1, \ldots, W_6$ are defined for $1 \leq l \leq L$ by

$$(W_3)_{ll} = 0 \qquad (W_4)_{ll} = 0 \qquad (W_5)_{ll} = 0 \qquad (W_6)_{ll} = 0$$

and for $1 \leq l \neq n \leq L$ by

$$(W_3)_{ln} = -\frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \left(\sum_{i=1}^{N} m_{ij} x_i\right), \sum_{j \in B(n)} \left(\sum_{i=1}^{N} m_{ij} x_i\right)\right),$$

$$(W_4)_{ln} = -\frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \delta\psi_j'(x_j), \sum_{j \in B(n)} \delta\psi_j'(x_j)\right),$$

$$(W_5)_{ln} = -\frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \left(\sum_{i=1}^{N} m_{ij} x_i\right), \sum_{j \in B(n)} \delta\psi_j'(x_j)\right),$$

$$(W_6)_{ln} = -\frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \delta\psi_j'(x_j), \sum_{j \in B(n)} \left(\sum_{i=1}^{N} m_{ij} x_i\right)\right).$$

The main tool for the estimation of $W_3, \ldots, W_6$ is the covariance estimate of Theorem 2.9. We estimate each matrix separately and start with W_3. Because the LSI(ϱ) implies the SG(ϱ) (cf. Lemma A.6), the hypotheses of the criterion of Otto & Reznikoff (cf. Theorem A.4) are stronger than the hypotheses of Theorem 2.9. In the proof of Proposition 2.3 we applied Theorem A.4 to the measure $\mu(dx|y)$. Hence we can apply Theorem 2.9 to the measure $\mu(dx|y)$ and get for $1 \leq l \neq n \leq L$ the estimate

$$\begin{aligned}
-(W_3)_{ln} &= \frac{1}{K} \operatorname{cov}_{\mu(dx|y)}\left(\sum_{j \in B(l)} \left(\sum_{i=1}^{N} m_{ij} x_i\right), \sum_{j \in B(n)} \left(\sum_{i=1}^{N} m_{ij} x_i\right)\right) \\
&\leq \sum_{s_1, s_2 = 1}^{L} \left(A^{-1}\right)_{s_1 s_2} \left(\sum_{i \in B(l), j \in B(s_1)} m_{ij}^2\right)^{\frac{1}{2}} \left(\sum_{i \in B(n), j \in B(s_2)} m_{ij}^2\right)^{\frac{1}{2}},
\end{aligned}$$

where the matrix A is defined as in (11). By equivalence of norms in finite dimensional vector spaces we get from the last estimate that

$$-(W_3)_{ln} \leq C \sum_{s_1, s_2 = 1}^{L} \|M_{l s_1}\| \left(A^{-1}\right)_{s_1 s_2} \|M_{s_2 n}\|.$$

Here and later on in this proof $0 < C < \infty$ denotes a generic constant depending only on K and c_1. Before we can continue the estimation of $W_3, \ldots, W_6$ we need some basic estimates for A. Using the Neumann representation of A^{-1} one sees that

$$\operatorname{diag}\left(\left(A^{-1}\right)_{11}, \ldots, \left(A^{-1}\right)_{LL}\right) \geq \frac{1}{\tilde{\varrho}} \operatorname{Id}, \tag{18}$$

in the sense of quadratic forms. Because for sufficiently small ε (cf. (13))

$$A \geq \tilde{\varrho}\,\mathrm{Id} - (\|M_{ln}\|)_{L\times L} > 0,$$

it follows that

$$A^{-1} \leq \left(\tilde{\varrho}\,\mathrm{Id} - (\|M_{ln}\|)_{L\times L}\right)^{-1} = \frac{1}{\tilde{\varrho}} \sum_{k=0}^{\infty} \left(\frac{(\|M_{ln}\|)_{L\times L}}{\tilde{\varrho}} \right)^{k}. \tag{19}$$

A combination of (18), (19) and (12) yields

$$\begin{aligned} A^{-1} - \mathrm{diag}\left(\left(A^{-1}\right)_{11}, \ldots, \left(A^{-1}\right)_{LL}\right) &\leq \frac{1}{\tilde{\varrho}} \sum_{k=1}^{\infty} \left(\frac{(\|M_{ln}\|)_{L\times L}}{\tilde{\varrho}} \right)^{k} \\ &\overset{\mathrm{CS}(\varepsilon)}{\leq} \frac{1}{\tilde{\varrho}} \frac{\varepsilon}{\tilde{\varrho} - \varepsilon}\ \mathrm{Id} \end{aligned} \tag{20}$$

and

$$(\|M_{s_1 s_2}\|)_{L\times L}\, A^{-1}\, (\|M_{s_1 s_2}\|)_{L\times L} \leq \frac{1}{\tilde{\varrho}} \sum_{k=2}^{\infty} \left(\frac{(\|M_{ln}\|)_{L\times L}}{\tilde{\varrho}} \right)^{k} \leq \frac{1}{\tilde{\varrho}} \frac{\varepsilon^2}{\tilde{\varrho} - \varepsilon}\ \mathrm{Id}\,. \tag{21}$$

Now, we turn back to the estimation of W_3. It follows from (21) that

$$-W_3 \leq (\|M_{s_1 s_2}\|)_{L\times L}\, A^{-1}\, (\|M_{s_1 s_2}\|)_{L\times L} \leq C\varepsilon,$$

for a generic constant $C < \infty$ that depends only on K and c_1.
Let us turn to the estimation of W_4. By an application of Theorem 2.9 we have for $1 \leq l \neq n \leq L$

$$\begin{aligned} -(W_4)_{ln} &\leq \left(A^{-1}\right)_{ln} \max_{i\in\{1,\ldots,N\}} \max_{x\in\mathbb{R}} |\delta\psi_i''(x)|^2 \\ &\leq C\left(A^{-1}\right)_{ln}. \end{aligned}$$

Hence (20) yields

$$-W_4 \leq A^{-1} - \mathrm{diag}\left(\left(A^{-1}\right)_{11}, \ldots, \left(A^{-1}\right)_{LL}\right) \leq C\varepsilon.$$

With an argument of the same type one can also estimate the matrices W_5 and W_6 as

$$-W_5 - W_6 \leq C\varepsilon,$$

which together with the estimates of W_3 and W_4 yields

$$-W_2 \leq C\varepsilon.$$

□

2.3.3 Estimation of the diagonal elements of the Hessian of $\bar{H}$

In this section we will deduce the strict positivity of the diagonal elements of the Hessian of $\bar{H}$ for sufficiently large block sizes K and sufficiently small interaction ε. More precisely, we will show the following statement.

Lemma 2.10. *Assume that the Hamiltonian H and the coarse-grained Hamiltonian $\bar{H}$ are given by* (2) *and* (9) *respectively. Additionally, assume that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. Then there exist $K_0 \in \mathbb{N}$ depending only on c_1 such that:*
If the block size $K \geq K_0$ and the interaction matrix M satisfies CS(ε), then there are constants $\lambda > 0$ and $C < \infty$ independent of N, m, and s (depending only on K and c_1) such that for all $1 \leq l \leq L$ and $y \in Y$

$$\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{ll} \geq \lambda - C\varepsilon.$$

Therefore it holds

$$\operatorname{diag}\left(\left(\operatorname{Hess}_Y \bar{H}(y)\right)_{11}, \ldots, \left(\operatorname{Hess}_Y \bar{H}(y)\right)_{LL}\right) \geq (\lambda - C\varepsilon)\operatorname{Id}$$

in the sense of quadratic forms.

For the proof of Lemma 2.10 we use a conditioning technique, which allows to apply a perturbation argument for small ε independent of N, m, and s. Let us consider an arbitrary but fixed block $B(l)$, $1 \leq l \leq L$. Recall that for a given configuration $x \in \mathbb{R}^N$ the spins inside the block $B(l)$ are denoted by $x^l := (x_i)_{i \in B(l)}$ and the spins outside the block $B(l)$ are denoted by $\bar{x}^l := (x_i)_{i \notin B(l)}$ (cf. Figure 3). As in the proof of Proposition 2.3 disintegration of the microscopic measure $\mu(dx|y)$ with respect to x^l yields

$$\mu(dx|y) = \mu(dx^l|\bar{x}^l, y)\, \bar{\mu}(d\bar{x}^l|y),$$

where $\mu(dx^l|\bar{x}^l, y)$ and $\bar{\mu}(d\bar{x}^l|y)$ denotes the conditional measure and the corresponding marginal respectively (cf. (10)). In the proof of Proposition 2.3 we have shown that the conditional measures $\mu(dx^l|\bar{x}^l, y)$ are given by

$$\mu(dx^l|\bar{x}^l, y) = \frac{1}{Z}\, 1_{\left\{\frac{1}{K}\sum_{i \in B(l)} x_i = y_l\right\}} \exp\left(-H(x^l|M, s_c)\right) \mathcal{H}(dx), \tag{22}$$

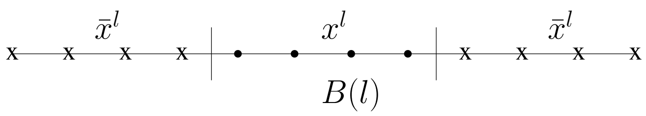

Figure 3: Conditioning on spins outside of the block $B(l)$

where the s_c and $H(x^l|M, s_c)$ are defined by

$$s_{c,i} := s_i + \sum_{j \notin B(l)} m_{ij} x_j \qquad \text{for } i \in B(l) \tag{23}$$

and by

$$H(x^l|M, s^*) := \sum_{i \in B(l)} \psi_i(x_i) + \frac{1}{2} \sum_{i,j \in B(l)} m_{ij} x_i x_j + \sum_{i \in B(l)} s_i^* x_i \tag{24}$$

for an arbitrary vector $s^* \in \mathbb{R}^{B(l)}$. We introduce the coarse-grained Hamiltonian of $H(x^l|M, s^*)$ as usual i.e. for $y_l \in \mathbb{R}$

$$\bar{H}(y_l|M, s^*) := -\frac{1}{K} \log \int_{\left\{\frac{1}{K}\sum_{i \in B(l)} x_i = y_l\right\}} \exp\left(-H(x^l|M, s^*)\right) \mathcal{H}(dx^l). \tag{25}$$

The next lemma shows that uniform positivity of

$$\frac{d^2}{dy_l^2} \bar{H}(y_l|M, s^*)$$

yields uniform positivity of $(\mathrm{Hess}_Y \, \bar{H}(y))_{ll}$ for small ε . This observation is one of the main insights in order to apply a perturbation argument for small ε independent of the system size N. The advantage of $\bar{H}(y_l|M, s^*)$ over $\bar{H}(y)$ is that in (25) one integrates only over sites of the block $B(l)$, whereas in the definition (9) of the coarse-grained Hamiltonian $\bar{H}(y)$ one integrates over all sites of the spin system.

Lemma 2.11. *Assume that the Hamiltonian H and the coarse-grained Hamiltonian $\bar{H}$ are given by* (2) *and* (9) *respectively. Additionally, assume that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1 < \infty$ independent of the system size $N \in \mathbb{N}$, the mean spin $m \in \mathbb{R}$, and the boundary data $s \in \mathbb{R}^N$. The vector s_c and the Hamiltonian $H(x^l|M, s_c)$ are given by* (23) *and* (24) *respectively. Then:*
If the interaction matrix M satisfies CS(ε), then for all $1 \leq l \leq L$ and $y \in Y$

$$(\mathrm{Hess}_Y \, \bar{H}(y))_{ll} \geq \int \frac{d^2}{dy_l^2} \bar{H}(y_l|M, s_c) \bar{\mu}(d\bar{x}^l|y) - C\varepsilon,$$

where the constant $C < \infty$ is independent of N, m, and s (depending only on the block size K and c_1).

The proof of Lemma 2.11 consists of two steps. In the first step we show that the disintegration (10) yields the identity

$$\begin{aligned} \left(\mathrm{Hess}_Y \, \bar{H}(y)\right)_{ll} &= \int \frac{d^2}{dy_l^2} \bar{H}(y_l|M, s_c) \; \bar{\mu}(d\bar{x}^l|y) \\ &\quad - \frac{1}{K} \mathrm{var}_{\bar{\mu}(d\bar{x}^l|y)} \left(\int \sum_{j \in B(l)} \left(\sum_{i=1}^{N} m_{ij} x_i \right) + \delta\psi_j'(x_j) \; \mu(dx^l|\bar{x}^l, y) \right). \end{aligned} \tag{26}$$

In the second step we show that the variance term on the right hand side can be estimated by using the covariance estimate of Theorem 2.9 as

$$\frac{1}{K}\operatorname{var}_{\bar\mu(d\bar x^l|y)}\left(\int \sum_{j\in B(l)}\left(\sum_{i=1}^{N} m_{ij}x_i\right)+\delta\psi_j'(x_j)\ \mu(dx^l|\bar x^l,y)\right)\leq C\varepsilon. \tag{27}$$

We will state the full proof of Lemma 2.11 in the next subsection. The next lemma provides the last remaining ingredient of the proof of Lemma 2.10 which is the uniform positivity of $\frac{d^2}{dy_l^2}\bar H(y_l|M,s^*)$.

Lemma 2.12. *Assume that the single-site potentials ψ_i satisfy* (1) *with a constant $c_1<\infty$ independent of the system size $N\in\mathbb{N}$, the mean spin $m\in\mathbb{R}$, and the boundary data $s\in\mathbb{R}^N$. Then there is $K_0\in\mathbb{N}$ such that:*
If the block size $K\geq K_0$ and the interaction matrix M satisfies CS(ε), then there are constants $\lambda>0$ and $C<\infty$ independent of N, m, and s (depending only on K and c_1) such that for all $1\leq l\leq L$, $y_l\in\mathbb{R}$, and $s^*\in\mathbb{R}^{B(l)}$

$$\frac{d^2}{dy_l^2}\bar H(y_l|M,s^*)\geq\lambda-C\varepsilon. \tag{28}$$

Let us explain the main idea of the proof of Lemma 2.12. If the block size K is large enough, the generalized local Cramér theorem (cf. Corollary 3.2 and Remark 3.3) yields

$$\frac{d^2}{dy_l^2}\bar H(y_l|0,\tilde s)\geq\lambda>0 \tag{29}$$

for all $y_l\in\mathbb{R}$ and $\tilde s\in\mathbb{R}^{B(l)}$. The strategy is to derive (28) from (29) by a perturbation argument. More precisely, we will show that for a clever choice of $\tilde s=\tilde s(s^*)\in\mathbb{R}^{B(l)}$

$$\left|\frac{d^2}{dy_l^2}\bar H(y_l|M,s^*)-\frac{d^2}{dy_l^2}\bar H(y_l|0,\tilde s)\right|\leq C\varepsilon, \tag{30}$$

for all y_l and s^*. The constant $C<\infty$ just depends on K and c_1. For the proof of Lemma 2.10 it is crucial that the last inequality holds uniformly in $s^*\in\mathbb{R}^{B(l)}$. Because we consider unbounded spins with quadratic interaction, the latter is a difficult task and leads to the specific choice of $\tilde s=\tilde s(s^*)$ given by (41). It would be a lot easier to derive (30) for bounded spin-values with finite-range interaction. In this case one could also deduce the estimate (30) for the choice $\tilde s=0$. Then the standard version of the local Cramér theorem [GOVW09, Proposition 31] would be sufficient for the perturbation argument at least for homogeneous single-site potentials $\psi_i=\psi$. The reason is that [GOVW09, Proposition 31] already yields

$$\frac{d^2}{dy_l^2}\bar H(y_l|0,0)\geq\lambda>0$$

in this case. We will state the full proof of Lemma 2.12 in the next subsection.

Proof of Lemma 2.10. The desired statement follows directly from a combination of Lemma 2.11 and Lemma 2.12. □

Proof of auxiliary results

Proof of Lemma 2.11. Let us deduce the identity (26). Recall that by Lemma 2.7 we have

$$\begin{aligned}\left(\mathrm{Hess}_Y \bar{H}(y)\right)_{ll} &= 1 + \frac{1}{K}\sum_{i,j\in B(l)} m_{ij} + \frac{1}{K}\int \sum_{j\in B(l)} \delta\psi_j(x_j)''\mu(dx|y) \\ &\quad - \frac{1}{K}\operatorname{var}_{\mu(dx|y)}\left(\sum_{j\in B(l)}\left(\sum_{i=1}^N m_{ij}x_i\right) + \delta\psi_j'(x_j)\right).\end{aligned}$$

Applying the disintegration rule (10) to variances yields that for an arbitrary function $f(x) = f(x^l, \bar{x}^l)$

$$\begin{aligned}&\operatorname{var}_{\mu(dx|y)}(f(x)) \\ &\quad = \int \operatorname{var}_{\mu(dx^l|\bar{x}^l,y)}(f(x^l,\bar{x}^l))\, \bar{\mu}(d\bar{x}^l|y) + \operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)}\left(\int f(x^l,\bar{x}^l)\mu(dx^l|\bar{x}^l,y)\right).\end{aligned}$$

An application of the last identity to

$$f(x) = \sum_{j\in B(l)}\left(\sum_{i=1}^N m_{ij}x_i\right) + \delta\psi_j'(x_j)$$

and the disintegration rule (10) yield the identity

$$\begin{aligned}\left(\mathrm{Hess}_Y \bar{H}(y)\right)_{ll} = \int \Bigg[\int &\left(1 + \frac{1}{K}\sum_{i,j\in B(l)} m_{ij} + \frac{1}{K}\int\sum_{j\in B(l)}\delta\psi_j(x_j)''\right)\mu(dx^l|\bar{x}^l,y) \\ &- \frac{1}{K}\operatorname{var}_{\mu(dx^l|\bar{x}^l,y)}\left(\sum_{j\in B(l)}\left(\sum_{i=1}^N m_{ij}x_i\right) + \delta\psi_j'(x_j)\right)\Bigg]\bar{\mu}(d\bar{x}^l|y) \\ &- \frac{1}{K}\operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)}\left(\int\sum_{j\in B(l)}\left(\sum_{i=1}^N m_{ij}x_i\right) + \delta\psi_j'(x_j)\ \mu(dx^l|\bar{x}^l,y)\right).\end{aligned}$$

Note that the Hamiltonian $H(x^l|M, s^*)$ defined by (24) has the same structure as the Hamiltonian $H(x)$ given by (2). Therefore an application of Lemma 2.7 yields that

$$\begin{aligned}\frac{d^2}{dy_l^2}\bar{H}(y_l|M, s_c) &= 1 + \frac{1}{K}\sum_{i,j\in B(l)} m_{ij} + \frac{1}{K}\int \sum_{j\in B(l)} \delta\psi_j''(x_j)\mu(dx^l|\bar{x}^l, y) \\ &\quad - \frac{1}{K}\operatorname{var}_{\mu(dx^l|\bar{x}^l,y)}\left(\sum_{j\in B(l)}\left(\sum_{i\in B(l)} m_{ij}x_i\right) + \delta\psi_j'(x_j)\right) \\ &= 1 + \frac{1}{K}\sum_{i,j\in B(l)} m_{ij} + \frac{1}{K}\int \sum_{j\in B(l)} \delta\psi_j''(x_j)\mu(dx^l|\bar{x}^l, y) \\ &\quad - \frac{1}{K}\operatorname{var}_{\mu(dx^l|\bar{x}^l,y)}\left(\sum_{j\in B(l)}\left(\sum_{i=1}^{N} m_{ij}x_i\right) + \delta\psi_j'(x_j)\right), \qquad (31)\end{aligned}$$

where we applied in the second step the fact that adding constants does not change the variance. Finally, the desired identity (26) follows from the last two equations.
It remains to derive the estimate (27) of the variance term of the right hand side of (26). By Young's inequality

$$\begin{aligned}&\frac{1}{K}\operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)}\left(\int \sum_{j\in B(l)}\left(\sum_{i=1}^{N} m_{ij}x_i\right) + \delta\psi_j'(x_j)\ \mu(dx^l|\bar{x}^l, y)\right) \\ &\leq \frac{2}{K}\operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)}\left(\int \sum_{j\in B(l)}\sum_{i=1}^{N} m_{ij}x_i\ \mu(dx^l|\bar{x}^l, y)\right) \\ &\quad + \frac{2}{K}\operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)}\left(\int \sum_{j\in B(l)} \delta\psi_j'(x_j)\ \mu(dx^l|\bar{x}^l, y)\right). \qquad (32)\end{aligned}$$

Let us consider the first term of the right hand side of (32). By the disintegration rule (10) we have for any function $\xi(\bar{x}^l)$

$$\begin{aligned}\int \xi(\bar{x}^l)\bar{\mu}(d\bar{x}^l|y) &= \int \xi(\bar{x}^l)\underbrace{\int 1\ \mu(dx^l|\bar{x}^l, y)}_{=1}\ \bar{\mu}(d\bar{x}^l|y) \\ &= \int \xi(\bar{x}^l)\ \mu(dx|y).\end{aligned}$$

It follows that

$$\frac{2}{K}\operatorname{var}_{\bar\mu(d\bar x^l|y)}\left(\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right)$$
$$=\frac{2}{K}\operatorname{var}_{\mu(dx|y)}\left(\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right).$$

Therefore an application of the covariance estimate of Theorem 2.9 (cf. proof of Lemma 2.8) yields

$$\frac{2}{K}\operatorname{var}_{\bar\mu(d\bar x^l|y)}\left(\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right)$$
$$\leq\frac{2}{\varrho K}\sum_{s_1,s_2=1}^L\left(A^{-1}\right)_{s_1s_2}\left(\int\sum_{k\in B(s_1)}\left|\frac{d}{dx_k}\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right|^2\bar\mu(d\bar x^l|y)\right)^{\frac12}$$
$$\times\left(\int\sum_{k\in B(s_2)}\left|\frac{d}{dx_k}\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right|^2\bar\mu(d\bar x^l|y)\right)^{\frac12}. \tag{33}$$

It follows from the definition $x^l=(x_k)_{k\in B(l)}$ that for $k\in B(l)$

$$\frac{d}{dx_k}\left(\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right)=0. \tag{34}$$

Using the definition (24) of $H(x^l|M,s_c)$ we see by direct calculation that for $k\notin B(l)$

$$\left|\frac{d}{dx_k}\int\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ \mu(dx^l|\bar x^l,y)\right|$$
$$=\left|\sum_{j\in B(l)}m_{kj}-\operatorname{cov}_{\mu(dx^l|\bar x^l,y)}\left(\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\quad\frac{d}{dx_k}H(x^l|M,s_c)\right)\right|$$
$$\leq\sum_{j\in B(l)}|m_{kj}|+\left|\operatorname{cov}_{\mu(dx^l|\bar x^l,y)}\left(\sum_{j\in B(l)}\sum_{i=1}^N m_{ij}x_i\ ,\ \sum_{s\in B(l)}m_{ks}x_s\right)\right|.$$

From now on, let $C < \infty$ denote a generic constant depending only on K and c_1. Because $\mu(dx^l|\bar{x}^l, y)$ satisfies LSI($\tilde{\varrho}$) with $\tilde{\varrho} > 0$ depending only on K and c_1 (cf. proof of Proposition 2.3), an application of Lemma A.6 yields

$$\begin{aligned} &\left| \frac{d}{dx_k} \int \sum_{j \in B(l)} \sum_{i=1}^{N} m_{ij} x_i \, \mu(dx^l|\bar{x}^l, y) \right| \\ &\quad \leq \; C \left(\sum_{j \in B(l)} m_{kj}^2 \right)^{\frac{1}{2}} + \frac{1}{\tilde{\varrho}} \underbrace{\left(\sum_{i,j \in B(l)} m_{ij}^2 \right)^{\frac{1}{2}}}_{\leq C \|M_{ll}\|} \left(\sum_{j \in B(l)} m_{kj}^2 \right)^{\frac{1}{2}} \\ &\quad \overset{\text{CS}(\varepsilon)}{\leq} \; \left(C + \frac{C}{\tilde{\varrho}} \, \varepsilon \right) \left(\sum_{j \in B(l)} m_{kj}^2 \right)^{\frac{1}{2}} . \end{aligned} \tag{35}$$

A combination of the estimates (33), (34) and (35) yields the estimate of the first term on the right hand side of (32). More precisely,

$$\begin{aligned} &\frac{2}{K} \operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)} \left(\int \sum_{j \in B(l)} \sum_{i=1}^{N} m_{ij} x_i \, \mu(dx^l|\bar{x}^l, y) \right) \\ &\quad \leq C \sum_{s_1, s_2 = 1}^{L} \left(A^{-1} \right)_{s_1 s_2} \left(\sum_{i \in B(s_1), \, j \in B(l)} m_{ij}^2 \right)^{\frac{1}{2}} \left(\sum_{i \in B(s_2), \, j \in B(l)} m_{ij}^2 \right)^{\frac{1}{2}} \\ &\quad \leq C \sum_{s_1, s_2 = 1}^{L} \left(A^{-1} \right)_{s_1 s_2} \; \|M_{l s_1}\| \, \|M_{s_2 l}\| \overset{(21)}{\leq} C\varepsilon. \end{aligned}$$

The second term on the right hand side of (32), namely

$$\frac{2}{K} \operatorname{var}_{\bar{\mu}(d\bar{x}^l|y)} \left(\int \sum_{j \in B(l)} \delta\psi_j'(x_j) \, \mu(dx^l|\bar{x}^l, y) \right),$$

can be estimated with the same argument as we used for the first term. The only different ingredient is the estimation of

$$\begin{aligned} \left| \frac{d}{dx_k} \int \sum_{j \in B(l)} \delta\psi_j'(x_j) \, \mu(dx^l|\bar{x}^l, y) \right| &= \left| \operatorname{cov}_{\mu(dx^l|\bar{x}^l, y)} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j) \, , \; \sum_{s \in B(l)} m_{ks} x_s \right) \right| \\ &\leq \frac{C}{\tilde{\varrho}} \left(\sum_{j \in B(l)} m_{kj}^2 \right)^{\frac{1}{2}} , \end{aligned}$$

where we applied the estimate of Lemma A.6 and the uniform bound (1) of the functions $\delta\psi_i$. □

Proof of Lemma 2.12. Because the estimate (29) follows directly from the generalized local Cramér theorem (cf. Corollary 3.2 and Remark 3.3), it is only left to deduce the estimate (30). Let $\nu(dx^l|M,s^*)$ denote the Gibbs measure on

$$\left\{ x^l \in \mathbb{R}^{B(l)} \;\middle|\; \frac{1}{K}\sum_{i\in B(l)} x_i = y_l \right\}$$

associated to the Hamiltonian $H(dx^l|M,s^*)$. More precisely,

$$\nu(dx^l|M,s^*) = \frac{1}{Z}\, 1_{\left\{\frac{1}{K}\sum_{i\in B(l)} x_i = y_l\right\}} \exp(-H(dx^l|M,s^*)\mathcal{H}(dx^l).$$

The same reason as for (31) yields that

$$\begin{aligned}
\frac{d^2}{dy_l^2}\bar{H}(y_l|M,s^*) &= 1 + \frac{1}{K}\sum_{i\in B(l),\, j\in B(l)} m_{ij} + \int \frac{1}{K}\sum_{j\in B(l)} \delta\psi_j''(x_j)\, \nu(dx^l|M,s^*) \\
&\quad - \frac{1}{K}\operatorname{var}_{\nu(dx^l|B,s^*)}\left(\sum_{j\in B(l)}\left(\sum_{i\in B(l)} m_{ij}x_i\right) + \delta\psi_j'(x_j)\right) \\
&= 1 + \frac{1}{K}\sum_{i\in B(l),\, j\in B(l)} m_{ij} + \int \frac{1}{K}\sum_{j\in B(l)} \delta\psi_j''(x_j)\, \nu(dx^l|M,s^*) \\
&\quad - \frac{1}{K}\operatorname{var}_{\nu(dx^l|B,s^*)}\left(\sum_{i,j\in B(l)} m_{ij}x_i\right) - \frac{1}{K}\operatorname{var}_{\nu(dx^l|B,s^*)}\left(\sum_{j\in B(l)} \delta\psi_j'(x_j)\right) \\
&\quad - \frac{2}{K}\operatorname{cov}_{\nu(dx^l|B,s^*)}\left(\sum_{i,j\in B(l)} m_{ij}x_i\,,\ \sum_{j\in B(l)} \delta\psi_j'(x_j)\right).
\end{aligned}$$

An application of this formula to $\bar{H}(y_l|0,\tilde{s})$ with arbitrary $\tilde{s}\in\mathbb{R}^{B(l)}$ yields

$$\frac{d^2}{dy_l^2}\bar{H}(y_l|0,\tilde{s}) = 1 + \int \frac{1}{K}\sum_{j\in B(l)} \delta\psi_j''(x_j)\, \nu(dx^l|0,\tilde{s}) - \frac{1}{K}\operatorname{var}_{\nu(dx^l|0,\tilde{s})}\left(\sum_{j\in B(l)} \delta\psi_j'(x_j)\right).$$

It follows from the last two equations that

$$
\begin{aligned}
&\left| \frac{d^2}{dy_l^2} \bar{H}(y_l|M,s) - \frac{d^2}{dy_l^2} \bar{H}(y_l|0,\tilde{s}) \right| \leq \frac{1}{K} \left| \sum_{i,j \in B(l)} m_{ij} \right| \\
&\quad + \frac{1}{K} \left| \operatorname{var}_{\nu(dx^l|M,s^*)} \left(\sum_{i,j \in B(l)} m_{ij} x_i \right) \right| \\
&\quad + \frac{2}{K} \left| \operatorname{cov}_{\nu(dx^l|M,s^*)} \left(\sum_{i,j \in B(l)} m_{ij} x_i \ , \ \delta\psi_j'(x_j) \right) \right| \\
&\quad + \frac{1}{K} \left| \int \sum_{j \in B(l)} \delta\psi_j''(x_j) \ \nu(dx^l|M,s^*) - \int \sum_{j \in B(l)} \delta\psi_j''(x_j) \ \nu(dx^l|0,\tilde{s}) \right| \\
&\quad + \frac{1}{K} \left| \operatorname{var}_{\nu(dx^l|M,s^*)} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j) \right) - \operatorname{var}_{\nu(dx^l|0,\tilde{s})} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j) \right) \right| \\
&=: T_1 + T_2 + T_3 + T_4 + T_5.
\end{aligned}
$$

By the same argument as in the proof of Proposition 2.3 it follows that the measure $\nu(dx^l|M,s^*)$ satisfies LSI($\tilde{\varrho}$) with $\tilde{\varrho} > 0$ depending only on K and c_1. Therefore by using the CS(ε) and the basic covariance estimate of Lemma A.6 it is easy to deduce

$$
T_1 + T_2 + T_3 \leq C\varepsilon
$$

for a constant $C < \infty$ depending only on K and c_1.

The interesting part is the estimation of T_4 and T_5, for which the right choice of $\tilde{s} = \tilde{s}(s^*) \in \mathbb{R}^{B(l)}$ plays an important role. Therefore let us motivate how to choose $\tilde{s} = \tilde{s}(s^*)$ for a given vector $s^* \in \mathbb{R}^{B(l)}$. Let ξ be an arbitrary bounded function (think of $\xi = \delta\psi_i$). Then the structure of T_4 and T_5 is given by

$$
\left| \int \xi(x^l) \ \nu(dx^l|M,s^*) - \int \xi(x^l) \ \nu(dx^l|0,\tilde{s}) \right|.
$$

We want to estimate this term uniformly in the unbounded parameters $y_l \in \mathbb{R}$ and $s^* \in \mathbb{R}^{B(l)}$. Therefore let us take a closer look at the dependence of

$$
\int \xi(x^l) \ \nu(dx^l|M,s^*) = \int_{\left\{ \frac{1}{K} \sum_{i \in B(l)} x_i = y_l \right\}} \xi(x^l) \frac{1}{Z} \exp\left(-H(x^l|M,s^*) \right) \mathcal{H}(dx^l) \quad (36)
$$

on the parameters y_l and s^*. On the block $B(l)$ we define the coarse-graining operator $P_l : \mathbb{R}^{B(l)} \to \mathbb{R}$ by

$$P_l x^l = \frac{1}{K} \sum_{i \in B(l)} x_i.$$

As for the original coarse-graining operator P we have the identity

$$P_l K P_l^t = \mathrm{Id}_{\mathbb{R}^{B(l)}},$$

where P_l^t denotes the transpose of P. Using the last identity we see that the orthogonal projection Π of $\mathbb{R}^{B(l)}$ on $\ker P_l = \{P_l x^l = 0\}$ is given by

$$\Pi = \mathrm{Id} - K P_l P_l^t. \tag{37}$$

The dependence of the integration space on y_l is abolished by the translation $x \mapsto x - K P_l^t y_l$, which maps

$$\left\{ \frac{1}{K} \sum_{i \in B(l)} x_i = y_l \right\} = \left\{ P_l x^l = y_l \right\} \to \left\{ P_l x^l = 0 \right\} = \left\{ \frac{1}{K} \sum_{i \in B(l)} x_i = 0 \right\}.$$

Using the definition (24) of $H(x^l|M, s^*)$ we get by direct calculation that

$$\begin{aligned} &H(x^l + K P_l^t y_l | M, s^*) \\ &\quad = \frac{1}{2} \left\langle x^l, (\mathrm{Id} + M_{ll}) x^l \right\rangle + \left\langle s^* + (\mathrm{Id} + M_{ll}) K P_l^t y_l, x^l \right\rangle + \sum_{i \in B(l)} \delta\psi_i(x_i + y_l) \\ &\qquad + \frac{1}{2} \left\langle K P_l^t y_l, (\mathrm{Id} + M_{ll}) K P_l^t y_l \right\rangle + \left\langle s^*, K P_l^t y_l \right\rangle. \end{aligned}$$

Because we can cancel all terms that are independent of x^l with terms of the normalization constant Z, the translation $x \mapsto x - K P_l^t y_l$ applied to the right hand side of (36) yields the identity

$$\begin{aligned} &\int \xi(x^l) \nu(dx^l | M, s^*) = \frac{1}{Z} \int_{\left\{P_l x^l = 0\right\}} \xi(x^l + K P_l^t y_l) \times \\ &\exp\left(-\frac{1}{2} \left\langle x^l, (\mathrm{Id} + M_{ll}) x^l \right\rangle - \left\langle s^* + M_{ll} K P_l^t y_l, x^l \right\rangle - \sum_{i \in B(l)} \delta\psi_i(x_i + y_l) \right) \mathcal{H}(dx^l), \end{aligned} \tag{38}$$

where we used the fact that $\left\langle K P_l^* y_l, x^l \right\rangle = 0$ for $x^l \in \left\{ P_l x^l = 0 \right\}$. Note that in (38) only the linear term $\left\langle s^* + M_{ll} K P_l^t y_l, x^l \right\rangle$ depends on the parameters y_l and s^*. The idea is to get rid of this term by a second translation $x^l \mapsto x^l + v$, which leaves the integration space $\{P_l x^l = 0\}$ invariant. The key observation is that for any $x^l \in \left\{P_l x^l = 0\right\}$ and $z \in \mathbb{R}^{B(l)}$

$$\left\langle z, x^l \right\rangle = \left\langle \Pi z, x^l \right\rangle,$$

where the orthogonal projection $\Pi : \mathbb{R}^{B(l)} \to \{P_l x^l = 0\}$ is defined by (37). Hence we can rewrite the Gaussian part of the Hamiltonian in (38) as

$$\begin{aligned}\frac{1}{2}\left\langle x^l, (\mathrm{Id} + M_{ll}) x^l \right\rangle + \left\langle s^* + M_{ll} K P_l^t y_l, x^l \right\rangle \\ = \frac{1}{2}\left\langle x^l, (\mathrm{Id} + \Pi M_{ll}) x^l \right\rangle + \langle \underbrace{(\Pi s^* + \Pi M_{ll} K P_l^t y_l)}_{\in \left\{P_l x^l = 0\right\}},\ x^l \rangle.\end{aligned}$$

Note that if M satisfies CS(ε) with $\varepsilon < 1$, then the map $(\mathrm{Id} + \Pi M_{ll}) : \left\{P_l x^l = 0\right\} \to \left\{P_l x^l = 0\right\}$ is invertible. We define v by

$$v = (\mathrm{Id} - \Pi M_{ll})^{-1} (\Pi s^* + \Pi M_{ll} K P_l^t y_l). \tag{39}$$

It follows from $v \in \left\{P_l x^l = 0\right\}$ that the transformation $x^l \mapsto x^l + v$ leaves the integration space $\left\{P_l x^l = 0\right\}$ invariant. Direct calculation using the definition of v yields

$$\begin{aligned}&\frac{1}{2}\left\langle x^l - v, (\mathrm{Id} + M_{ll})(x^l - v) \right\rangle + \left\langle s^* + M_{ll} K P_l^t y_l, x^l - v \right\rangle \\ &= \frac{1}{2}\left\langle x^l - v, (\mathrm{Id} + \Pi M_{ll})(x^l - v) \right\rangle + \langle (\Pi s^* + \Pi M_{ll} K P_l^t y_l),\ x^l - v \rangle \\ &= \frac{1}{2}\left\langle x^l, (\mathrm{Id} + \Pi M_{ll}) x^l \right\rangle - \langle (\Pi s^* + \Pi M_{ll} K P_l^t y_l),\ v \rangle + \frac{1}{2}\langle v, (\mathrm{Id} + \Pi M_{ll}) v \rangle .\end{aligned}$$

Because we can cancel all terms that are independent of x^l with terms of the normalization constant Z, the transformation $x^l \mapsto x^l + v$ applied to the right hand side of (38) yields

$$\begin{aligned}\int \xi(x^l)\, \nu(dx^l | M, s^*) &= \int_{\left\{P_l x^l = 0\right\}} \xi(x^l + N P^* y_l - v) \\ &\times \frac{1}{Z} \exp\left(-\frac{1}{2}\left\langle x^l, (\mathrm{Id} + M_{ll}) x^l \right\rangle - \sum_{i \in B(l)} \delta\psi_i (x_i + y_l - v_i) \right) \mathcal{H}(dx^l).\end{aligned} \tag{40}$$

Note that we have gained compactness by this representation: The unbounded parameters y_l and s^* only enter (40) as an argument of the bounded functions ξ and $\delta\psi_i$. This observation is crucial for the estimation of T_4 and T_5. The calculation also reveals that it is natural to choose

$$\tilde{s}(s^*) = \Pi s^* + \Pi M_{ll} K P_l^t y_l = \left(\mathrm{Id} - K P_l^t P_l\right) \left(s^* + M_{ll} K P_l^t y_l\right) \tag{41}$$

in $\frac{d^2}{dy_l^2}\bar{H}(y_l|0,\tilde{s})$. The reason is that carrying out the two translations from above yields

$$\int \xi(x^l)\,\nu(dx^l|0,\tilde{s}) = \int_{\left\{\frac{1}{K}\sum_{i\in B(l)} x_i=0\right\}} \xi(x^l + KP_l^t y_l - v)$$
$$\times \frac{1}{Z}\exp\left(-\frac{1}{2}\left\langle x^l, x^l\right\rangle - \sum_{i\in B(l)} \delta\psi_i(x_i + y_l - v_i)\right)\mathcal{H}(dx^l). \qquad (42)$$

Note that the right hand side of (40) and (42) coincide except of the interaction term $\left\langle x^l, M_{ll}x^l\right\rangle$. The latter is very helpful to apply a perturbation argument for the uniform estimation of T_4 and T_5, because they have the same structure as

$$\left|\int \xi(x^l)\,\nu(dx^l|M,s^*) - \int \xi(x^l)\,\nu(dx^l|0,\tilde{s})\right|.$$

Now, we deduce the uniform estimation of T_4 and T_5. Let us choose $\tilde{s} = \tilde{s}(s^*)$ as in (41). Firstly, we estimate the term

$$T_4 = \frac{1}{K}\left|\int \sum_{j\in B(l)} \delta\psi_j''(x_j)\,\nu(dx^l|M,s^*) - \int \sum_{j\in B(l)} \delta\psi_j''(x_j)\nu(dx^l|0,\tilde{s})\right|.$$

For $0 \leq \lambda \leq 1$ we define the probability measure ν_λ on $\left\{P_l x^l = 0\right\}$ by

$$\nu_\lambda(dx^l) := \frac{1}{Z}\,\exp\left(-\frac{1}{2}\left\langle x^l, (\mathrm{Id} + \lambda M_{ll})x^l\right\rangle - \sum_{j\in B(l)} \delta\psi_j\,(x_j + y_l - v_j)\right)\,\mathcal{H}(dx^l),$$

where the vector v is defined by (39). Applying the translation $x^l \mapsto x^l - KP_l^t y_l + v$ on the integrals of T_4 yields (cf. (38), (40), and (42))

$$T_4 = \frac{1}{K}\left|\int \sum_{j\in B(l)} \delta\psi_j''(x_j + y_l - v_j)\,\nu_1(dx^l) - \int \sum_{j\in B(l)} \delta\psi_j''(x_j + y_l - v_j)\,\nu_0(dx^l)\right|$$
$$\leq \frac{1}{K}\sup_{0\leq\lambda\leq 1}\left|\frac{d}{d\lambda}\int \sum_{j\in B(l)} \delta\psi_j''(x_j + y_l - v_j)\,\nu_\lambda(dx^l)\right|. \qquad (43)$$

Because M satisfies CS(ε), we may assume w.l.o.g. that

$$-\frac{1}{2}\,\mathrm{Id} \leq M_{ll} \leq \frac{1}{2}\,\mathrm{Id}\,. \qquad (44)$$

Let $C < \infty$ denote a generic constant depending only on K and c_1. By direct calculation we get that for any $0 \leq \lambda \leq 1$

$$
\begin{aligned}
&\left| \frac{d}{d\lambda} \int \sum_{j \in B(l)} \delta\psi_j''(x_j + y_l - v_j)\ \nu_\lambda(dx^l) \right| \\
&= \frac{1}{2} \left| \operatorname{cov}_{\nu_\lambda(dx^l)} \left(\sum_{j \in B(l)} \delta\psi_j''(x_j + y_l - v_j)\ ,\ \langle x^l, M_{ll} x^l \rangle \right) \right| \\
&= \frac{1}{2} \left| \int \left(\sum_{j \in B(l)} \delta\psi_j''(x_j + y_l - v_j) - \int \delta\psi_j''(x_j + y_l - v_j)\ \nu_\lambda(dx^l) \right) \ \langle x^l, M_{ll} x^l \rangle\ \nu_\lambda(dx^l) \right| \\
&\leq K \ \max_{j \in B(l)} \sup_{x \in \mathbb{R}} \left| \delta\psi_j''(x) \right| \ \int \left| \langle x^l, M_{ll} x^l \rangle \right|\ \nu_\lambda(dx^l) \\
&\overset{\mathrm{CS}(\varepsilon)}{\leq} K c_1\ \varepsilon\ \frac{\int_{\{P_l x^l = 0\}} \left| x^l \right|^2 \exp\left(-\frac{1}{2} \langle x^l, (\mathrm{Id} + \lambda M_{ll})\, x^l \rangle - \sum_{j \in B(l)} \delta\psi_j\,(x_j + y_l - v_j) \right)\ \mathcal{H}(dx)}{\int_{\{P_l x^l = 0\}} \exp\left(-\frac{1}{2} \langle x^l, (\mathrm{Id} + \lambda M_{ll})\, x^l \rangle - \sum_{j \in B(l)} \delta\psi_j\,(x_j + y_l - v_j) \right)\ \mathcal{H}(dx)} \\
&. \overset{(44)}{\leq} K\ c_1\ \varepsilon\ \exp\left(K \max_{j \in B(l)} \left(\sup_x \delta\psi_j(x) - \inf_x \delta\psi_j(x) \right) \right) \\
&\qquad\qquad \times \frac{\int_{\{P_l x^l = 0\}} \left| x^l \right|^2 \exp\left(-\frac{1}{2} \langle x^l, x^l \rangle \right)\ \mathcal{H}(dx)}{\int_{\{P_l x^l = 0\}} \exp\left(-\frac{3}{2} \langle x^l, x^l \rangle \right)\ \mathcal{H}(dx)} \\
&\leq C\ \varepsilon. \qquad (45)
\end{aligned}
$$

A combination of (43) and (45) yields the estimate

$$T_4 \leq C\ \varepsilon.$$

The same argument also yields

$$T_5 = \frac{1}{K} \left| \operatorname{var}_{\nu(dx^l | M, s^*)} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j) \right) - \operatorname{var}_{\nu(dx^l | 0, \bar{s})} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j) \right) \right| \leq C\ \varepsilon.$$

There is only one difference compared to the estimation of T_4. It is the identity

$$
\begin{aligned}
&\left| \frac{d}{d\lambda} \operatorname{var}_{\nu_\lambda(dx^l)} \left(\sum_{j \in B(l)} \delta\psi_j'(x_j + y_l - v_j) \right) \right| \\
&= \left| \frac{d}{d\lambda} \int \left(\sum_{j \in B(l)} \delta\psi_j'(x_j + y_l - v_j) - \int \delta\psi_j'(x_j + y_l - v_j)\ \nu_\lambda(dx^l) \right)^2 \nu_\lambda(dx^l) \right|.
\end{aligned}
$$

Because

$$\int \left[\frac{d}{d\lambda} \left(\sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j) - \int \delta\psi_j'(x_j + y_l - v_j)\, \nu_\lambda(dx^l) \right)^2 \right] \nu_\lambda(dx^l)$$

$$= -2 \int \left(\sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j) - \int \delta\psi_j'(x_j + y_l - v_j)\, \nu_\lambda(dx^l) \right) \nu_\lambda(dx^l)$$

$$\times \ \frac{d}{d\lambda} \int \sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j)\, \nu_\lambda(dx^l)$$

$$= 0,$$

it follows by direct calculation that

$$\left| \frac{d}{d\lambda} \operatorname{var}_{\nu_\lambda(dx^l)} \left(\sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j) \right) \right|$$

$$= \left| \int \left(\sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j) - \int \delta\psi_j'(x_j + y_l - v_j)\, \nu_\lambda(dx^l) \right)^2 \left[\frac{d}{d\lambda} \nu_\lambda(dx^l) \right] \right|$$

$$= \left| \operatorname{cov}_{\nu_\lambda(dx^l)} \left(\left(\sum_{j\in B(l)} \delta\psi_j'(x_j + y_l - v_j) - \left\langle \delta\psi_j'(x_j + y_l - v_j) \right\rangle_{\nu_\lambda(dx^l)} \right)^2, \ \left\langle x^l, M_{ll} x^l \right\rangle \right) \right|.$$

However, the covariance term on the right hand side can be estimated as in (45). Therefore we have deduced (30) uniformly in $y_l \in \mathbb{R}$ and $s^* \in \mathbb{R}^{B(l)}$, which completes the proof of Lemma 2.12. □

3 Generalized local Cramér theorem

In this section we generalize the local Cramér theorem [GOVW09, Proposition 31] to a broader class of Hamiltonians. We recommend to read Section 5.4 and 5.5 of [GOVW09] for better understanding of this section. In the first step (cf. Theorem 3.1) we deduce the local Cramér theorem for Hamiltonians of the form

$$H(x) := \sum_{i=1}^{K} \frac{1}{2} x_i^2 + \delta\psi_i(x_i). \tag{46}$$

The difference to [GOVW09] is that the non-convex functions

$$\delta\psi_i : \mathbb{R} \to \mathbb{R}$$

are allowed to depend on the site $i \in \{1, \dots, N\}$. In the second step (cf. Corollary 3.2) we extend the local Cramér theorem to Hamiltonians that also contain a linear term given by a vector $s \in \mathbb{R}^K$, namely

$$H(x) := \sum_{i=1}^{K} \frac{1}{2} x_i^2 + s_i x_i + \delta\psi_i(x_i). \tag{47}$$

Corollary 3.2 is one of the main ingredients to derive the macroscopic LSI (cf. Proposition 2.4, Lemma 2.12 and Remark 3.3). The coarse-grained Hamiltonian of H is defined by

$$\bar{H}(m) := \psi_K(m) := -\frac{1}{K} \log \int_{X_{K,m}} \exp\left(-H\left(x\right)\right) \mathcal{H}\left(dx\right), \tag{48}$$

where the hyper-plane $X_{K,m}$ is given by

$$X_{K,m} := \left\{ x \in \mathbb{R}^K \mid \frac{1}{K} \sum_{i=1}^{K} x_i = m \right\}.$$

Let $\varphi_K(m)$ be defined as the Cramér transform of $\exp(-H(x))$ (cf. [GOVW09, (73)]). More precisely,

$$\varphi_K(m) := \sup_{\sigma \in \mathbb{R}} \left(\sigma m - \frac{1}{K} \log \int_{\mathbb{R}^K} \exp\left(-H(x) + \sum_{i=1}^{K} \sigma x_i \right) dx \right). \tag{49}$$

By elementary properties of the Legendre transform for every $m \in \mathbb{R}$ there exits a unique $\sigma_m \in \mathbb{R}$ such that

$$\varphi_K(m) = \sigma_m m - \varphi_K^*(\sigma_m). \tag{50}$$

It is also well-known that σ_m is determined by the equation

$$m = \frac{d}{d\sigma} \varphi_K^*(\sigma_m). \tag{51}$$

The first generalization of the local Cramér theorem is

Theorem 3.1 (Local Cramér theorem)**.** *Let H be defined by* (46) *and assume that the functions $\delta\psi_i : \mathbb{R} \to \mathbb{R}$, $i \in \{1, \dots, N\}$, satisfy the uniform bound* (1) *i.e.*

$$\|\delta\psi_i\|_{C^2} \leq c_1 < \infty, \qquad i \in \{1, \dots, K\}.$$

Then the function φ_K is strictly convex independent of K, and

$$\|\bar{H}(m) - \varphi_K(m)\|_{C^2} \to 0 \qquad \text{as } K \to \infty.$$

The convergence depends only on c_1.

The following heuristic argument shows that this generalization of the local Cramér theorem [GOVW09, Proposition 31] is not surprising. By Cramér's representation (see [GOVW09, (125)] and Lemma 3.4) the proof of [GOVW09, Proposition 31] relies on a local central limit theorem for the independent random variables X_i, $i \in \{1, \ldots, K\}$, identically distributed as (cf. [GOVW09, (74)])

$$Z^{-1} \exp\left(\sigma_m x_i - \frac{x_i^2}{2} - \delta\psi(x_i)\right) dx_i.$$

However, for the classical central limit theorem it is not important that the random variables X_i are identically distributed. It suffices that the expectation and variance of X_i is uniformly bounded. The latter is true for the independent random variables X_i distributed as (cf. Lemma 3.4)

$$Z^{-1} \exp\left(\sigma_m x_i - \frac{x_i^2}{2} - \delta\psi_i(x_i)\right) dx_i,$$

if the functions $\delta\psi_i$ satisfy the uniform bound (1). As a consequence we can apply the same strategy as for [GOVW09, Proposition 31]. We just have to pay attention that every step does not rely on the specific form of $\delta\psi_i$ but on the uniform bound (1). Because the complete proof of Theorem 3.1 is a bit lengthy, we will state the details in the next section. The second generalization of local Cramér theorem is

Corollary 3.2. *For an arbitrary vector $s \in \mathbb{R}^K$ let H be defined by* (47) *and assume that the functions $\delta\psi_i : \mathbb{R} \to \mathbb{R}$, $i \in \{1, \ldots, N\}$, satisfy the uniform bound* (1) *i.e.*

$$\|\delta\psi_i\|_{C^2} \leq c_1 < \infty, \qquad i \in \{1, \ldots, K\}.$$

Then the function φ_K is strictly convex independent of K and s. It holds that

$$\|\bar{H}(m) - \varphi_K(m)\|_{C^2} \to 0 \qquad \text{as } K \to \infty.$$

The convergence depends only on c_1.

Again, this generalization of the local Cramér theorem is intuitively clear: A linear term should not influence the local Cramér theorem by the equivalence of ensembles (see [GOVW09]). The proof is also straightforward. Corollary 3.2 follows from Theorem 3.1 by a linear transformation.

Proof of Corollary 3.2. Let the Hamiltonian H be defined by (47). We get rid of the linear term $\langle s, x\rangle$ in the coarse-grained Hamiltonian $\bar{H}$ by the linear transformation $x \mapsto x + s$. More precisely, let

$$\tilde{H}(x) := \sum_{i=1}^{K} \frac{x_i^2}{2} + \delta\psi_i(x_i - s_i).$$

Then direct calculation shows that

$$H(x) = \tilde{H}(x+s) - \frac{1}{2}\langle s, s\rangle,$$

which yields by the transformation $x \mapsto x + s$ that

$$\begin{aligned}\bar{H}(m) &= -\frac{1}{K}\log\int_{X_{K,m}} \exp\left(-H\left(x\right)\right)\,\mathcal{H}\left(dx\right)\\ &= \underbrace{-\frac{1}{K}\log\int_{X_{K,m+\frac{1}{K}\sum_{i=1}^K s_i}} \exp\left(-\tilde{H}\left(x\right)\right)\,\mathcal{H}\left(dx\right)}_{=:\tilde{\psi}_K(m+\frac{1}{K}\sum_{i=1}^K s_i)} - \frac{1}{2K}\langle s, s\rangle.\end{aligned}$$

Applying Theorem 3.1 to $\tilde{\psi}_K(m + \frac{1}{K}\sum_{i=1}^K s_i)$ yields

$$\left\|\bar{H}(m) + \frac{1}{2K}\langle s, s\rangle - \tilde{\varphi}_K\left(m + \frac{1}{K}\sum_{i=1}^K s_i\right)\right\|_{C^2} \to 0, \qquad \text{for } K \to \infty,$$

where $\tilde{\varphi}_K$ is the Cramér transform of $\exp(-\tilde{H}(x))$ defined by (49). Then Corollary 3.2 follows from

$$\begin{aligned}&\tilde{\varphi}_K\left(m + \frac{1}{K}\sum_{i=1}^K s_i\right)\\ &\quad= \sup_{\sigma\in\mathbb{R}}\left(\sigma\left(m + \frac{1}{K}\sum_{i=1}^K s_i\right) - \frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-\tilde{H}(x) + \sum_{i=1}^K \sigma x_i\right)\,dx\right)\\ &\quad= \sup_{\sigma\in\mathbb{R}}\left(\sigma m - \frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-H(x) + \sum_{i=1}^K \sigma x_i\right)\,dx\right) + \frac{1}{2K}\langle s, s\rangle\\ &\quad= \varphi_K(m) + \frac{1}{2K}\langle s, s\rangle,\end{aligned}$$

where we used the identity

$$\begin{aligned}&-\frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-\tilde{H}(x) + \sum_{i=1}^K \sigma x_i\right)\,dx\\ &\quad= -\frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-\tilde{H}(x+s) + \sum_{i=1}^K \sigma(x_i + s_i)\right)\,dx\\ &\quad= -\frac{\sigma}{K}\sum_{i=1}^K s_i - \frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-\tilde{H}(x+s) + \frac{1}{2}\langle s, s\rangle + \sum_{i=1}^K \sigma x_i\right)\,dx + \frac{1}{2K}\langle s, s\rangle\\ &\quad= -\frac{\sigma}{K}\sum_{i=1}^K s_i - \frac{1}{K}\log\int_{\mathbb{R}^K}\exp\left(-H(x) + \sum_{i=1}^K \sigma x_i\right)\,dx + \frac{1}{2K}\langle s, s\rangle.\end{aligned}$$

□

Remark 3.3. *By Corollary 3.2 it follows that there is $K_0 \in \mathbb{N}$ and $\lambda > 0$ such that for all $K \geq K_0$ and $m \in \mathbb{R}$*

$$\frac{d^2}{dm^2} \bar{H}(m) \geq \lambda.$$

This observation shows that the convexification of a non-interacting coarse-grained Hamiltonian is untouched by a linear term and inhomogeneous single-site potentials ψ_i that satisfy (1).

Proof of Theorem 3.1

Except of small adjustments, the argument for Theorem 3.1 is the same as for the local Cramér theorem of [GOVW09]. Therefore some elements of the proof may also be found in [Fel71, Chapter XVI], [KL99, Appendix 2], [GPV88, Section 3] and [LPY02, p. 752 and Section 5].

We show that φ_K is strictly convex independent of K. Note that by (49) the function φ_K is the Legendre transform of

$$\varphi_K^*(\sigma) := \frac{1}{K} \log \int_{\mathbb{R}^K} \exp\left(\sum_{i=1}^K \sigma x_i - \frac{x_i^2}{2} - \delta \psi_i(x_i) \right) \, dx.$$

Because the Legendre transform of a strictly convex function is also strictly convex, it suffices to show that $\varphi_K^*(\sigma)$ is strictly convex. This fact follows from the decomposition of $\varphi_K^*(\sigma)$ into

$$\varphi_K^*(\sigma) = \frac{1}{K} \sum_{i=1}^K \underbrace{\log \int_{\mathbb{R}} \exp\left(\sigma x_i - \frac{x_i^2}{2} - \delta \psi_i(x_i) \right) \, dx_i}_{=: \varphi_{K,i}^*(\sigma)} = \frac{1}{K} \sum_{i=1}^K \varphi_{K,i}^*(\sigma) \qquad (52)$$

and the observation that by Lemma 3.7 from below the functions $\varphi_{K,i}^*$ are uniformly strictly convex in $i \in \{1, \ldots, K\}$.

Let us consider now the convergence of $|\varphi_K(m) - \psi_K(m)|$. The next lemma represents the difference $(\varphi_K(m) - \psi_K(m))$ in the same way as [GOVW09, (120)].

Lemma 3.4. *Assume that X_i, $i \in \{1, \ldots, K\}$, are independent random variables distributed as*

$$d\mu_{\sigma_m, i} := \exp\left(-\varphi_{K,i}^*(\sigma_m) + \sigma_m x_i - \frac{x_i^2}{2} - \delta \psi_i(x_i) \right) \, dx_i,$$

where σ_m is given by (51). *For $i \in \{1, \dots, K\}$ we define $m_i = m_i(\sigma_m) \in \mathbb{R}$ by*

$$m_i = \frac{d}{d\sigma} \varphi^*_{K,i}(\sigma_m). \tag{53}$$

Additionally, let the random variable X be defined as

$$X := K^{-\frac{1}{2}} \sum_{i=1}^{K} X_i - m_i.$$

The density of X with respect to the Lebesgue measure is denoted by $\frac{d}{d\mathcal{L}^1}\, g_{K,m}(\xi)$. Then

$$\frac{d}{d\mathcal{L}^1}\, g_{K,m}(0) = \exp(K\varphi_K(m) - K\psi_K(m)). \tag{54}$$

Proof of Lemma 3.4. Using (52), (51), and (53) we can decompose m into

$$m = \frac{d}{d\sigma}\varphi^*_K(\sigma_m) = \frac{1}{K}\sum_{i=1}^{K} \frac{d}{d\sigma}\varphi^*_{K,i}(\sigma_m) = \frac{1}{K}\sum_{i=1}^{K} m_i. \tag{55}$$

The density $\frac{d}{d\mathcal{L}^1} g_{K,m}(\xi)$ of X at $\xi = 0$ can be written as

$$\frac{d}{d\mathcal{L}^1} g_{K,m}(0) = \int_{\left\{K^{-\frac{1}{2}}\sum_{i=1}^{K} x_i - m_i = 0\right\}} \exp\left(\sum_{i=1}^{K} -\varphi^*_{K,i}(\sigma_m) + \sigma_m x_i - \frac{x_i^2}{2} - \delta\psi_i(x_i)\right) \mathcal{H}(dx).$$

By (52) and (55) we get

$$\frac{d}{d\mathcal{L}^1} g_{K,m}(0) = \int_{\left\{\sum_{i=1} x_i = Km\right\}} \exp\left(-K\varphi^*_K(\sigma_m) + K\sigma_m m - \left(\sum_{i=1}^{K} \frac{x_i^2}{2} + \delta\psi_i(x_i)\right)\right) \mathcal{H}(dx).$$

Using (50) the right hand side becomes

$$\frac{d}{d\mathcal{L}^1} g_{K,m}(0) = \exp\left(K\varphi_K(m)\right) \int_{\left\{\sum_{i=1}^{K} x_i = Km\right\}} \exp\left(-\sum_{i=1}^{K} \frac{x_i^2}{2} + \delta\psi_i(x_i)\right) \mathcal{H}(dx).$$

Now, applying the definition (48) of $\psi_K(m)$ yields the desired formula. □

Note that Theorem 3.1 is proved once we established the following bounds

$$\frac{1}{C} \leq \left| \frac{d}{d\mathcal{L}^1} g_{K,m}(0) \right| \leq C \quad \text{and} \quad \left| \frac{d^2}{dm^2} \frac{d}{d\mathcal{L}^1} g_{K,m}(0) \right| \leq C$$

for a constant $0 < C < \infty$ independent of K and m (cf. [GOVW09, (121)]). Because the density of a sum of independent random variables is the product of the densities, an application of the Inversion Lemma (see for example [Shi96]) to (54) yields

$$\frac{d}{d\mathcal{L}^1} g_{K,m}(0) = \frac{1}{2\pi} \int_{\mathbb{R}} \Pi_{i=1}^{K} h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi,$$

where the function $h_i(m_i, \xi)$ is defined as

$$h_i(m_i, \xi) := \int_{\mathbb{R}} \exp\left(-i\xi m_i + i\xi x - \varphi_{K,i}^{*}(\sigma_m) + \sigma_m x - \frac{x^2}{2} - \delta\psi_i(x) \right) dx. \quad (56)$$

Note the representation of $\frac{d}{d\mathcal{L}^1} g_{K,m}(0)$ from above is the analog of [GOVW09, (122)] and that

$$h_i(m_i, \xi) = h(m_i, \xi),$$

where $h(m_i, \xi)$ is defined as in [GOVW09, (126)] (one has to set $\delta\psi = \delta\psi_i$). Hence it suffices to deduce the following estimates:

$$\frac{1}{C} \leq \left| \int_{\mathbb{R}} \Pi_{i=1}^{K} h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \right| \leq C \quad (57)$$

and

$$\left| \frac{d^2}{dm^2} \int_{\mathbb{R}} \Pi_{i=1}^{K} h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \right| \leq C, \quad (58)$$

where the constant $0 < C < \infty$ is independent of K and m.

The inequalities (57) and (58) are established by the same argument as in the proof of [GOVW09, Proposition 31]. The latter works fine because all steps in [GOVW09] do not rely on the particular form of the non-convexity $\delta\psi_i$. They just require a uniform bound on $\|\delta\psi_i\|_{C^2}$, $i \in \{1, \ldots, K\}$. For the sake of completeness, we outline all steps in full detail below.

From now on let $C < \infty$ denote a generic constant independent of K and m. We verify (57) and (58) by splitting the integrals into "inner" and "outer" integrals. On the one hand we will show that there exists $\delta > 0$ and $K_0 \in \mathbb{N}$ such that for all $K \geq K_0$ and all

$m \in \mathbb{R}$

$$\left| \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \right| \leq C, \tag{59}$$

$$\mathrm{Re} \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \geq \frac{1}{C}, \tag{60}$$

$$\left| \frac{d^2}{dm^2} \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \right| \leq C. \tag{61}$$

On the other hand we will show that for any $\delta > 0$

$$\lim_{K \uparrow \infty} \int_{K^{-\frac{1}{2}}|\xi| \geq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi = 0, \tag{62}$$

$$\lim_{K \uparrow \infty} \frac{d^2}{dm^2} \int_{K^{-\frac{1}{2}}|\xi| \geq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi = 0. \tag{63}$$

Firstly, we will consider the limits (62) and (63) for the "outer" integrals. Note that (62) and (63) is the analog of [GOVW09, (137) and (138)]. The next lemma is the analog statement of [GOVW09, Lemma 39], which was the essential ingredient to deduce [GOVW09, (137) and (138)].

Lemma 3.5. *Let functions h_i, $i \in \{1, \ldots, K\}$, be defined by* (56)*. Then for any $\delta > 0$ there exists a positive constant C_δ such that for all $i \in \{1, \ldots, K\}$, $|\xi| > \delta$ and $m \in \mathbb{R}$*

$$(i) \qquad |h_i(m_i, \xi)| \leq \frac{1}{1 + \frac{|\xi|}{C_\delta}},$$

$$(ii) \qquad \left| \frac{d}{dm}\, h_i(m_i, \xi) \right| \leq C_\delta\, |\xi|,$$

$$(iii) \qquad \left| \frac{d^2}{dm^2}\, h_i(m_i, \xi) \right| \leq C_\delta\, |\xi|^2,$$

where $m_i(m) \in \mathbb{R}$ is given by (53).

The proof of Lemma 3.5 is given in the next subsection. Provided Lemma 3.5 holds, (62) and (63) can be deduced in exactly the same way as [GOVW09, (137) and (138)].

Let us consider the "inner" integrals (59), (60), and (61) which are the analog of [GOVW09, (134),(135) and (136)]. As in [GOVW09, (139)] we have that for all $i \in \{1, \ldots, K\}$ and $m_i \in \mathbb{R}$

$$h_i(m_i, 0) = 1, \qquad \frac{d}{d\xi} h_i(m_i, 0) = 0 \qquad \text{and}$$

$$-\frac{d^2}{d\xi^2} h_i(m_i, 0) = \int (x - m_i)^2 \mu_{\sigma_m, i}(dx) > 0.$$

According to Lemma 3.7 from the next subsection we have the uniform bound

$$\frac{1}{C} \leq \int (x - m_i)^2 \mu_{\sigma_m,i}(dx) \leq C \qquad \text{for all } i \in \{1, \ldots, K\}.$$

By using the lower bound and Taylor's theorem one can show the existence of functions $h_{2,i}(m_i, \xi)$, $i \in \{1, \ldots, K\}$ defined on a uniform δ-neighborhood of $\xi = 0$ such that for all $i \in \{1, \ldots, K\}$ and $m_i \in \mathbb{R}$ (cf. [GOVW09, (141)])

$$h_i(m_i, \xi) = \exp\left(-\xi^2 h_{2,i}(m_i, \xi)\right) \tag{64}$$

and

$$\frac{1}{C} \leq h_{2,i}(m_i, 0) \leq C. \tag{65}$$

The next lemma is the analog of [GOVW09, Lemma 40] and provides strong control on the functions $h_{2,i}$ and its derivatives. We will prove it in the next subsection.

Lemma 3.6. *There exists $\delta > 0$ and $C < \infty$ such that for all $i \in \{1, \ldots, K\}$, $|\xi| \leq \delta$, and m_i*

$$(i) \quad \left|\frac{d}{d\xi} h_{2,i}(m_i, \xi)\right| \leq C, \quad (ii) \quad \left|\frac{d}{dm} h_{2,i}(m_i, \xi)\right| \leq C, \quad (iii) \quad \left|\frac{d^2}{dm^2} h_{2,i}(m_i, \xi)\right| \leq C,$$

where $m_i(m) \in \mathbb{R}$ is given by (53).

Now, we are able to verify (59) with the same argument as for [GOVW09, (134)]. It holds that for any $|\tilde{\xi}| \leq \delta$ and m_i (cf. [GOVW09, (139)])

$$\operatorname{Re} h_{2,i}(m_i, \tilde{\xi}) \geq \frac{1}{C}, \tag{66}$$

uniformly in $i \in \{1, \ldots, K\}$. Thus we can estimate by using (64) and (66) that for $K^{-\frac{1}{2}}|\xi| \leq \delta$

$$\begin{aligned}
\left| \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi)\, d\xi \right| &\leq \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \exp\left(-\frac{\xi^2}{K} \sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right)\, d\xi \\
&\leq \int_{K^{-\frac{1}{2}}|\xi| \leq \delta} \exp\left(-\frac{\xi^2}{C}\right)\, d\xi \\
&\leq C.
\end{aligned}$$

The argument for (61) is almost the same as for [GOVW09, (135)]. By (64) we have

$$\Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi) = \exp\left(-\frac{\xi^2}{K} \sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right),$$

which yields the identity

$$\frac{d^2}{dm^2}\Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi) = -\frac{\xi^2}{K}\sum_{i=1}^K \frac{d^2}{dm^2}\left(h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right)\exp\left(-\frac{\xi^2}{K}\sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right)$$
$$+\frac{\xi^4}{K^2}\left(\sum_{i=1}^K \frac{d}{dm}\left(h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right)\right)^2 \exp\left(-\frac{\xi^2}{K}\sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right).$$

By using the estimates of Lemma 3.6 we get

$$\frac{d^2}{dm^2}\Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}}\xi) \le C_\delta\left(\xi^2+\xi^4\right)\exp\left(-\frac{\xi^2}{K}\sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right).$$

This inequality and (66) yields

$$\left|\frac{d^2}{dm^2}\int_{\{K^{-\frac{1}{2}}|\xi|\le\delta\}}\Pi_{i=1}^K h_i(m_i, K^{-\frac{1}{2}})\xi)\, d\xi\right|$$
$$\le C\int_{\{K^{-\frac{1}{2}}|\xi|\le\delta\}}\left(\xi^2+\xi^4\right)\exp\left(-\frac{\xi^2}{C}\right)\, d\xi \le C.$$

Finally, we verify the last remaining estimate (60). The argument is essentially the same as for [GOVW09, (141)]. We introduce $h_{3,i}$, $i \in \{1,\dots,K\}$ via

$$h_{2,i}(m_i,\hat{\xi}) = h_{2,i}(m_i,0) + \hat{\xi}h_{3,i}(m_i,\hat{\xi}),$$

which according to Taylor and Lemma 3.6 i) satisfies

$$\sup_{|\hat{\xi}|\le\delta}|h_{3,i}(m_i,\hat{\xi})| \le \sup_{|\hat{\xi}|\le\delta}\left|\frac{d}{d\hat{\xi}}h_{2,i}(m_i,\hat{\xi})\right| \le C, \tag{67}$$

uniformly in $i \in \{1,\dots,K\}$. By the definition of $h_{3,i}$ and (65) we have

$$\left|\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^K h_{2,i}(m_i, K^{-\frac{1}{2}}\xi)\right) - \exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^K h_{2,i}(m_i,0)\right)\right|$$
$$= \exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^K h_{2,i}(m_i,0)\right)\left|\exp\left(-K^{-\frac{1}{2}}\xi^3\frac{1}{K}\sum_{i=1}^K h_{3,i}(m_i, K^{-\frac{1}{2}}\xi)\right)-1\right|$$
$$\le \exp\left(-\frac{\xi^2}{C}\right)\left|\exp\left(-K^{-\frac{1}{2}}\xi^3\frac{1}{K}\sum_{i=1}^K h_{3,i}(m_i, K^{-\frac{1}{2}}\xi)\right)-1\right|.$$

We apply the fact

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

to

$$z = -K^{-\frac{1}{2}}\xi^3\frac{1}{K}\sum_{i=1}^{K}h_{3,i}(m_i,K^{-\frac{1}{2}})$$

and conclude from the last estimate that

$$\begin{aligned}&\left|\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,K^{-\frac{1}{2}}\xi)\right)-\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,0)\right)\right|\\&\qquad\le \exp\left(-\frac{\xi^2}{C}\right)\left|\exp\left(K^{-\frac{1}{2}}|\xi|^3\frac{1}{K}\sum_{i=1}^{K}|h_{3,i}(m_i,K^{-\frac{1}{2}}\xi)|\right)-1\right|.\end{aligned}$$

If $K^{-\frac{1}{2}}|\xi|\le\delta$, we can continue the last estimation using (67) as

$$\begin{aligned}&\left|\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,K^{-\frac{1}{2}}\xi)\right)-\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,0)\right)\right|\\&\qquad\le \exp\left(-\frac{\xi^2}{C}\right)\left(\exp\left(C\delta\xi^2\right)-1\right).\end{aligned}$$

Hence we get as in [GOVW09, (144)] that for δ sufficiently small

$$\begin{aligned}&\left|\int_{\{K^{-\frac{1}{2}}|\xi|\le\delta\}}\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,K^{-\frac{1}{2}}\xi)\right)-\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,0)\right)\,d\xi\right|\\&\quad\le\left|\int_{\mathbb{R}}\exp\left(-\xi^2\left(\frac{1}{C}-C\delta\right)\right)-\exp\left(-\frac{\xi^2}{C}\right)\,d\xi\right|\\&\quad\le C\left(\frac{1}{\sqrt{\frac{1}{C}-C\delta}}-\frac{1}{\sqrt{\frac{1}{C}}}\right)\\&\quad\le C\delta.\end{aligned}$$

By (65) we have

$$
\begin{aligned}
\int_{\{K^{-\frac{1}{2}}|\xi|\leq\delta\}} &\exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,0)\right)\,d\xi \\
&\geq \int_{\{K^{-\frac{1}{2}}|\xi|\leq\delta\}} \exp\left(-\frac{\xi^2}{C}\right)\,d\xi \\
&= \int_{\mathbb{R}} \exp\left(-\frac{\xi^2}{C}\right)\,d\xi - \int_{\{K^{-\frac{1}{2}}|\xi|\geq\delta\}} \exp\left(-\frac{\xi^2}{C}\right)\,d\xi \\
&\geq C\sqrt{\frac{1}{C}} - C\exp\left(-\frac{K\delta^2}{C}\right).
\end{aligned}
$$

Finally, the last two estimates yield

$$
\begin{aligned}
\mathrm{Re}\int_{\{K^{-\frac{1}{2}}|\xi|\leq\delta\}} &\Pi_{i=1}^{K} h_i(m_i, K^{-\frac{1}{2}}\xi)\,d\xi \\
&= \mathrm{Re}\int_{\{K^{-\frac{1}{2}}|\xi|\leq\delta\}} \exp\left(-\xi^2\frac{1}{K}\sum_{i=1}^{K}h_{2,i}(m_i,K^{-\frac{1}{2}}\xi)\right)\,d\xi \\
&\geq C\sqrt{\frac{1}{C}} - C\exp\left(-\frac{K\delta^2}{C}\right) - C\delta \\
&\geq C > 0,
\end{aligned}
$$

if δ is sufficiently small and K is sufficiently large. Hence we verified (60), which completes the proof of Theorem 3.1.

Proof of auxiliary lemmas

In this section we will state the proof of Lemma 3.5 and Lemma 3.6. The next statement is the analog of [GOVW09, Lemma 41].

Lemma 3.7. *Assume that the functions $\delta\psi_i : \mathbb{R} \to \mathbb{R}$, $i \in \{1,\ldots,N\}$, satisfy the uniform bound* (1) *i.e.*

$$
\|\delta\psi_i\|_{C^2} \leq c_1 < \infty, \qquad i \in \{1,\ldots,K\}.
$$

Consider the change of variables

$$
m_i(\sigma) = \frac{d}{d\sigma}\,\varphi^*_{K,i}(\sigma),
$$

where $\varphi^*_{K,i}(\sigma)$ is defined by (52)*. The corresponding measure $\mu_{\sigma,i}$ is defined by*

$$
d\mu_{\sigma,i} := \exp\left(-\varphi^*_{K,i}(\sigma) + \sigma x_i - \frac{x_i^2}{2} - \delta\psi_i(x_i)\right)\,dx_i.
$$

Note that $\int_{\mathbb{R}} x\mu_{\sigma,i} = m_i$. Then:

(i) The first two derivatives of m_i are related to the moments of $\mu_{\sigma,i}$ as:

$$\frac{d}{d\sigma}\, m_i = \frac{d^2}{d\sigma^2}\, \varphi^*_{K,i} = \int_{\mathbb{R}} (x - m_i)^2\, \mu_{\sigma,i}(dx),$$
$$\frac{d^2}{d\sigma^2}\, m_i = \frac{d^3}{d\sigma^3}\, \varphi^*_{K,i} = \int_{\mathbb{R}} (x - m_i)^3\, \mu_{\sigma,i}(dx).$$

(ii) The moments of $\mu_{\sigma,i}$ satisfy the bounds

$$\frac{1}{C} \leq \int_{\mathbb{R}} (x - m_i)^2\, \mu_{\sigma,i}(dx) \leq C,$$
$$\left| \int_{\mathbb{R}} (x - m_i)^3\, \mu_{\sigma,i}(dx) \right| \leq C,$$
$$\int_{\mathbb{R}} (x - m_i)^4\, \mu_{\sigma,i}(dx) \leq C,$$

uniformly in i and m_i.

(iii) The second derivatives of the inverse map are bounded uniformly in i and m_i:

$$\left| \frac{d^2}{dm_i^2}\, \sigma \right| \leq C.$$

(iv) The map is close to the identity uniformly in i and m_i: $|\sigma - m_i| \leq C$.

We do not state the proof of Lemma 3.7 because one could copy the proof [GOVW09, Lemma 41] using the uniform bound (1). From Lemma 3.7 we are able to deduce the next statement, which is the only new ingredient of the proof of Lemma 3.5 & 3.6 compared to the proof of [GOVW09, Lemma 39 & 40].

Lemma 3.8. *Assume that the functions $\delta\psi_i : \mathbb{R} \to \mathbb{R}$, $i \in \{1, \ldots, N\}$, satisfy the uniform bound* (1) *i.e.*

$$\|\delta\psi_i\|_{C^2} \leq c_1 < \infty, \qquad i \in \{1, \ldots, K\}.$$

Consider change of variables

$$m_i(\sigma) = \frac{d}{d\sigma}\, \varphi^*_{K,i}(\sigma),$$

where $\varphi^*_{K,i}(\sigma)$ is defined by (52). *For $m \in \mathbb{R}$ let $\sigma_m \in \mathbb{R}$ be defined by* (51). *Then:*

$$(i) \quad \left| \frac{d}{dm}\, \sigma_m \right| \leq C, \qquad (ii) \quad \left| \frac{d^2}{dm^2}\, \sigma_m \right| \leq C,$$
$$(iii) \quad \left| \frac{d}{dm}\, m_i(\sigma_m) \right| \leq C, \qquad (iv) \quad \left| \frac{d^2}{dm^2}\, m_i(\sigma_m) \right| \leq C,$$

uniformly in i and m.

Proof of Lemma 3.8. Argument for (i): It suffices to show the uniform bound

$$0 < \frac{1}{C} \leq \left| \frac{d}{d\sigma} \, m \right| \leq C. \tag{68}$$

By (51) and (52) we have

$$\frac{d}{d\sigma} \, m = \frac{d^2}{d\sigma^2} \, \varphi_K^*(\sigma_m) = \frac{1}{K} \sum_{i=1}^{K} \frac{d^2}{d\sigma^2} \, \varphi_{K,i}^*(\sigma_m).$$

Applying Lemma 3.7 (i) & (ii) yields the estimate (68).
Argument for (ii): By direct calculation we have

$$\begin{aligned} \frac{d^2}{dm^2} \, \sigma_m &= \frac{d}{dm} \, \left(\frac{d}{d\sigma} \, m(\sigma_m) \right)^{-1} \\ &= \left[\frac{d}{d\sigma} \left(\frac{d}{d\sigma} \, m(\sigma_m) \right)^{-1} \right] \frac{d}{dm} \, \sigma_m \\ &= \left[- \left(\frac{d}{d\sigma} \, m(\sigma_m) \right)^{-2} \frac{d^2}{d\sigma^2} \, m(\sigma_m) \right] \frac{d}{dm} \, \sigma_m. \end{aligned}$$

By the uniform bound (68) and Lemma 3.8 (i) it suffices to show

$$\left| \frac{d^2}{d\sigma^2} \, m \right| \leq C.$$

By (51) and (52) we have

$$\frac{d^2}{d\sigma^2} \, m = \frac{d^3}{d\sigma^3} \, \varphi_K^*(\sigma_m) = \frac{1}{K} \sum_{i=1}^{K} \frac{d^3}{d\sigma^3} \, \varphi_{K,i}^*(\sigma_m).$$

The desired bound follows from Lemma 3.7 (i) & (ii).
Argument for (iii): By the definition of m_i we have

$$\frac{d}{dm} \, m_i(\sigma_m) = \frac{d}{dm} \, \left(\frac{d}{d\sigma} \, \varphi_{K,i}^*(\sigma_m) \right) = \frac{d^2}{d\sigma^2} \, \varphi_{K,i}^*(\sigma_m) \, \frac{d}{dm} \, \sigma_m.$$

The desired estimate follows from Lemma 3.7 (i) & (ii) and Lemma 3.8 (i).
Argument for (iv): By the definition of m_i it holds

$$\frac{d^2}{dm^2} \, m_i(\sigma_m) = \frac{d^2}{d\sigma^2} \, \varphi_{K,i}^*(\sigma_m) \, \frac{d^2}{dm^2} \, \sigma_m + \frac{d^3}{d\sigma^3} \, \varphi_{K,i}^*(\sigma_m) \, \left(\frac{d}{dm} \, \sigma_m \right)^2.$$

The desired estimate follows from a combination of Lemma 3.7 (i) & (ii) and Lemma 3.8 (i) & (ii). □

Now, we can proceed to the proof of Lemma 3.5 and Lemma 3.6.

Proof of Lemma 3.5. Note that [GOVW09, Lemma 41] was the crucial ingredient in the proof of [GOVW09, Lemma 39]. Therefore we can show by using Lemma 3.8 in the same way as [GOVW09, Lemma 41] that for any $\delta > 0$ there exists a positive constant C_δ such that for all $i \in \{1, \ldots, K\}$, $|\xi| > \delta$, and $m_i \in \mathbb{R}$

$$(i) \qquad |h_i(m_i, \xi)| \leq \frac{1}{1 + \frac{|\xi|}{C_\delta}},$$

$$(ii) \qquad \left| \frac{d}{dm_i} \, h_i(m_i, \xi) \right| \leq C_\delta \, |\xi|,$$

$$(iii) \qquad \left| \frac{d^2}{dm_i^2} \, h_i(m_i, \xi) \right| \leq C_\delta \, |\xi|^2,$$

This statement differs from [GOVW09, Lemma 39] in one aspect: The bound C_δ is uniform in i. Note that the last statement almost yields Lemma 3.5 except of one detail: Instead of considering derivatives w.r.t m it considers derivatives w.r.t. m_i. It follows that the statement (i) of Lemma 3.5 is already verified.
The statement (ii) of Lemma 3.5 follows from a combination of the identity

$$\frac{d}{dm} \, h_i(m_i, \xi) = \frac{d}{dm_i} \, h_i(m_i, \xi) \, \frac{d}{dm} \, m_i,$$

the estimate (ii) from above, and Lemma 3.8 (iii).
The statement (iii) of Lemma 3.5 follows from a combination of the identity

$$\frac{d^2}{dm^2} \, h_i(m_i, \xi) = \frac{d^2}{dm_i^2} \, h_i(m_i, \xi) \, \left(\frac{d}{dm} m_i \right)^2 + \frac{d}{dm_i} \, h_i(m_i, \xi) \, \frac{d^2}{dm^2} \, m_i,$$

the estimate (iii) from above, Lemma 3.8 (iii) & (iv), and Lemma 3.5 (ii). □

Proof of Lemma 3.6. Note that [GOVW09, Lemma 41] was the crucial ingredient in the proof of [GOVW09, Lemma 40]. Therefore we can show by using Lemma 3.8 in the same way as [GOVW09, Lemma 41] that there exists $\delta > 0$ and $C < \infty$ such that for all $i \in \{1, \ldots, K\}$, $|\xi| \leq \delta$, and $m_i \in \mathbb{R}$

$$(i) \qquad \left| \frac{d}{d\xi} \, h_{2,i}(m_i, \xi) \right| \leq C,$$

$$(ii) \qquad \left| \frac{d}{dm_i} \, h_{2,i}(m_i, \xi) \right| \leq C,$$

$$(iii) \qquad \left| \frac{d^2}{dm_i^2} \, h_{2,i}(m_i, \xi) \right| \leq C.$$

The last statement already yields Lemma 3.6 by the same consideration as in the proof of Lemma 3.5. □

A Basic facts about the LSI

In this section we quote some basic facts about the LSI, that are needed in our arguments. For a general introduction to LSI we refer to [Led01, Roy99, GZ03]. There are several standard criteria for LSI. The Tensorization principle shows that LSI is compatible with products (cf. [Gro75]).

Theorem A.1 (Tensorization principle)**.** *Let μ_1 and μ_2 be probability measures on Euclidean spaces X_1 and X_2 respectively. If μ_1 and μ_2 satisfy LSI(ϱ_1) and LSI(ϱ_2) respectively, then the product measure $\mu_1 \otimes \mu_2$ satisfies LSI($\min\{\varrho_1, \varrho_2\}$).*

The next criterion contains how the LSI behaves under perturbations (cf. [HS87]). Note that it is not well suited for high dimensions.

Theorem A.2 (Criterion of Holley & Stroock)**.** *Let μ be a probability measure on the Euclidean space X and let $\delta\psi : X \to \mathbb{R}$ be a bounded function. Let the probability measure $\tilde{\mu}$ be defined as*

$$\tilde{\mu}(dx) = \frac{1}{Z} \exp\left(-\delta\psi(x)\right)\ \mu(dx).$$

Then $\tilde{\mu}$ satisfies LSI($\tilde{\varrho}$) with

$$\tilde{\varrho} = \varrho\ \exp\left(-\left(\sup \delta\psi - \inf \delta\psi\right)\right).$$

The criterion of Bakry & Émery connects the convexity of the Hamiltonian to the LSI constant (cf. [BE85, OV00]).

Theorem A.3 (Criterion of Bakry & Émery)**.** *Let X be a n-dimensional Euclidean space and let $H \in C^2(X)$. The probability measure μ on X is defined via*

$$\mu(dx) = \frac{1}{Z} \exp\left(-H(x)\right)\ dx.$$

If there is a constant $\varrho > 0$ such that for all $x, v \in X$

$$\langle v, \operatorname{Hess} H(x) v\rangle \geq \varrho |v|^2,$$

then μ satisfies LSI(ϱ).

More recently, Otto & Reznikoff [OR07] deduced a criterion that is capable to deal with certain non-convex Hamiltonians in high dimensions.

Theorem A.4 (Criterion of Otto & Reznikoff)**.** *Let $d\mu := \frac{1}{Z} \exp(-H(x))\ dx$ be a probability measure on a direct product of Euclidean spaces $X = X_1 \times \cdots \times X_M$. We assume that*

- *the conditional measures* $\mu(dx^l|x^n \in X_n\ n \neq l)$, $1 \leq l \neq n \leq M$, *satisfy a uniform LSI with constant* $\varrho_l > 0$,

- *the numbers* κ_{ln} *satisfy*

$$\kappa_{ln} := |\nabla_l \nabla_n H(x)| \leq C < \infty,$$

uniformly in $1 \leq l \neq n \leq M$*; here* $|\cdot|$ *denotes the operator norm of a bilinear form.*

- *the matrix* $A = (a_{ij})_{M\times M}$ *defined by*

$$a_{ij} = \begin{cases} \varrho_i & \text{if } i = j, \\ -\kappa_{ij} & \text{else}, \end{cases}$$

satisfies in the sense of quadratic forms

$$A \geq \varrho \, \mathrm{Id} \qquad \text{for a constant } \varrho > 0.$$

Then μ *satisfies LSI(*ϱ*).*

One can understand Theorem A.4 as a comparison principle. Via the matrix A, a Gaussian measure $\mu_A(dx) = \exp(-\langle x, Ax\rangle)\, dx$ is associated to the original Gibbs measure $\mu(dx) = \exp(-H(x))\, dx$. Because for Gaussian measures the property of positive definiteness of A and the LSI are equivalent (see for example [OR07]), the criterion of Otto & Reznikoff becomes:

Theorem. *If* μ_A *satisfies LSI(*ϱ*), then also* μ *does.*

Due to this example one could hope that μ inherits further features from μ_A. Theorem 2.9 shows that this happens for covariances (cf. [MO10b]). In the proof of the main result we also need the linearized version of the LSI, which is known as spectral gap inequality (SG).

Definition A.5. *A probability measure* μ *satisfies SG(*ϱ*),* $\varrho > 0$*, if for all functions* f

$$\mathrm{var}_\mu(f) := \int \left(f - \int f \, d\mu\right)^2 d\mu \leq \frac{1}{\varrho} \int |\nabla f|^2 \, d\mu.$$

We need the following well-known facts about SG.

Lemma A.6.

- *If* μ *satisfies LSI(*ϱ*), then* μ *also satisfies SG(*ϱ*).*

- *If* μ *satisfies SG(*ϱ*), then for all functions* f *and* g

$$\mathrm{cov}_\mu(f, g) \leq \frac{1}{\varrho} \left(\int |\nabla f|^2 \, d\mu\right)^{\frac{1}{2}} \left(\int |\nabla g|^2 \, d\mu\right)^{\frac{1}{2}}.$$

Acknowledgement

The author thanks Professor Felix Otto for bringing this topic into his view and for the helpful guidance. He thanks Maria Westdickenberg for the helpful comments. Finally, he also thanks the DFG for financial support through the Gottfried Wilhelm Leibniz program and the Bonn International Graduate School in Mathematics for partial support.

References

[BE85] D. Bakry and M. Émery. Diffusions hypercontractives. *Sem. Probab. XIX, Lecture Notes in Math, Springer-Verlag*, 1123:177–206, 1985.

[Cap03] P. Caputo. Uniform Poincaré inequalities for unbounded conservative spin systems: the non-interacting case. *Stochastic Process. Appl.*, 106(2):223–244, 2003.

[Cha03] D. Chafaï. Glauber versus Kawasaki for spectral gap and logarithmic Sobolev inequalities of some unbounded conservative spin systems. *Markov Process. Related Fields*, 9(3):341–362, 2003.

[CMR02] N. Cancrini, F. Martinelli, and C. Roberto. The logarithmic Sobolev constant of Kawasaki dynamics under a mixing condition revisited. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(4):385 – 436, 2002.

[Fel71] W. Feller. *An introduction to probability theory and its applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley and Sons, Inc. XXIV, 1971.

[GOVW09] N. Grunewald, F. Otto, C. Villani, and M. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. He. Poincaré Probab. Statist.*, 45(2):302–351, 2009.

[GPV88] M.Z. Guo, G.C. Papanicolau, and S.R.S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, 118:31–59, 1988.

[Gro75] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97:1061–1083, 1975.

[GZ03] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. 1801:1–134, 2003.

[HS87] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46:1159–1194, 1987.

[KL99] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems.* Grundlehren der Mathematischen Wissenschaften. 320. Berlin: Springer., 1999.

[Kos01] E. Kosygina. The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.*, 29(3):1086–110, 2001.

[Led01] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisted. *Sem. Probab. XXXV, Lecture Notes in Math, Springer- Verlag*, 1755:167–194, 2001.

[LPY02] C. Landim, G. Panizo, and H. T. Yau. Spectral gap and logarithmic Sobolev inequality for unbounded conservative spin systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 38(5):739–777, 2002.

[LY93] S. L. Lu and H-T. Yau. Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics. *Comm. Math. Phys.*, 156(2):399–433, 1993.

[MO10a] G. Menz and F. Otto. LSI for Kawasaki dynamics with quartic Hamiltonians. *in preparation*, 2010.

[MO10b] G. Menz and F. Otto. A new covariance estimate. *in preparation*, 2010.

[OR07] F. Otto and M. Reznikoff. A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.*, 243(1):121–157, 2007.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173:361–400, 2000.

[Roy99] G. Royer. Une initiation aux inégalités de Sobolev logarithmiques. *Cours Spéc., Soc. Math. France*, 1999.

[Shi96] A. N. Shiryaev. *Probability*. Springer-Verlag New York, 2nd, edition, 1996.

[SZ92a] D. Stroock and B. Zegarlinski. The equivalence of the logarithmic Sobolev inequality and the dobrushin-shlosman mixing condition. *Comm. Math. Phys.*, 144:303–323, 1992.

[SZ92b] D. Stroock and B. Zegarlinski. The logarithmic Sobolev inequality for discrete spin systems on the lattice. *Comm. Math. Phys.*, 149:175–193, 1992.

[SZ95] D. Stroock and B. Zegarlinski. On the ergodic properties of glauber dynamics. *J. Stat. Phys.*, 81:1007–1019, 1995.

[Yau91] H.-T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.*, 22:63–80, 1991.

[Yau96] H.-T. Yau. Logarithmic Sobolev inequality for lattice gases with mixing conditions. *Comm. Math. Phys.*, 181(2):367–408, 1996.

[Yos00] N. Yoshida. Application of log-Sobolev inequality to the stochastic dynamics of unbounded spin systems on the lattice. *J. Funct. Anal.*, 173:74–102, 2000.

[Yos01] N. Yoshida. The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):223–243, 2001.

[Zeg96] B. Zegarlinski. The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice. *Comm. Math. Phys.*, 175:401–432, 1996.