

Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig

Uniform logarithmic Sobolev inequalities for  
conservative spin systems with super-quadratic  
single-site potential.

by

*Georg Menz, and Felix Otto*

Preprint no.: 5

2011





# Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential.

Georg Menz & Felix Otto \*

February 7, 2011

We consider a non-interacting unbounded spin system with conservation of the mean spin. We derive a uniform logarithmic Sobolev inequality (LSI) provided the single-site potential is a bounded perturbation of a strictly convex function. The scaling of the LSI constant is optimal in the system size. The argument adapts the two-scale approach of Grunewald, Otto, Westdickenberg, and Villani from the quadratic to the general case. Using an asymmetric Brascamp-Lieb type inequality for covariances we reduce the task of deriving a uniform LSI to the convexification of the coarse-grained Hamiltonian, which follows from a general local Cramèr theorem.

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

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

## 1. Introduction and main result

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

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

---

\*Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany.  
Georg.Menz@mis.mpg.de, Felix.Otto@mis.mpg.de.

Throughout the article,  $Z$  denotes a generic normalization constant. The value of  $Z$  may change from line to line or even within a line. The non-interacting Hamiltonian  $H : \mathbb{R}^N \rightarrow \mathbb{R}$  is given by a sum of single-site potentials  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  that are specified later i.e.

$$H(x) := \sum_{i=1}^N \psi(x_i). \quad (1)$$

For a real number  $m$ , we consider the  $N - 1$  dimensional hyper-plane  $X_{N,m}$  given by

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

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

$$\langle x, \tilde{x} \rangle := \sum_{i=1}^N x_i \tilde{x}_i.$$

The restriction of  $\mu$  to  $X_{N,m}$  is called canonical ensemble  $\mu_{N,m}$  i.e.

$$\mu_{N,m}(dx) := \frac{1}{Z} \exp(-H(x)) \mathcal{H}_{|X_{N,m}}^{N-1}(dx). \quad (2)$$

Here,  $\mathcal{H}_{|X_{N,m}}^{N-1}$  denotes the  $N - 1$  dimensional Hausdorff measure restricted to the hyperplane  $X_{N,m}$ . For convenience, we introduce the notation

$$\begin{aligned} a \lesssim b &\Leftrightarrow \text{there is a uniform constant } C > 0 \text{ such that } a \leq Cb, \\ a \sim b &\Leftrightarrow \text{it holds that } a \lesssim b \text{ and } b \lesssim a. \end{aligned}$$

In 1993, Varadhan [Var93] posed the question for which kind of single-site potential  $\psi$  the canonical ensemble  $\mu_{N,m}$  satisfies a spectral gap inequality (SG) uniformly in the system size  $N$  and the mean spin  $m$ . A partial answer was given by Caputo [Cap03]:

**Theorem 1.1** (Caputo). *Assume that for the single-site potential  $\psi$  exist a splitting  $\psi = \psi_c + \delta\psi$  and constants  $\beta_-, \beta_+ \in [0, \infty)$  such that for all  $x \in [0, \infty)$*

$$\psi_c''(x) \sim |x|^{\beta_+} + 1, \quad \psi_c''(-x) \sim |x|^{\beta_-} + 1, \quad \text{and} \quad |\delta\psi| + |\delta\psi'| + |\delta\psi''| \lesssim 1. \quad (3)$$

*Then the canonical ensemble  $\mu_{N,m}$  satisfies the SG with constant  $\varrho > 0$  uniformly in the system size  $N$  and the mean spin  $m$ . More precisely, for any function  $f$*

$$\text{var}_{\mu_{N,m}}(f) = \int \left( f - \int f d\mu_{N,m} \right)^2 d\mu_{N,m} \leq \frac{1}{\varrho} \int |\nabla f|^2 d\mu_{N,m}.$$

*Here,  $\nabla$  denotes the gradient determined by the Euclidean structure of  $X_{N,m}$ .*

In this article, we give a full answer to the question by Varadhan [Var93] and also consider the question if the statement of the last theorem can be strengthened to the logarithmic Sobolev inequality (LSI).

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

$$\int f \log f \, d\mu - \int f \, d\mu \log \left( \int f \, d\mu \right) \leq \frac{1}{2\varrho} \int \frac{|\nabla f|^2}{f} \, d\mu. \quad (4)$$

Here,  $\nabla$  denotes the gradient determined by the Euclidean structure of  $X$ .

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

$$|\nabla f|^2 = \sum_{i=1}^N \left| \frac{d}{dx_i} f \right|^2.$$

The LSI was originally introduced by Gross [Gro75]. It yields the SG and can be used as a powerful tool for studying spin systems. Like the SG, the LSI implies exponential convergence to equilibrium of the naturally associated conservative diffusion process. The rate of convergence is given by the LSI constant  $\varrho$  (cf. [SZ92a, SZ92b, SZ95, Yos99, Zeg96] and Remark 1.7). Therefore, an appropriate scaling of the LSI constant in the system size indicates the absence of phase transitions. The SG yields convergence in the sense of variances in contrast to the LSI, which yields convergence in the sense of relative entropies. The SG and the LSI are also useful for deducing the hydrodynamic limit (see [Var93, LY93, Kos01] for the SG and [GOVW09] for the LSI).

We consider three cases of different potentials: sub-quadratic, quadratic, and super-quadratic single-site potentials. In the case of sub-quadratic single-site potentials, Barthe & Wolff [BW09] gave a counterexample where the scaling in the system size of the SG and the LSI constant of the canonical ensemble differs in the system size. More precisely, they showed:

**Theorem 1.4 (Barthe & Wolff).** *Assume that the single-site potential  $\psi$  is given by*

$$\psi(x) = \begin{cases} x, & \text{for } x > 0, \\ \infty, & \text{else.} \end{cases}$$

*Then the SG constant  $\varrho_1$  and the LSI constant  $\varrho_2$  of the canonical ensemble  $\mu_{N,m}$  satisfy*

$$\varrho_1 \sim \frac{1}{m^2} \quad \text{and} \quad \varrho_2 \sim \frac{1}{Nm^2}.$$

In the case of perturbed quadratic single-site potentials it is known that Theorem 1.1 can be improved to the LSI. More precisely, several authors (cf. [LY93, LPY02, Cha03, GOVW09]) deduced the following statement by different methods:

**Theorem 1.5** (Landim, Panizo, and Yau). *Assume that the single-site potential  $\psi$  is perturbed quadratic in the following sense: There exists a splitting  $\psi = \psi_c + \delta\psi$  such that*

$$\psi_c'' = 1 \quad \text{and} \quad |\delta\psi| + |\delta\psi'| + |\delta\psi''| \lesssim 1. \quad (5)$$

*Then the canonical ensemble  $\mu_{N,m}$  satisfies the LSI with constant  $\varrho > 0$  uniformly in the system size  $N$  and the mean spin  $m$ .*

There is only left to consider the super-quadratic case. It is conjectured that the optimal scaling LSI also holds, if the single-site potential  $\psi$  is a bounded perturbation of a strictly convex function (cf. [LPY02, p. 741], [Cha03, Theorem 0.3 f.], and [Cap03, p. 226]). Heuristically, this conjecture seems reasonable: Because the LSI is closely linked to convexity (consider for example the Bakry-Émery criterion), a perturbed strictly convex potential should behave no worse than a perturbed quadratic one. However technically, the methods for the quadratic case are not able to handle the perturbed strictly convex case, because they require an upper bound on the second derivative of the Hamiltonian. In the main result of the article we show that the conjecture from above is true:

**Theorem 1.6.** *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense that there is a splitting  $\psi = \psi_c + \delta\psi$  such that*

$$\psi_c'' \gtrsim 1 \quad \text{and} \quad |\delta\psi| + |\delta\psi'| \lesssim 1. \quad (6)$$

*Then the canonical ensemble  $\mu_{N,m}$  satisfies the LSI with constant  $\varrho > 0$  uniformly in the system size  $N$  and the mean spin  $m$ .*

**Remark 1.7** (From Glauber to Kawasaki). *The bound on the r.h.s. 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. [GOVW09, Remark 15] or [Cap03]) in the sense that one endows  $X_{N,m}$  with the Euclidean structure coming from the discrete  $H^{-1}$ -norm. More precisely, if  $\Lambda$  is a cubic lattice in any dimension of width  $L$ , then Theorem 1.6 yields the LSI for Kawasaki dynamics with constant  $L^{-2}\varrho$ , which is the optimal scaling in  $L$  (cf. [Yau96]).*

Note that the standard criteria for the SG and the LSI (cf. Appendix A) fail for the canonical ensemble  $\mu_{N,m}$ :

- The **Tensorization principle** for the SG and the LSI does not apply because of the restriction to the hyper-plane  $X_{N,m}$  (cf. for example [GZ03, Theorem 4.4] or Theorem A.1).

- The **Bakry-Émery** criterion does not apply because the Hamiltonian  $H$  is not strictly convex (cf. [BE85, Proposition 3 and Corollary 2] or Theorem A.3).
- The **Holley-Stroock** criterion does not help because the LSI constant  $\varrho$  has to be independent of the system size  $N$  (cf. [HS87, p. 1184] or Theorem A.2).

Therefore, a more elaborated machinery was needed for the proof of Theorem 1.1 and Theorem 1.5. The approach of Caputo to Theorem 1.1 seems to be restricted to the SG, because it relies on the spectral nature of the SG. The most common approach for the proof of Theorem 1.5 is the Lu-Yau martingale method (see [LY93, LPY02, Cha03]). Recently, Grunewald, Otto, Villani, and Westdickenberg [GOVW09] provided a new technique for deducing Theorem 1.5 called the two-scale approach. We follow this approach in the proof of Theorem 1.6.

The limiting factor for extending Theorem 1.5 to more general single-site potentials is almost the same for the Lu-Yau martingale method and for the two-scale approach: It is the estimation of a covariance term w.r.t. the measure  $\mu_{N,m}$  conditioned on a special event (cf. [LPY02, (4.6)] and [GOVW09, (42)]). In the two-scale approach one has to estimate for some large but fixed  $K \gg 1$  and any non-negative function  $f$  the covariance

$$\left| \text{cov}_{\mu_{K,m}} \left( f, \frac{1}{K} \sum_{i=1}^K \psi'(x_i) \right) \right|.$$

In [GOVW09] this term was estimated by using a standard estimate (cf. Lemma 2.9 and [GOVW09] [Lemma 22]), which only can be applied for perturbed quadratic single-site potentials  $\psi$ . We get around this difficulty by making the following adaptations: Instead of one-time coarse-graining of big blocks we consider iterative coarse-graining of pairs. As a consequence we only have to estimate the covariance term from above in the case  $K = 2$ . Because  $\mu_{2,m}$  is a one-dimensional measure, we are able to apply the more robust asymmetric Brascamp-Lieb inequality (cf. Lemma 2.10), which can also be applied for perturbed strictly convex single-site potentials  $\psi$ .

Recently, the optimal scaling LSI was established in [Men10] for a weakly interacting Hamiltonian with perturbed quadratic single-site potentials  $\psi$  i.e.

$$H(x) = \sum_{i=1}^N \psi(x_i) + \varepsilon \sum_{1 \leq i < j \leq N} b_{ij} x_i x_j.$$

Because the original two-scale approach was used, it is an interesting question if one could extend this result to perturbed strictly convex single-site potentials. A direct transfer of the argument of [Men10] fails because of the iterative structure of the proof of Theorem 1.6.

The remaining part of this article is organized as follows. In Section 2.1 we prove of the main result. The auxiliary results of Section 2.1 are proved in Section 2.2. There is one exception: The convexification of the single-site potential by iterated renormalization (see Theorem 2.6) is proved in Section 3. In the short Appendix A we state the standard criteria for the SG and the LSI.

## 2. Adapted two-scale approach

### 2.1. Proof of the main result

In this section we state the proof of Theorem 1.6, which is based based on an adaptation of the two-scale approach of [GOVW09]. We start with introducing the concept of coarse-graining of pairs. We recommend to read [GOVW09, Chapter 2.1] as a guideline. We assume that the number  $N$  of sites is given by  $N = 2^K$  for some large number  $K \in \mathbb{N}$ . The step to arbitrary  $N$  is not difficult (cf. Remark 2.7 below).

We decompose the spin system into blocks each containing two spins. The coarse-graining operator  $P : X_{N,m} \rightarrow X_{\frac{N}{2},m}$  assigns to each block the mean spin of the block. More precisely,  $P$  is given by

$$P(x) := \left( \frac{1}{2}(x_1 + x_2), \frac{1}{2}(x_3 + x_4), \dots, \frac{1}{2}(x_{N-1} + x_N) \right). \quad (7)$$

Due to the coarse-graining operator  $P$  we can decompose the canonical ensemble  $\mu_{N,m}$  into

$$\mu_{N,m}(dx) = \mu(dx|y)\bar{\mu}(dy), \quad (8)$$

where  $\bar{\mu} := P_{\#}\mu_{N,m}$  denotes the push forward of the Gibbs measure  $\mu$  under  $P$  and  $\mu(dx|y)$  is the conditional measure of  $x$  given  $Px = y$ . The last equation has to be understood in a weak sense i.e. for any test function  $\xi$

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

Now, we are able to state the first ingredient of the proof of Theorem 1.6.

**Proposition 2.1** (Hierarchic criterion for the LSI). *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). If the marginal  $\bar{\mu}$  satisfies the LSI with constant  $\varrho_1 > 0$  uniformly in the system size  $N$  and the mean spin  $m$ , then the canonical ensemble  $\mu_{N,m}$  also satisfies the LSI with constant  $\varrho_2 > 0$  uniformly in the system size  $N$  and the mean spin  $m$ .*

The proof of this statement is given in Section 2.2. Due to the last proposition it suffices to deduce the LSI for the marginal  $\bar{\mu}$ . Hence, let us have a closer look at the structure of  $\bar{\mu}$ . We will characterize the Hamiltonian of the marginal  $\bar{\mu}$  with the help of the renormalization operator  $\mathcal{R}$ , which is introduced as follows.

**Definition 2.2.** Let  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  be a single-site potential. Then the renormalized single-site potential  $\mathcal{R}\psi : \mathbb{R} \rightarrow \mathbb{R}$  is defined by

$$\mathcal{R}\psi(y) := -\log \int \exp(-\psi(x+y) - \psi(-x+y)) dx. \quad (9)$$

**Remark 2.3.** The renormalized single-site potential  $\mathcal{R}\psi$  can be interpreted in the following way: A change of variables (cf. [EG92, Section 3.3.3]) and the invariance of the Hausdorff measure under translation yield the identity

$$\begin{aligned} \exp(-\mathcal{R}\psi(y)) &= \int \exp(-\psi(x+m) - \psi(-x+m)) dx \\ &= \frac{1}{\sqrt{2}} \int \exp(-\psi(x_1) - \psi(x_2)) \mathcal{H}_{\{x_1+x_2=2y\}}^1(dx). \end{aligned}$$

Therefore, the renormalized single-site potential  $\mathcal{R}\psi$  describes the free energy of two independent spins  $X_1$  and  $X_2$  (identically distributed according to  $Z^{-1} \exp(-\psi)$ ) conditioned on a fixed mean value  $\frac{1}{2}(X_1 + X_2) = y$ .

**Lemma 2.4** (Invariance under renormalization). Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). Then the renormalized Hamiltonian  $\mathcal{R}\psi$  is also perturbed strictly convex in the sense of (6).

Direct calculation using the coarea formula (cf. [EG92, Section 3.4.2]) reveals the following structure of the marginal  $\bar{\mu}$ .

**Lemma 2.5.** The marginal  $\bar{\mu}$  is given by

$$\bar{\mu}(dy) := \frac{1}{Z} \exp\left(-\sum_{i=1}^{\frac{N}{2}} \mathcal{R}\psi(y_i)\right) \mathcal{H}_{[X_{\frac{N}{2},m}]}^{\frac{N}{2}-1}(dy).$$

It follows from the last two lemmas that the marginal  $\bar{\mu}$  has the same structure as the canonical ensemble  $\mu_{N,m}$ . The single-site potential of  $\bar{\mu}$  is given by the renormalized single-site potential  $\mathcal{R}\psi$ . Hence, one can iterate the coarse-graining of pairs. The next statement shows that after finitely many iterations the renormalized single-site potential  $\mathcal{R}^M\psi$  becomes uniformly strictly convex. Therefore, the Bakry-Émery criterion (cf. Theorem A.3) yields that the corresponding marginal satisfies the LSI with constant  $\tilde{\varrho} > 0$ , uniformly in the system size  $N$  and the mean spin  $m$ . Then an iterated application of the hierarchic criterion of the LSI (cf. Proposition 2.1) yields Theorem 1.6 and the proof of the main result for  $N = 2^K$  is finished.

**Theorem 2.6** (Convexification by renormalization). *Let  $\psi$  be a perturbed strictly convex single-site potential in the sense of (6). Then there is an integer  $M_0$  such that for all  $M \geq M_0$  the  $M$ -times renormalized single-site potential  $\mathcal{R}^M \psi$  is uniformly strictly convex independently of the system size  $N$  and the mean spin  $m$ .*

We conclude this section by giving some remarks and pointing out the central tools needed for the proof of the auxiliary results. The next remark shows how Theorem 1.6 is proved in the case of an arbitrary number  $N$  of sites.

**Remark 2.7.** *Note that an arbitrary number of sites  $N$  can be written as*

$$N = \tilde{K}2^K + R$$

*for some number  $\tilde{K}$ , a large but fixed number  $K$ , and a bounded number  $R < 2^K$ . Hence, one can decompose the spin system into  $\tilde{K}$  blocks of  $2^K$  spins and one block of  $R$  spins. The big blocks of  $2^K$  spins are coarse-grained by pairs, whereas the small block of  $R$  spins is not coarse-grained at all. After iterating this procedure sufficiently often, the renormalized single-site potentials of the big blocks are uniformly strictly convex. On the remaining block of  $R$  spins, the corresponding single-site potentials are unchanged. Because  $\psi$  is a bounded perturbation of a strictly convex function, it follows from a combination of the Bakry-Émery criterion (cf. Theorem A.3) and the Holley-Stroock criterion (cf. Theorem A.2) that the marginal of the whole system satisfies the LSI with constant*

$$\varrho \gtrsim \exp(-R(\sup \delta\psi - \inf \delta\psi)),$$

*which is independent on  $N$  and  $m$ . Therefore, an iterated application of the hierarchic criterion of the LSI (cf. Proposition 2.1) yields Theorem 1.6.*

The proof of Proposition 2.1 and Lemma 2.4 is given in Section 2.2, whereas the proof of Theorem 2.6 is stated in Section 3.

Starting point for the proof of Theorem 2.6 is the observation that the  $M$ -times renormalized single-site potential  $\mathcal{R}^M \psi$  corresponds to the coarse-grained Hamiltonian related to coarse-graining with block size  $2^M$  (cf. [GOVW09]).

**Lemma 2.8.** *For  $K \in \mathbb{N}$  let the coarse-grained Hamiltonian  $\bar{H}_K$  be defined by*

$$\bar{H}_K(m) = -\frac{1}{K} \log \int \exp(-H(x)) \mathcal{H}_{[X_K, m]}^{K-1}(dx). \quad (10)$$

*Let  $M \in \mathbb{N}$ . Then there is a constant  $0 < C(2^M) < \infty$  depending only on  $2^M$  such that*

$$\mathcal{R}^M \psi = 2^M \bar{H}_{2^M} + C(2^M).$$

Because the last statement is verified by a straight-forward application of the area and coarea formula, we omit the proof. In Lemma 2.8 one could easily determine the exact value of the constant  $C(2^M)$ . However, the exact value is not important because we are only interested in the convexity of  $\mathcal{R}^M\psi$ . In [GOVW09], the convexification of  $\bar{H}_K$  was deduced from a local Cramér theorem (cf. [GOVW09][Proposition 31]). For the proof of Theorem 2.6 we follow the same strategy generalizing the argument to perturbed strictly convex single-site potentials  $\psi$ .

Now, we make some comments on the proof of Proposition 2.1 and Lemma 2.4. One of the limiting factors in the proof of Theorem 1.5 is the application of a classical covariance estimate (cf. [GOVW09][Lemma 22]). In our framework this estimate can be formulated as:

**Lemma 2.9.** *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). Let  $\nu$  be a probability measure on  $\mathbb{R}$  given by*

$$\nu(dx) = \frac{1}{Z} \exp(-\psi(x)) dx.$$

*Then for any function  $f \geq 0$  and  $g$*

$$|\text{cov}_\nu(f, g)| \lesssim \sup_x |g'(x)| \left( \int f d\nu \right)^{\frac{1}{2}} \left( \int \frac{|f'|^2}{f} d\nu \right)^{\frac{1}{2}}.$$

In [GOVW09], the last estimate was applied to the function  $g = \psi'$ . Note that  $|g'(x)| = |\psi''(x)|$  is only bounded in the case of a perturbed quadratic single-site potential  $\psi$ . The main new ingredient for the proof of the hierarchic criterion for the LSI (cf. Proposition 2.1) and the invariance principle (cf. Lemma 2.4) is an asymmetric Brascamp-Lieb inequality, which does not exhibit this restriction.

**Lemma 2.10.** *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). Let  $\nu$  be a probability measure on  $\mathbb{R}$  given by*

$$\nu(dx) = \frac{1}{Z} \exp(-\psi(x)) dx.$$

*Then for any function  $f$  and  $g$*

$$|\text{cov}_\nu(f, g)| \leq \exp(-3 \text{osc } \delta\psi) \sup_x \left| \frac{g'(x)}{\psi_c''(x)} \right| \int |f'| d\nu,$$

*where  $\text{osc } \delta\psi := \sup_x \delta\psi(x) - \inf_x \delta\psi(x)$ .*

We call the last inequality asymmetric, because compared to the original Brascamp-Lieb inequality [BL76]  $L^2 \times L^2$  is replaced by  $L^1 \times L^\infty$  and the factor  $\frac{1}{\sqrt{\psi_c''}}$  is not evenly distributed. It is an interesting question if an analog statement also holds for higher dimensions. The proof of Lemma 2.10 is based on a kernel representation of the covariance. All steps are elementary.

*Proof of Lemma 2.10.* Let us consider a Gibbs measure  $\mu$  associated to the Hamiltonian  $H : \mathbb{R} \rightarrow \mathbb{R}$ . More precisely,  $\mu$  is given by

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

We start by deriving the following integral representation of the covariance of  $\mu$ :

$$\text{cov}_\mu(f, g) = \int \int f'(x) K_\mu(x, y) g'(y) dx dy, \quad (11)$$

where the non-negative kernel  $K_\mu(x, y)$  is given by

$$K_\mu(x, y) := \begin{cases} M_\mu(x)(1 - M_\mu)(y) & \text{for } y \geq x \\ (1 - M_\mu)(x)M_\mu(y) & \text{for } y \leq x \end{cases},$$

and  $M_\mu(x) := \mu((-\infty, x))$  so that  $(1 - M_\mu)(x) = \mu((x, \infty))$ . Indeed, we start by noting that

$$\text{cov}_\mu(f, g) = \int \int (f(z) - f(x))\mu(x) dx \int (g(z) - g(y))\mu(y) dy \mu(z) dz, \quad (12)$$

where we do not distinguish between the measure  $\mu(dx)$  and its Lebesgue density  $\mu(x)$  in our notation. Using  $M'_\mu(x) = \mu(x)$ , we can use integration by parts to rewrite each factor in terms of the derivative:

$$\begin{aligned} & \int (f(z) - f(x))\mu(x) dx \\ &= \int_{-\infty}^z (f(z) - f(x))M'_\mu(x) dx - \int_z^\infty (f(z) - f(x))(1 - M_\mu)'(x) dx \\ &= \int_{-\infty}^z f'(x)M_\mu(x) dx - \int_z^\infty f'(x)(1 - M_\mu)(x) dx \\ &= \int f'(x)(I(x < z)M_\mu(x) - I(x > z)(1 - M_\mu)(x)) dx, \end{aligned}$$

where  $I(x < z)$  assumes the value 1 if  $x < z$  and zero otherwise. Inserting this and the corresponding identity for  $g(y)$  into (12), we obtain

$$\begin{aligned} & \text{cov}_\mu(f, g) \\ &= \int \int f'(x)(I(x < z)M_\mu(x) - I(x > z)(1 - M_\mu)(x)) dx \\ & \quad \times \int g'(y)(I(y < z)M_\mu(y) - I(y > z)(1 - M_\mu)(y)) dy \mu(z) dz \\ &= \int \int f'(x)K_\mu(x, y)g'(y) dx dy \end{aligned} \quad (13)$$

with kernel  $K_\mu(x, y)$  as desired given by

$$\begin{aligned}
K_\mu(x, y) &= M_\mu(x)M_\mu(y) \int I(x < z)I(y < z)\mu(z) dz \\
&\quad - M_\mu(x)(1 - M_\mu(y)) \int I(x < z)I(y > z)\mu(z) dz \\
&\quad - (1 - M_\mu(x))M_\mu(y) \int I(x > z)I(y < z)\mu(z) dz \\
&\quad + (1 - M_\mu(x))(1 - M_\mu(y)) \int I(x > z)I(y > z)\mu(z) dz \\
&= M_\mu(x)M_\mu(y)(1 - M_\mu(\max\{x, y\})) \\
&\quad - M_\mu(x)(1 - M_\mu(y))I(y > x)(M_\mu(y) - M_\mu(x)) \\
&\quad - (1 - M_\mu(x))M_\mu(y)I(y < x)(M_\mu(x) - M_\mu(y)) \\
&\quad + (1 - M_\mu(x))(1 - M_\mu(y))M_\mu(\min\{x, y\}) \\
&= I(y > x)(M_\mu(x)M_\mu(y)(1 - M_\mu(y)) - M_\mu(x)(1 - M_\mu(y))(M_\mu(y) - M_\mu(x)) \\
&\quad + (1 - M_\mu(x))(1 - M_\mu(y))M_\mu(x)) \\
&\quad + I(y \leq x)(M_\mu(x)M_\mu(y)(1 - M_\mu(x)) - (1 - M_\mu(x))M_\mu(y)(M_\mu(x) - M_\mu(y)) \\
&\quad + (1 - M_\mu(x))(1 - M_\mu(y))M_\mu(y)) \\
&= I(y > x)M_\mu(x)(1 - M_\mu(y)) + I(y \leq x)(1 - M_\mu(x))M_\mu(y).
\end{aligned}$$

We now establish the following identity for the above kernel:

$$\int K_\mu(x, y)H''(y)dy = \mu(x). \tag{14}$$

Indeed, we have by integrations by part

$$\begin{aligned}
&\int K_\mu(x, y)H''(y) dy \\
&= (1 - M_\mu(x)) \int_{-\infty}^x M_\mu(y)H''(y) dy + M_\mu(x) \int_x^\infty (1 - M_\mu(y))H''(y) dy \\
&= (1 - M_\mu(x)) \left( M_\mu(x)H'(x) - \int_{-\infty}^x M'_\mu(y)H'(y) dy \right) \\
&\quad + M_\mu(x) \left( -(1 - M_\mu(x))H'(x) + \int_x^\infty M'_\mu(y)H'(y) dy \right) \\
&= -(1 - M_\mu(x)) \int_{-\infty}^x \exp(-H(y))H'(y) dy \\
&\quad + M_\mu(x) \int_x^\infty \exp(-H(y))H'(y) dy \\
&= (1 - M_\mu(x))\mu(x) + M_\mu(x)\mu(x) = \mu(x).
\end{aligned}$$

Let us now consider the Gibbs measures  $\nu(dx)$  and  $\nu_c(dx)$  given by

$$\nu(dx) = \frac{1}{Z} \exp(-\psi_c(x) - \delta\psi(x)) dx \quad \text{and} \quad \nu_c(dx) = \frac{1}{Z} \exp(-\psi_c(x)) dx.$$

By the integral representation (11) of the covariance we have the estimate

$$|\text{cov}_\nu(f, g)| \leq \int \int |f'(x)| K_\nu(x, y) |g'(y)| dx dy.$$

By a straight-forward calculation we can estimate

$$\begin{aligned} M_\nu(x) &= \frac{\int_{-\infty}^x \exp(-\psi_c(x) - \delta\psi(x)) dx}{\int \exp(-\psi_c(x) - \delta\psi(x)) dx} \\ &\leq \exp(-\text{osc } \delta\psi) \frac{\int_{-\infty}^x \exp(-\psi_c(x)) dx}{\int \exp(-\psi_c(x)) dx} \\ &= \exp(-\text{osc } \delta\psi) M_{\nu_c}(x). \end{aligned}$$

Together with a similar estimate for  $(1 - M_\nu(y))$  this yields the kernel estimate

$$K_\nu(x, y) \leq \exp(-2 \text{osc } \delta\psi) K_{\nu_c}(x, y).$$

Applying this to the covariance estimate from above yields

$$|\text{cov}_\nu(f, g)| \leq \exp(-2 \text{osc } \delta\psi) \int \int |f'(x)| K_{\nu_c}(x, y) |g'(y)| dx dy.$$

Using the identity (14) for  $\mu = \nu_c$  we may easily conclude:

$$\begin{aligned} |\text{cov}_\nu(f, g)| &\leq \exp(-2 \text{osc } \delta\psi) \sup_y \frac{|g'(y)|}{\psi_c''(y)} \int |f'(x)| \int K_{\nu_c}(x, y) \psi_c''(y) dy dx \\ &= \exp(-2 \text{osc } \delta\psi) \sup_y \frac{|g'(y)|}{\psi_c''(y)} \int |f'(x)| \nu_c(dx) \\ &\leq \exp(-3 \text{osc } \delta\psi) \sup_y \frac{|g'(y)|}{\psi_c''(y)} \int |f'(x)| \nu(dx). \end{aligned}$$

□

For the entertainment of the reader, let us argue how the identity (14) also yields the traditional Brascamp-Lieb inequality in the case  $H'' > 0$ . Indeed, by the symmetry of the kernel  $K_\mu(x, y)$  the identity (14) yields for all  $x$  and  $y$

$$\int K_\mu(x, y) H''(y) dy = \mu(x) \quad \text{and} \quad \int K_\mu(x, y) H''(x) dx = \mu(y). \quad (15)$$

The integral representation of the covariance (11) yields

$$\begin{aligned} \text{var}_\mu(f) &= \int \int f'(x) K_\mu(x, y) f'(y) dx dy \\ &= \int \int f'(x) \left( \frac{K_\mu(x, y) H''(y)}{H''(x)} \right)^{\frac{1}{2}} f'(y) \left( \frac{K_\mu(x, y) H''(x)}{H''(y)} \right)^{\frac{1}{2}} dx dy. \end{aligned}$$

Then a combination of Hoelder's inequality and the identity (15) for the kernel  $K_\mu(x, y)$  yields the Brascamp-Lieb inequality:

$$\begin{aligned}
& \text{var}_\mu(f) \\
& \leq \left( \int \int \frac{|f'(x)|^2}{H''(x)} K_\mu(x, y) H''(y) dy dx \right)^{\frac{1}{2}} \left( \int \int \frac{|f'(y)|^2}{H''(y)} K_\mu(x, y) H''(x) dx dy \right)^{\frac{1}{2}} \\
& = \left( \int \frac{|f'(x)|^2}{H''(x)} \mu(x) dx \right)^{\frac{1}{2}} \left( \int \frac{|f'(y)|^2}{H''(y)} \mu(y) dy \right)^{\frac{1}{2}} \\
& = \int \frac{|f'(x)|^2}{H''(x)} \mu(x) dx. \tag{16}
\end{aligned}$$

## 2.2. Proof of auxiliary results

In this section we outline the proof of Proposition 2.1 and Lemma 2.4. We start with Proposition 2.1, which is the hierarchic criterion for the LSI. Unfortunately, we cannot directly apply the two-scale criterion of [GOVW09][Theorem 3]. The reason is that the number

$$\kappa := \{ \langle \text{Hess } H(x)u, v \rangle, u \in \text{im}(2P^t P), v \in \text{im}(\text{id}_X - 2P^t P); |u| = |v| = 1 \}, \tag{17}$$

which measures the interaction between the microscopic and macroscopic scales, can be infinite for a perturbed strictly convex single-site potential  $\psi$ . However, we follow the proof of [GOVW09][Theorem 3] with only one major difference: Instead of applying the classical covariance estimate (cf. Lemma 2.9), we apply the asymmetric Brascamp-Lieb inequality (cf. Lemma 2.10). Let us assume for the rest of this section that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6).

For convenience we set  $X := X_{N,m}$  and  $Y := X_{\frac{N}{2},m}$ . We choose on  $X$  and  $Y$  the standard Euclidean structure given by

$$\langle x, y \rangle = \sum_{i=1}^N x_i y_i.$$

The coarse-graining operator  $P : X \rightarrow Y$  given by (7) satisfies the identity

$$2PP^t = \text{id}_Y,$$

where  $P^t : Y \rightarrow X$  is the adjoint operator of  $P$ . Note that our  $P^t$  differs from the  $P^t$  of [GOVW09], because the Euclidean structure on  $Y$  differs from the Euclidean structure used in [GOVW09]. The last identity yields that  $2P^t P$  is the orthogonal projection of

$X$  to  $\text{im } P^t$ . Hence, one can decompose  $X$  into the orthogonal sum of *microscopic fluctuations* and *macroscopic variables* according to

$$\begin{aligned} X &= \ker P \oplus \text{im } P^t \quad \text{and} \\ x &= (\text{id}_X - 2P^t P) x + 2P^t P x. \end{aligned}$$

We apply this decomposition to the gradient  $\nabla f$  of a smooth function  $f$  on  $X$ . The gradient  $\nabla f$  is decomposed into a macroscopic gradient and a fluctuation gradient satisfying

$$\begin{aligned} \nabla f(x) &= (\text{id}_X - 2P^t P) \nabla f(x) + 2P^t P \nabla f(x) \quad \text{and} \\ |\nabla f(x)|^2 &= |(\text{id}_X - 2P^t P) \nabla f(x)|^2 + |2P^t P \nabla f(x)|^2. \end{aligned} \quad (18)$$

Note that  $\ker P$  is the tangent space of the fiber  $\{Px = y\}$ . Hence, the gradient of  $f$  on  $\{Px = y\}$  is given by  $(\text{id}_X - 2P^t P) \nabla f(x)$ . The first main ingredient of the proof of Proposition 2.1 is the following statement.

**Lemma 2.11.** *The conditional measure  $\mu(dx|y)$  given by (8) satisfies the LSI with constant  $\varrho > 0$  uniformly in the system size  $N$ , the macroscopic profile  $y$ , and the mean spin  $m$ . More precisely, for any non-negative function  $f$*

$$\begin{aligned} \int f \log f \mu(dx|y) - \int f \mu(dx|y) \log \left( \int f \mu(dx|y) \right) \\ \leq \frac{1}{2\varrho} \int \frac{|(\text{id}_X - 2P^t P) \nabla f|^2}{f} \mu(dx|y). \end{aligned}$$

*Proof of Lemma 2.11.* Observe that the conditional measures  $\mu(dx|y)$  have a product structure: We decompose  $\{Px = y\}$  into a product of Euclidean spaces. Namely for

$$X_{2,y_i} := \left\{ (x_{2i-1}, x_{2i}), x_{2i-1} + x_{2i} = 2y_i \right\}, \quad i \in \left\{ 1, \dots, \frac{N}{2} \right\}$$

we have

$$\{Px = y\} = X_{2,y_1} \times \cdots \times X_{2,y_{\frac{N}{2}}}.$$

It follows from the coarea formula (cf. [EG92, Section 3.4.2]) that

$$\begin{aligned} \int_{\{Px=y\}} f(x) \mu(dx|y) \\ = \int f(x) \bigotimes_{i=1}^{\frac{N}{2}} \frac{1}{Z} \exp(-\psi(x_{2i-1}) - \psi(x_{2i})) \mathcal{H}_{[X_{2,y_i}]}^1(dx_{2i-1}, dx_{2i}). \end{aligned}$$

Hence  $\mu(dx|y)$  is the product measure

$$\mu(dx|y) = \bigotimes_{i=1}^{\frac{N}{2}} \mu_{2,y_i}(dx_{2i-1}, dx_{2i}), \quad (19)$$

where we make use of the notation introduced in (2). Because the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6), a combination of the Bakry-Émery criterion (cf. Theorem A.3) and the Holley-Stroock criterion (cf. Theorem A.2) yield that the measure  $\mu_{2,m}(dx_1, dx_2)$  satisfies the LSI with constant  $\varrho > 0$  uniformly in  $m$ . Then the tensorization principle (cf. Theorem A.1) implies the desired statement.  $\square$

For convenience, let us introduce the following notation: Let  $f$  be an arbitrary function. Then its conditional expectation  $\bar{f}$  is defined by

$$\bar{f}(y) := \int f(x) \mu(dx|y).$$

The second main ingredient of the proof of Proposition 2.1 is the following proposition, which is the analogue statement of [GOVW09, Proposition 20].

**Proposition 2.12.** *Assume that the marginal  $\bar{\mu}(dy)$  given by (8) satisfies the LSI with constant  $\lambda > 0$  uniformly in the system size  $N$  and the mean spin  $m$ . Then for any non-negative function  $f$*

$$\frac{|\nabla \bar{f}(y)|^2}{\bar{f}(y)} \lesssim \int \frac{|\nabla f(x)|^2}{f(x)} \mu(dx|y),$$

*uniformly in the macroscopic profile  $y$  and the system size  $N$ .*

Before we verify Proposition 2.12, let us show how it can be used in the proof of Proposition 2.1.

*Proof of Proposition 2.1.* Under the assumption that Lemma 2.11 and Proposition 2.12 hold, the argument is exactly the same as in the proof of [GOVW09, Theorem 3]: Let  $\phi$  denote the function

$$\phi(x) := x \log x.$$

First, the additive property of the entropy implies

$$\begin{aligned} \int \phi(f) d\mu_{N,m} - \phi \left( \int f d\mu_{N,m} \right) &= \int \left[ \int \phi(f(x)) \mu(dx|y) - \phi(\bar{f}(y)) \right] \bar{\mu}(dy) \\ &\quad + \left[ \int \phi(\bar{f}(y)) \bar{\mu}(dy) - \phi \left( \int \bar{f}(y) \bar{\mu}(dy) \right) \right]. \end{aligned}$$

An application of Lemma 2.11 yields the estimate

$$\begin{aligned} \int \left[ \int \phi(f(x)) \mu(dx|y) - \phi(\bar{f}(y)) \right] \bar{\mu}(dy) \\ \leq \frac{1}{2\varrho} \int \int \frac{|(\text{id}_X - 2P^t P) \nabla f(x)|^2}{f(x)} \mu(dx|y) \bar{\mu}(dy). \end{aligned}$$

By assumption the marginal  $\bar{\mu}$  satisfies the LSI with constant  $\lambda > 0$ . Together with Proposition 2.12 this yields the estimate

$$\begin{aligned} \int \phi(\bar{f}(y)) \bar{\mu}(dy) - \phi \left( \int \bar{f}(y) \bar{\mu}(dy) \right) &\leq \frac{1}{2\lambda} \int \frac{|\nabla \bar{f}(y)|^2}{\bar{f}(y)} \bar{\mu}(dy) \\ &\lesssim \int \int \frac{|\nabla f(x)|^2}{f(x)} \mu(dx|y) \bar{\mu}(dy). \end{aligned}$$

A combination of the last three formulas and the observations (8) and (18) yield

$$\begin{aligned} \int \phi(f) d\mu_{N,m} - \phi \left( \int f d\mu_{N,m} \right) \\ \lesssim \int \frac{|(\text{id}_X - 2P^t P) \nabla f(x)|^2}{f(x)} \mu_{N,m}(dx) + \int \frac{|\nabla f(x)|^2}{f(x)} \mu_{N,m}(dx) \\ \lesssim \int \frac{|\nabla f(x)|^2}{f(x)} \mu_{N,m}(dx), \end{aligned}$$

uniformly in the system size  $N$  and the mean spin  $m$ . □

Because the hierarchic criterion for the LSI is an important ingredient in the proof of the main result, we outline the proof of Proposition 2.12 in full detail. We follow the proof of [GOVW09][Proposition 20], which is based on two lemmas. We directly take over the first lemma (cf. [GOVW09, Lemma 21]), which in our notation becomes:

**Lemma 2.13.** *For any function  $f$  on  $X$  and any  $y \in Y$  it holds*

$$\int P \nabla f(x) \mu(dx|y) = \frac{1}{2} \nabla \bar{f}(y) + P \text{cov}_{\mu(dx|y)}(f, \nabla H).$$

**Remark 2.14.** *The notational difference compared to [GOVW09, Lemma 21] is based on our choice of the Euclidean structure on  $Y = X_{\frac{N}{2}, m}$ . Compared to the notation in Lemma 21 of [GOVW09] we have*

$$\nabla_Y \bar{f}(y) = \frac{N}{2} \nabla \bar{f}(y).$$

*Hence we omit the proof, which is a straight-forward calculation.*

The more interesting ingredient of the proof of [GOVW09, Proposition 20] is the estimate (see [GOVW09, (42),(43)])

$$|2P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2 \leq \frac{\sqrt{2}\kappa^2}{\varrho^2} \bar{f}(y) \int \frac{|(\operatorname{id}_X - 2P^t P)\nabla f(x)|^2}{f(x)} \mu(dx|y). \quad (20)$$

The estimate (20) follows in [GOVW09] by direct calculation from the standard covariance estimate given by Lemma 2.9. In contrast to [GOVW09] we cannot use the estimate (20), because the constant  $\kappa$  given by (17) maybe infinite for a perturbed strictly convex single-site potential  $\psi$ . We avoid this problem by applying the more robust asymmetric Brascamp-Lieb inequality given by Lemma 2.10. Our substitute for (20) is:

**Lemma 2.15.** *For any non-negative function  $f$*

$$|2P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2 \lesssim \bar{f}(y) \int \frac{|\nabla f(x)|^2}{f(x)} \mu(dx|y),$$

*uniformly in the system size  $N$ , the macroscopic profile  $y$ , and the mean spin  $m$ .*

We postpone the proof of Lemma 2.15 and show how it is used in the proof of Proposition 2.12 (cf. proof of [GOVW09][Proposition 20]).

*Proof of Proposition 2.12.* Note that because for any  $a, b \in \mathbb{R}$

$$\frac{1}{2}(a + b)^2 \leq a^2 + b^2,$$

it follows from the definition (7) of  $P$  that for any  $x$

$$|Px|^2 \leq |x^2|. \quad (21)$$

By successively using Lemma 2.13 and Jensen's inequality (with the convex function  $(a, b) \mapsto |b|^2/a$ ), we have

$$\begin{aligned} \frac{|\nabla \bar{f}(y)|^2}{\bar{f}(y)} &= \frac{4}{\bar{f}(y)} \left| P \int \nabla f(x) \mu(dx|y) - P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H) \right|^2 \\ &\lesssim \frac{1}{\bar{f}(y)} \left| \int P \nabla f(x) \mu(dx|y) \right|^2 + \frac{1}{\bar{f}(y)} |P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2 \\ &\lesssim \int \frac{|P \nabla f(x)|^2}{f(x)} \mu(dx|y) + \frac{1}{\bar{f}(y)} |2P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2. \end{aligned}$$

On the first term on the r.h.s. we apply the estimate (21). On the second term we apply Lemma 2.15, which yields the desired estimate.  $\square$

Now, we state the proof of Lemma 2.15, which also represents one of the main differences compared to the two-scale approach of [GOVW09]. The main ingredients are the product structure (19) of  $\mu(dx|y)$  and the asymmetric Brascamp-Lieb inequality (cf. Lemma 2.10).

*Proof of Lemma 2.15.* We have to estimate the covariance

$$|2P \operatorname{cov}_{\mu(dx|y)}(f, \nabla H)|^2 = \sum_{j=1}^{\frac{N}{2}} |\operatorname{cov}_{\mu(dx|y)}(f, (2P\nabla H)_j)|^2. \quad (22)$$

Therefore, let us consider for  $j \in \{1, \dots, \frac{N}{2}\}$  the term  $\operatorname{cov}_{\mu(dx|y)}(f, (2P\nabla H)_j)$ . Note that the function

$$(2P\nabla H(x))_j = \psi'(x_{2j-1}) + \psi'(x_{2j})$$

only depends of the variables  $x_{2j-1}$  and  $x_{2j}$ . Hence, the product structure (19) of  $\mu(dx|y)$  yields the identity

$$\begin{aligned} & \operatorname{cov}_{\mu(dx|y)}(f, 2(P\nabla H)_j) \\ &= \int \operatorname{cov}_{\mu_{2,y_j}(dx_{2j-1}, dx_{2j})}(f, (2P\nabla H)_j) \bigotimes_{i=1, i \neq j}^{\frac{N}{2}} \mu_{2,y_i}(dx_{2i-1}, dx_{2i}). \end{aligned} \quad (23)$$

As we will show below, we obtain by using the asymmetric Brascamp-Lieb inequality of Lemma 2.10 and the Csiszár-Kullback-Pinsker inequality the estimate

$$\begin{aligned} & \left| \operatorname{cov}_{\mu_{2,y_j}(dx_{2j-1}, dx_{2j})}(f, (2P\nabla H)_j) \right| \lesssim \left( \int f(x) \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \right)^{\frac{1}{2}} \\ & \times \left( \int \frac{|\frac{d}{dx_{2j-1}} f(x)|^2 + |\frac{d}{dx_{2j}} f(x)|^2}{f(x)} \mu_{2,y_j}(dx_{2j-1}, dx_{2j}) \right)^{\frac{1}{2}} \end{aligned} \quad (24)$$

uniformly in  $j$  and  $y_j$ . Therefore, a combination of the identity (23), the last estimate, and Hölder's inequality yield

$$\begin{aligned} & |\operatorname{cov}_{\mu(dx|y)}(f, (2P\nabla H)_j)|^2 \\ & \lesssim \int f(x) \mu(dx|y) \int \frac{|\frac{d}{dx_{2j-1}} f(x)|^2 + |\frac{d}{dx_{2j}} f(x)|^2}{f(x)} \mu(dx|y), \end{aligned}$$

which implies the desired estimate by the identity (22). It is only left to deduce the estimate (24). We assume w.l.o.g.  $j = 1$ . Recall the splitting  $\psi = \psi_c + \delta\psi$  given by (6). We use the bound on  $|\delta\psi'|$  to estimate

$$\begin{aligned} & \left| \operatorname{cov}_{\mu_{2,y_1}(dx_1, dx_2)}(f, (2P\nabla H)_1) \right| \lesssim \left| \operatorname{cov}_{\mu_{2,y_1}(dx_1, dx_2)}(f, \psi'_c(x_1) + \psi'_c(x_2)) \right| \\ & \quad + \int \left| f - \int f \mu_{2,y_1}(dx_1, dx_2) \right| \mu_{2,y_1}(dx_1, dx_2). \end{aligned} \quad (25)$$

Now, we consider the first term on the r.h.s. of the last estimate. For  $y_1 \in \mathbb{R}$  let the one-dimensional probability measure  $\nu(dz|y_1)$  be defined by the density

$$\nu(dz|y_1) := \frac{1}{Z} \exp(-(\psi(z+y_1) + \psi(-z+y_1))) dz.$$

A reparametrization of the one-dimensional Hausdorff measure implies

$$\int \xi(x_1, x_2) \mu_{2, y_1}(dx_1, dx_2) = \int \xi(-z+y_1, z+y_1) \nu(dz|y_1) \quad (26)$$

for any measurable function  $\xi$ . We may assume w.l.o.g. that the function  $f(x) = f(x_1, x_2)$  just depends on the variables  $x_1$  and  $x_2$ . Hence for

$$\tilde{f}(z, y_1) := f(-z+y_1, z+y_1) \quad \text{and} \quad \tilde{g}(z, y_1) := \psi'_c(-z+y_1) + \psi'_c(z+y_1)$$

the last identity yields

$$\text{cov}_{\mu_{2, y_1}(dx_1, dx_2)}(f, \psi'_c(x_1) + \psi'_c(x_2)) = \text{cov}_{\nu(dz|y_1)}(\tilde{f}, \tilde{g}).$$

Because

$$\left| \frac{\frac{d}{dz} \tilde{g}(z, y_1)}{\psi''_c(-z+y_1) + \psi''_c(z+y_1)} \right| = \left| \frac{-\psi''_c(-z+y_1) + \psi''_c(z+y_1)}{\psi''_c(-z+y_1) + \psi''_c(z+y_1)} \right| \leq 2,$$

an application of the asymmetric Brascamp-Lieb inequality (cf. Lemma 2.10) yields

$$\left| \text{cov}_{\nu(dz|y_1)}(\tilde{f}, \tilde{g}) \right| \lesssim \int \left| \frac{d}{dz} \tilde{f} \right| \nu(dz|y_1) \lesssim \left( \int \tilde{f} \nu(dz|y_1) \right)^{\frac{1}{2}} \left( \int \frac{|\frac{d}{dz} \tilde{f}|^2}{\tilde{f}} \nu(dz|y_1) \right)^{\frac{1}{2}}.$$

From the last inequality and (26) follows the estimate

$$\begin{aligned} & \left| \text{cov}_{\mu_{2, y_1}(dx_1, dx_2)}(f, \psi'_c(x_1) + \psi'_c(x_2)) \right| \\ & \lesssim \left( \int f \mu_{2, y_1}(dx_1, dx_2) \right)^{\frac{1}{2}} \left( \int \frac{|\frac{d}{dx_1} f|^2 + |\frac{d}{dx_2} f|^2}{f} \mu_{2, y_1}(dx_1, dx_2) \right)^{\frac{1}{2}}. \end{aligned} \quad (27)$$

We turn to the second term on the r.h.s. of (25). For convenience we write  $\tilde{f}(y_1) := \int f \mu_{2, y_1}(dx_1, dx_2)$ . An application of the (well-known) Csiszár-Kullback-Pinsker inequality (cf. [Csi67, Kul67]) yields

$$\begin{aligned} \int \left| f - \tilde{f}(y_1) \right| \mu_{2, y_1}(dx_1, dx_2) &= \tilde{f}(y_1) \int \left| \frac{f}{\tilde{f}(y_1)} - 1 \right| \mu_{2, y_1}(dx_1, dx_2) \\ &\lesssim \tilde{f}(y_1) \left( \int \frac{f}{\tilde{f}(y_1)} \log \frac{f}{\tilde{f}(y_1)} \mu_{2, y_1}(dx_1, dx_2) \right)^{\frac{1}{2}}. \end{aligned}$$

An application of the LSI for the measure  $\mu_{2,y_1}(dx_1, dx_2)$  implies

$$\begin{aligned} & \int \left| f - \int f \mu_{2,y_1}(dx_1, dx_2) \right| \mu_{2,y_1}(dx_1, dx_2) \\ & \lesssim \left( \int f \mu_{2,y_1}(dx_1, dx_2) \right)^{\frac{1}{2}} \left( \int \frac{|\frac{d}{dx_1} f|^2 + |\frac{d}{dx_2} f|^2}{f} \mu_{2,y_1}(dx_1, dx_2) \right)^{\frac{1}{2}}. \end{aligned}$$

A combination of (25), (27), and the last inequality yield the desired estimate (24).  $\square$

We turn to the proof of Lemma 2.4. Again, the main ingredient of the proof is the asymmetric Brascamp-Lieb inequality.

*Proof of Lemma 2.4.* We define

$$\bar{\psi}_c(m) := -\frac{1}{2} \log \int \exp(-\psi_c(-x+m) - \psi_c(x+m)) dx$$

and

$$\begin{aligned} \bar{\delta\psi}(m) &:= -\frac{1}{2} \log \int \exp(-\psi(-x+m) - \psi(x+m)) dx \\ &\quad + \frac{1}{2} \log \int \exp(-\psi_c(-x+m) - \psi_c(x+m)) dx. \end{aligned}$$

Now, we show that the splitting  $\mathcal{R}\psi = \bar{\psi}_c + \bar{\delta\psi}$  satisfies the conditions given by (6). Using the strict convexity of  $\psi_c$  it follows by a standard argument based on the Brascamp-Lieb inequality (cf. [BL76] and (16)) that the first condition is preserved i.e.

$$\bar{\psi}_c'' \gtrsim 1.$$

We turn to the perturbation  $\bar{\delta\psi}$ . For convenience, we introduce the measures

$$\nu(dx) := \frac{1}{Z} \exp(-\psi(-x+m) - \psi(x+m)) dx$$

and

$$\nu_c(dx) := \frac{1}{Z} \exp(-\psi_c(-x+m) - \psi_c(x+m)) dx$$

so that

$$\bar{\delta\psi}(m) = -\frac{1}{2} \log \int \exp(-\delta\psi(-x+m) - \delta\psi(x+m)) \nu_c(dx).$$

Direct calculation using the bound  $|\delta\psi| \lesssim 1$  yields

$$|\bar{\delta\psi}(m)| \lesssim 1.$$

We turn to the first derivative of  $\overline{\delta\psi}$ . A direct calculation based on the definition of  $\overline{\delta\psi}$  yields

$$\begin{aligned} 2\overline{\delta\psi}'(m) &= \int (\psi'(-x+m) + \psi'(x+m)) \nu(dx) \\ &\quad - \int (\psi'_c(-x+m) + \psi'_c(x+m)) \nu_c(dx). \end{aligned}$$

For  $s \in [0, 1]$  we define the measure

$$\nu^s(dx) := \frac{1}{Z} \exp(-\psi_c(-x+m) - \psi_c(x+m) - s\delta\psi(-x+m) - s\delta\psi(x+m)) dx$$

that interpolates between  $\nu^0 = \nu_c$  and  $\nu^1 = \nu$ . By the mean-value theorem there is  $s \in [0, 1]$  such that

$$\begin{aligned} 2\overline{\delta\psi}'(m) &= \frac{d}{ds} \int (\psi'_c(-x+m) + \psi'_c(x+m) + s\delta\psi'(-x+m) + s\delta\psi'(x+m)) \nu^s(dx) \\ &= \int (\delta\psi'(-x+m) + \delta\psi'(x+m)) \nu^s(dx) \\ &\quad + \text{cov}_{\nu^s} \left( \psi'_c(-x+m) + \psi'_c(x+m), \delta\psi(-x+m) + \delta\psi(x+m) \right) \\ &\quad + \text{cov}_{\nu^s} \left( s\delta\psi'(-x+m) + s\delta\psi'(x+m), \delta\psi(-x+m) + \delta\psi(x+m) \right). \end{aligned}$$

The first term on the r.h.s. is controlled by the assumption  $|\delta\psi'| \lesssim 1$ . We turn to the estimation of the first covariance term. An application of the asymmetric Brascamp-Lieb inequality of Lemma 2.10 and  $|\delta\psi| + |\delta\psi'| \lesssim 1$  yields the estimate

$$\begin{aligned} &\left| \text{cov}_{\nu^s} \left( \psi'_c(-x+m) + \psi'_c(x+m), \delta\psi(-x+m) + \delta\psi(x+m) \right) \right| \\ &\lesssim \sup_x \left| \frac{\psi''_c(-x+m) - \psi''_c(x+m)}{\psi''_c(-x+m) + \psi''_c(x+m)} \right| \int |-\delta\psi'(-x+m) + \delta\psi'(x+m)| \nu^s(dx) \\ &\lesssim 1. \end{aligned}$$

The second covariance term can be estimated using the assumption  $|\delta\psi| + |\delta\psi'| \lesssim 1$ . Summing up, we have deduced the desired estimate  $|\overline{\delta\psi}'| \lesssim 1$ .  $\square$

### 3. Convexification by iterated renormalization

In this section we prove Theorem 2.6 that states the convexification of a perturbed strictly convex single-site potential  $\psi$  by iterated renormalization. The proof relies on a

local Cramér theorem and some auxiliary results. The proof of Theorem 2.6 is given in the Subsection 3.1. The proofs of the auxiliary results are given in the Subsection 3.2.

### 3.1. Proof of Theorem 2.6

In view of Lemma 2.8 it suffices to show the strict convexity of the coarse-grained Hamiltonian  $\bar{H}_K$  defined by (10) for large  $K \gg 1$ . The strategy is the same as in [GOVW09, Proposition 31]. Let  $\varphi$  denote the Cramér transform of  $\psi$ , namely

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

Because  $\varphi$  is the Legendre transform of the strictly convex function

$$\varphi^*(\sigma) = \log \int \exp(\sigma x - \psi(x)) dx, \quad (28)$$

there exists for any  $m \in \mathbb{R}$  a unique  $\sigma = \sigma(m)$  such that

$$\varphi(m) = \sigma m - \varphi^*(\sigma). \quad (29)$$

From basic properties of the Legendre transform it follows that the  $\sigma$  is determined by the equation

$$\frac{d}{d\sigma} \varphi^*(\sigma) = \frac{\int x \exp(\sigma x - \psi(x)) dx}{\int \exp(\sigma x - \psi(x)) dx} = m. \quad (30)$$

The starting point of the proof of the convexification of the coarse-grained Hamiltonian  $\bar{H}_K(m)$  is the explicit representation

$$\tilde{g}_{K,m}(0) = \exp \left( K \varphi(m) - K \bar{H}_K(m) \right). \quad (31)$$

Here,  $\tilde{g}_{K,m}$  denotes the Lebesgue density of the distribution of the random variable

$$\frac{1}{\sqrt{K}} \sum_{i=1}^K (X_i - m),$$

where  $X_i$  are  $K$  real-valued independent random variables identically distributed as

$$\mu^\sigma(dx) := \exp \left( -\varphi^*(\sigma) + \sigma x - \psi(x) \right) dx. \quad (32)$$

We note that in view of (30) the mean of  $X_i$  is  $m$ . As in [GOVW09, (125)] the Cramér representation (31) follows from direct substitution and the coarea formula. As we will see in the proof of Lemma 3.3, the Cramér transform  $\varphi$  is strictly convex. The main idea of the proof is to transfer the convexity from  $\varphi$  to  $\bar{H}_K$  using the representation (31) and a local central limit type theorem for the density  $\tilde{g}_{K,m}$ , which is formulated in the next statement.

**Proposition 3.1.** *Let  $\psi(x)$  be a smooth function that is increasing sufficiently fast as  $|x| \uparrow \infty$  for all subsequent integrals to exist. Note that the probability measure  $\mu^\sigma$  defined by (32) depends on the field strength  $\sigma$ . We introduce its mean  $m$  and variance  $s^2$*

$$m := \int x \mu^\sigma(dx) \quad \text{and} \quad s^2 := \int (x - m)^2 \mu^\sigma(dx). \quad (33)$$

*We assume that uniformly in the field strength  $\sigma$ , the probability measure  $\mu^\sigma$  has its standard deviation  $s$  as unique length scale in the sense that*

$$\int |x - m|^k \mu^\sigma(dx) \lesssim s^k \quad \text{for } k = 1, \dots, 5, \quad (34)$$

$$\left| \int \exp(ix\xi) \mu^\sigma(dx) \right| \lesssim |s\xi|^{-1} \quad \text{for all } \xi \in \mathbb{R}. \quad (35)$$

*Consider  $K$  independent random variables  $X_1, \dots, X_K$  identically distributed according to  $\mu^\sigma$ . Let  $g_{K,\sigma}$  denote the Lebesgue density of the distribution of the normalized sum  $\frac{1}{\sqrt{K}} \sum_{i=1}^K \frac{X_i - m}{s}$ .*

*Then  $g_{K,\sigma}(0)$  converges for  $K \uparrow \infty$  to the corresponding value for the normalized Gaussian. This convergence is uniform in  $m$ , of order  $\frac{1}{\sqrt{K}}$ , and  $C^2$  in  $\sigma$ :*

$$|g_{K,\sigma}(0) - \frac{1}{\sqrt{2\pi}}| \lesssim \frac{1}{\sqrt{K}}, \quad (36)$$

$$\left| \frac{1}{s} \frac{d}{d\sigma} g_{K,\sigma}(0) \right| \lesssim \frac{1}{\sqrt{K}}, \quad (37)$$

$$\left| \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 g_{K,\sigma}(0) \right| \lesssim \frac{1}{\sqrt{K}}. \quad (38)$$

Let us comment a bit on this result: Quantitative versions of the central limit theorem like (36) are abundant in the literature, see for instance [Fel71][Chapter XVI], [KL99][Appendix 2], [GPV88][Section 3], and [LPY02][p. 752 an Section 5]. In his work on the spectral gap, Caputo appeals even to a finer estimate that makes the first terms in an error expansion in  $\frac{1}{\sqrt{K}}$  explicit [Cap03, Theorem 2.1]. The coefficients of the higher order terms are expressed in terms of moments of  $\mu^\sigma$ . However, following [GOVW09, Proposition 31], for our two-scale argument we need *pointwise* control of the Lebesgue density  $g_{K,\sigma}$  (in form of  $g_{K,\sigma}(0)$ ) and, in addition, control of derivatives of  $g_{K,\sigma}$  w.r.t. the field parameter  $\sigma$ , cf. (37), (38). Note that the derivative  $\frac{d}{d\sigma}$  has units of length (because  $\sigma$ , which multiplies  $x$  in the Hamiltonian, cf. (32), has units of inverse length) so that  $\frac{1}{s} \frac{d}{d\sigma}$  is the properly non-dimensionalized derivative. Pointwise control means that control of the moments, cf. (34), is not sufficient. One also needs to know that  $\mu^\sigma$  has no fine structure on scales much smaller than  $s$ . This property is ensured the upper bound (35).

As opposed to [GOVW09, Proposition 31], the Hamiltonian  $\psi$  we want Proposition 3.1 apply to is not a perturbation of the quadratic  $\frac{1}{2}x^2$  but of a general strictly convex potential  $\psi$ . As a consequence, the variance  $s^2$  can be a strongly varying function of the field strength  $\sigma$ . Nevertheless, Lemma 3.2 from below shows that every element  $\mu^\sigma$  in the family of measures is characterized by the single length scale  $s$ , uniformly in  $\sigma$  in the sense of (34) and (35). For the verification of (34) in Lemma 3.2, one could take over the argument of [Cap03, Lemma 2.2] that relies on a result by Bokbov [Bob99] stating that the SG constant  $\varrho$  of the measure  $\mu^\sigma$  can be estimated by its variance i.e.  $\varrho \gtrsim \frac{1}{s^2}$ . However, we provide a self-contained argument for the verification of (34) and (35) in Lemma 3.2 just using basic calculus of one variable. The merit of Proposition 3.1 consists in providing a version of the central limit theorem that is  $C^2$  in the field strength  $\sigma$  even if the variance  $s^2$  varies strongly with  $\sigma$ .

**Lemma 3.2.** *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). Then  $s \lesssim 1$  uniformly in  $m$ , and the conditions (34) and (35) of Proposition 3.1 are satisfied.*

Using Proposition 3.1, Lemma 3.2, and the Cramér representation (31) we could easily deduce a local Cramér theorem (cf. [GOVW09, Proposition 31]) for general perturbed strictly convex potentials  $\psi$ . However, because we are just interested in the convexification of  $\bar{H}_K$ , we just consider the convergence of the second derivatives of  $\varphi$  and  $\bar{H}_K$ .

**Lemma 3.3.** *Assume that the single-site potential  $\psi$  is perturbed strictly convex in the sense of (6). Then for all  $m \in \mathbb{R}$  it holds*

$$\left| \frac{d^2}{dm^2} \varphi(m) - \frac{d^2}{dm^2} \bar{H}_K(m) \right| \lesssim \frac{1}{Ks^2},$$

where  $s^2$  is defined as in Proposition 3.1.

*Proof of Theorem 2.6.* Because of Lemma 2.8 it suffices to show that there exists  $\delta > 0$  and  $K_0 \in \mathbb{N}$  such that for all  $K \geq K_0$  and  $m \in \mathbb{R}$

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

We start with some formulas on the derivatives of  $\varphi$ . Differentiation of the identity (29) yields

$$\begin{aligned} \frac{d}{dm} \varphi &= \frac{d}{dm} \sigma m + \sigma - \frac{d}{d\sigma} \varphi^* \frac{d}{dm} \sigma \\ &\stackrel{(30)}{=} \frac{d}{dm} \sigma m + \sigma - m \frac{d}{dm} \sigma \\ &= \sigma. \end{aligned}$$

A direct calculation reveals that (see (61) below)

$$\frac{d}{d\sigma}m = s^2,$$

where  $s^2$  is defined as in Proposition 3.1. Hence, a second differentiation of  $\varphi$  yields the identity

$$\frac{d^2}{dm^2}\varphi = \frac{d}{dm}\sigma = \left(\frac{d}{d\sigma}m\right)^{-1} = \frac{1}{s^2}. \quad (39)$$

By Lemma 3.3 we thus have

$$\begin{aligned} \frac{d^2}{dm^2}\bar{H}_K &= \frac{d^2}{dm^2}\varphi + \frac{d^2}{dm^2}(\bar{H}_K - \varphi) \\ &\geq \frac{1}{s^2} - \frac{C}{K} \frac{1}{s^2} \\ &\geq \frac{1}{2} \frac{1}{s^2}, \end{aligned}$$

if  $K \geq K_0$  for some large  $K_0$ . The statement follows from the uniform bound  $s \lesssim 1$  provided by Lemma 3.2.  $\square$

### 3.2. Proof of the local Cramér theorem and of the auxiliary results

In this section we prove the auxiliary statements of the last subsection. Before turning to the proof of Proposition 3.1 we sketch the strategy. For convenience we introduce the notation

$$\langle f \rangle := \int f(x)\mu^\sigma(dx) = \int f(x) \exp(-\varphi^*(\sigma) + \sigma x - \psi(x)) dx. \quad (40)$$

The definition of  $g_{K,\sigma}$  (cf. Proposition 3.1) suggests to introduce the shifted and rescaled variable

$$\hat{x} := \frac{x - m}{s}. \quad (41)$$

We note that by (33) the first and second moment in  $\hat{x}$  are normalized

$$\langle \hat{x} \rangle = 0, \quad \langle \hat{x}^2 \rangle = 1 \quad (42)$$

and that (34) turns into

$$\sum_{k=1}^5 \langle |\hat{x}|^k \rangle \lesssim 1. \quad (43)$$

Proposition 3.1 is a version of the central limit theorem that, like most others, is best proved with help of the Fourier transform. Indeed, since the random variables  $\hat{X}_1 := \frac{X_1 - m}{s}, \dots, \hat{X}_K := \frac{X_K - m}{s}$  in the statement of Proposition 3.1 are independent and identically distributed, the distribution of their sum is the  $K$ -fold convolution of the distribution of  $\hat{X}_1$ . Therefore, the Fourier transform of the distribution of the  $\sum_{n=1}^K \hat{X}_n$  is the  $K$ -th power of the Fourier transform of the distribution of  $\hat{X}$ . The latter is given by

$$\langle \exp(i\hat{x}\hat{\xi}) \rangle,$$

where  $\hat{\xi}$  denotes the variable dual to  $\hat{x}$ . Hence, the Fourier transform of the distribution of the normalized sum  $\frac{1}{\sqrt{K}} \sum_{n=1}^K \hat{X}_K$  is given by  $\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle^K$ . Applying the inverse Fourier transform, we obtain the representation

$$2\pi g_{K,\sigma}(0) = \int \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle^K d\hat{\xi}. \quad (44)$$

In order to make use of formula (44), we need estimates on  $\langle \exp(i\hat{x}\hat{\xi}) \rangle$ . Because of

$$\frac{d^k}{d\hat{\xi}^k} \langle \exp(i\hat{x}\hat{\xi}) \rangle = i^k \langle \hat{x}^k \exp(i\hat{x}\hat{\xi}) \rangle, \quad (45)$$

the moment bounds (43) translate into control of  $\langle \exp(i\hat{x}\hat{\xi}) \rangle$  for  $|\hat{\xi}| \ll 1$ . Together with the normalization (42), we obtain in particular

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle - (1 - \frac{1}{2}\hat{\xi}^2)| \lesssim |\hat{\xi}|^3.$$

We will use the latter in the following form: There exists a complex-valued function  $h(\hat{\xi})$  such that for  $|\hat{\xi}| \ll 1$ :

$$\langle \exp(i\hat{x}\hat{\xi}) \rangle = \exp(-h(\hat{\xi})) \quad \text{with} \quad |h(\hat{\xi}) - \frac{1}{2}\hat{\xi}^2| \lesssim |\hat{\xi}|^3. \quad (46)$$

This estimate, showing that the Fourier transform of the normalized probability  $\langle \cdot \rangle$  is close for  $|\hat{\xi}| \ll 1$  to the Fourier transform of the normalized Gaussian, is at the core of most proofs of the central limit theorem.

Estimate (46) provides good control over  $\langle \exp(i\hat{x}\hat{\xi}) \rangle$  for  $|\hat{\xi}| \ll 1$ . Another key ingredient is uniform decay for  $|\hat{\xi}| \gg 1$ . In our new variables, (35) takes on the form

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle| \lesssim |\hat{\xi}|^{-1}. \quad (47)$$

As usual in central limit theorems, we also need control of the characteristic function for intermediate values of  $|\hat{\xi}|$ . This can be inferred from (43) and (47) by a soft argument (in particular, it does not require the more intricate argument for [Cap03, (2.10)] from [Cap03, Lemma 2.5]):

**Lemma 3.4.** *Under the assumptions of Proposition 3.1 and for any  $\delta > 0$  there exists  $\lambda < 1$  such that for all  $\sigma$*

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle| \leq \lambda \quad \text{for all } |\hat{\xi}| \geq \delta.$$

So far, the strategy is standard; now comes the new ingredient: In view of formula (44), in order to control  $\sigma$ -derivatives of  $g_{K,\sigma}(0)$ , we need to control  $\frac{1}{s} \frac{1}{d\sigma} \langle \exp(i\hat{x}\hat{\xi}) \rangle$ . Relying on the identities

$$\frac{1}{s} \frac{1}{d\sigma} \langle f(x) \rangle = \langle \hat{x} f(x) \rangle, \quad (48)$$

$$\frac{1}{s} \frac{1}{d\sigma} \hat{x} = -1 - \frac{1}{2} \langle \hat{x}^3 \rangle \hat{x}, \quad (49)$$

that will be established in the proof of Lemma 3.5 below, we see that the estimate again follow from the moment control (43). Lemma 3.5 is the only new element of our analysis.

**Lemma 3.5.** *Under the assumptions of Proposition 3.1 we have*

$$\left| \frac{1}{s} \frac{1}{d\sigma} \langle \exp(i\hat{x}\hat{\xi}) \rangle \right| \lesssim (1 + |\hat{\xi}|) |\hat{\xi}|^3, \quad (50)$$

$$\left| \left( \frac{1}{s} \frac{1}{d\sigma} \right)^2 \langle \exp(i\hat{x}\hat{\xi}) \rangle \right| \lesssim (1 + \hat{\xi}^2) |\hat{\xi}|^3. \quad (51)$$

Before turning to the proof of Proposition 3.1, we prove Lemma 3.4 and Lemma 3.5.

*Proof of Lemma 3.4.* In view of (43) and (47), it suffices to show: For any  $C < \infty$  and  $\delta > 0$  there exists  $\lambda < 1$  with the following property: Suppose  $\langle \cdot \rangle$  is a probability measure (in  $\hat{x}$ ) such that

$$\langle |\hat{x}| \rangle \leq C, \quad (52)$$

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle| \leq \frac{C}{|\hat{\xi}|} \quad \text{for all } \hat{\xi}. \quad (53)$$

Then

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle| \leq \lambda \quad \text{for all } |\hat{\xi}| \geq \delta.$$

In view of (53), it is enough to show

$$|\langle \exp(i\hat{x}\hat{\xi}) \rangle| \leq \lambda \quad \text{for all } \delta \leq |\hat{\xi}| \leq \frac{1}{\delta}.$$

We give an indirect argument for this statement and thus assume that there is a sequence  $\{\langle \cdot \rangle_\nu\}$  of probability measures satisfying (52) & (53) and a sequence  $\{\hat{\xi}_\nu\}$  of numbers in  $[\delta, \frac{1}{\delta}]$  such that

$$\liminf_{\nu \uparrow \infty} |\langle \exp(i\hat{x}\hat{\xi}_\nu) \rangle_\nu| \geq 1. \quad (54)$$

In view of (52), after passage to a subsequence, we may assume that there exists a probability measure  $\langle \cdot \rangle_\infty$  and a number  $\hat{\xi}_\infty > 0$  such that

$$\lim_{\nu \uparrow \infty} \langle f \rangle_\nu = \langle f \rangle_\infty \quad \text{for all bounded and continuous } f(\hat{x}), \quad (55)$$

$$\lim_{\nu \uparrow \infty} \hat{\xi}_\nu = \hat{\xi}_\infty. \quad (56)$$

Since  $|\exp(i\hat{x}\hat{\xi}_\nu) - \exp(i\hat{x}\hat{\xi}_\infty)| \leq |\hat{x}| |\hat{\xi}_\nu - \hat{\xi}_\infty|$ , we obtain from (52), (55) & (56):

$$\lim_{\nu \uparrow \infty} \langle \exp(i\hat{x}\hat{\xi}_\nu) \rangle_\nu = \langle \exp(i\hat{x}\hat{\xi}_\infty) \rangle_\infty,$$

so that (54) saturates to

$$|\langle \exp(i\hat{x}\hat{\xi}_\infty) \rangle_\infty| \geq 1. \quad (57)$$

On the other hand, (53) is preserved under (55) so that we have in particular

$$\lim_{|\hat{\xi}| \uparrow \infty} |\langle \exp(i\hat{x}\hat{\xi}) \rangle_\infty| = 0. \quad (58)$$

We claim that (57) and (58) contradict each other. Indeed, since  $\hat{x} \mapsto \exp(i\hat{x}\hat{\xi}_\infty)$  is  $S^1$ -valued, it follows from (57) that there is a fixed  $\zeta \in S^1$  such that

$$\exp(i\hat{x}\hat{\xi}_\infty) = \zeta \quad \text{for } \langle \cdot \rangle_\infty - a. e. \hat{x}.$$

This implies for every  $n \in \mathbb{N}$

$$\exp(i\hat{x}(n\hat{\xi}_\infty)) = \zeta^n \quad \text{for } \langle \cdot \rangle_\infty - a. e. \hat{x}$$

and thus

$$|\langle \exp(i\hat{x}(n\hat{\xi}_\infty)) \rangle_\infty| = |\zeta^n| = 1, \quad (59)$$

which in view of  $\hat{\xi}_\infty \neq 0$  and thus  $|n\hat{\xi}_\infty| \uparrow \infty$  as  $n \uparrow \infty$  contradicts (58).  $\square$

*Proof of Lemma 3.5.* We restrict our attention to estimate (51); estimate (50) is easier and can be derived by the same arguments. We start with the identities (48) and (49). Deriving (40) w.r.t.  $\sigma$  yields

$$\frac{d}{d\sigma} \langle f(x) \rangle = \langle (x - \frac{d\varphi^*}{d\sigma}) f(x) \rangle \stackrel{(30)}{=} \langle (x - m) f(x) \rangle. \quad (60)$$

In view of definition (41), the latter turns into (48).

We now turn to identity (49) and note that in view of definitions (33) and (41), the identity (60) yields in particular

$$\frac{d}{d\sigma} m \stackrel{(33),(60)}{=} \langle (x - m)x \rangle \stackrel{(33)}{=} \langle (x - m)^2 \rangle \stackrel{(33)}{=} s^2, \quad (61)$$

$$\frac{d}{d\sigma} s^2 \stackrel{(33),(60)}{=} \langle (x - m)(x - m)^2 \rangle \stackrel{(41)}{=} s^3 \langle \hat{x}^3 \rangle, \quad (62)$$

which we rewrite as

$$\begin{aligned}\frac{1}{s} \frac{d}{d\sigma} m &= s, \\ \frac{1}{s} \frac{d}{d\sigma} s &= \frac{1}{2} s \langle \hat{x}^3 \rangle.\end{aligned}\quad (63)$$

These formulas imply as desired

$$\frac{1}{s} \frac{d}{d\sigma} \hat{x} \stackrel{(41)}{=} \frac{1}{s} \frac{d}{d\sigma} \frac{x-m}{s} = -1 - \frac{1}{2} \langle \hat{x}^3 \rangle \hat{x}.$$

We now combine formulas (48) and (49) to express derivatives of  $\langle f(\hat{x}) \rangle$ . We start with the first derivative:

$$\begin{aligned}\frac{1}{s} \frac{d}{d\sigma} \langle f(\hat{x}) \rangle &\stackrel{(48)}{=} \left\langle \frac{df}{d\hat{x}}(\hat{x}) \frac{1}{s} \frac{d}{d\sigma} \hat{x} + f(\hat{x}) \hat{x} \right\rangle \\ &\stackrel{(49)}{=} - \left\langle \frac{df}{d\hat{x}}(\hat{x}) \right\rangle - \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \hat{x} \frac{df}{d\hat{x}}(\hat{x}) \right\rangle + \langle \hat{x} f(\hat{x}) \rangle.\end{aligned}\quad (64)$$

(As a consistency check we note that  $\frac{1}{s} \frac{d}{d\sigma} \langle f(\hat{x}) \rangle \stackrel{(64)}{=} - \left\langle \left( \frac{d}{d\hat{x}} - \hat{x} \right) f \right\rangle - \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \hat{x} \frac{df}{d\hat{x}} \right\rangle$  vanishes if  $\psi$  is quadratic since then the distribution of  $\hat{x}$  under  $\langle \cdot \rangle$  is the normalized Gaussian so that both  $\left\langle \left( \frac{d}{d\hat{x}} - \hat{x} \right) f \right\rangle = 0$  and  $\langle \hat{x}^3 \rangle = 0$ .)

Iterating this formula, we obtain for the second derivative

$$\begin{aligned}\left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle f(\hat{x}) \rangle &\stackrel{(64)}{=} - \frac{1}{s} \frac{d}{d\sigma} \left\langle \frac{df}{d\hat{x}}(\hat{x}) \right\rangle - \frac{1}{2} \left( \frac{1}{s} \frac{d}{d\sigma} \langle \hat{x}^3 \rangle \right) \left\langle \hat{x} \frac{df}{d\hat{x}}(\hat{x}) \right\rangle \\ &\quad - \frac{1}{2} \langle \hat{x}^3 \rangle \left( \frac{1}{s} \frac{d}{d\sigma} \left\langle \hat{x} \frac{df}{d\hat{x}}(\hat{x}) \right\rangle \right) + \frac{1}{s} \frac{d}{d\sigma} \langle \hat{x} f(\hat{x}) \rangle \\ &\stackrel{(64)}{=} \left\langle \frac{d^2 f}{d\hat{x}^2} \right\rangle + \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \hat{x} \frac{d^2 f}{d\hat{x}^2} \right\rangle - \left\langle \hat{x} \frac{df}{d\hat{x}} \right\rangle \\ &\quad + \frac{1}{2} \left( 3 \langle \hat{x}^2 \rangle + \frac{3}{2} \langle \hat{x}^3 \rangle^2 - \langle \hat{x}^4 \rangle \right) \left\langle \hat{x} \frac{df}{d\hat{x}} \right\rangle \\ &\quad + \frac{1}{2} \langle \hat{x}^3 \rangle \left( \left\langle \frac{df}{d\hat{x}} + \hat{x} \frac{d^2 f}{d\hat{x}^2} \right\rangle + \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \hat{x} \frac{df}{d\hat{x}} + \hat{x}^2 \frac{d^2 f}{d\hat{x}^2} \right\rangle - \langle \hat{x}^2 \frac{df}{d\hat{x}} \right\rangle \right) \\ &\quad - \left\langle f + \hat{x} \frac{df}{d\hat{x}} \right\rangle - \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \hat{x} f + \hat{x}^2 \frac{df}{d\hat{x}} \right\rangle + \langle \hat{x}^2 f \rangle \\ &= \left\langle \frac{d^2 f}{d\hat{x}^2} \right\rangle + \langle \hat{x}^3 \rangle \left\langle \hat{x} \frac{d^2 f}{d\hat{x}^2} \right\rangle + \frac{1}{4} \langle \hat{x}^3 \rangle^2 \left\langle \hat{x}^2 \frac{d^2 f}{d\hat{x}^2} \right\rangle \\ &\quad + \frac{1}{2} \langle \hat{x}^3 \rangle \left\langle \frac{df}{d\hat{x}} \right\rangle - \frac{1}{2} (1 - 2 \langle \hat{x}^3 \rangle^2 + \langle \hat{x}^4 \rangle) \left\langle \hat{x} \frac{df}{d\hat{x}} \right\rangle - \langle \hat{x}^3 \rangle \left\langle \hat{x}^2 \frac{df}{d\hat{x}} \right\rangle \\ &\quad - \langle f \rangle - \frac{1}{2} \langle \hat{x}^3 \rangle \langle \hat{x} f \rangle + \langle \hat{x}^2 f \rangle.\end{aligned}$$

Because of (45) we have for any  $k \in \mathbb{N}$

$$\frac{d^k}{d\hat{\xi}^k} \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \exp(i\hat{\xi}\hat{x}) \rangle = \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \frac{d^k}{d\hat{\xi}^k} \langle \exp(i\hat{\xi}\hat{x}) \rangle = i^k \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \hat{x}^k \exp(i\hat{\xi}\hat{x}) \rangle. \quad (65)$$

This formula and the normalization (42) yield that  $\left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \exp(i\hat{\xi}\hat{x}) \rangle$  vanishes to second order in  $\hat{\xi}$ . More precisely, for  $k \in \{0, 1, 2\}$

$$\frac{d^k}{d\hat{\xi}^k} \Big|_{\hat{\xi}=0} \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \exp(i\hat{\xi}\hat{x}) \rangle = i^k \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \hat{x}^k \rangle = 0. \quad (66)$$

Therefore, we consider the third derivative w.r.t.  $\hat{\xi}$  given by (65). For this purpose we apply the formula for  $\left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle f(\hat{x}) \rangle$  from above to the function  $f = \hat{x}^3 \exp(i\hat{\xi}\hat{x})$ . Using the abbreviation  $e := \exp(i\hat{\xi}\hat{x})$  we obtain

$$\begin{aligned} \frac{d^3}{d\hat{\xi}^3} \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle e \rangle &= i^3 \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \hat{x}^3 e \rangle \\ &= i^3 \left( 6 \langle \hat{x} e \rangle + i6\hat{\xi} \langle \hat{x}^2 e \rangle - \hat{\xi}^2 \langle \hat{x}^3 e \rangle \right. \\ &\quad + \langle \hat{x}^3 \rangle \left( 6 \langle \hat{x}^2 e \rangle + i6\hat{\xi} \langle \hat{x}^3 e \rangle - \hat{\xi}^2 \langle \hat{x}^4 e \rangle \right) \\ &\quad + \frac{1}{4} \langle \hat{x}^3 \rangle^2 \left( 6 \langle \hat{x}^3 e \rangle + i6\hat{\xi} \langle \hat{x}^4 e \rangle - \hat{\xi}^2 \langle \hat{x}^5 e \rangle \right) \\ &\quad + \frac{1}{2} \langle \hat{x}^3 \rangle \left( 3 \langle \hat{x}^2 e \rangle + i\hat{\xi} \langle \hat{x}^3 e \rangle \right) \\ &\quad - \frac{1}{2} \left( 1 - 2 \langle \hat{x}^3 \rangle^2 + \langle \hat{x}^4 \rangle \right) \left( 3 \langle \hat{x}^3 e \rangle + i\hat{\xi} \langle \hat{x}^4 e \rangle \right) \\ &\quad - \langle \hat{x}^3 \rangle \left( 3 \langle \hat{x}^4 e \rangle + i\hat{\xi} \langle \hat{x}^5 e \rangle \right) \\ &\quad \left. - \langle \hat{x}^3 e \rangle - \frac{1}{2} \langle \hat{x}^3 \rangle \langle \hat{x}^4 e \rangle + \langle \hat{x}^5 e \rangle \right). \end{aligned}$$

From this formula and the moment estimates (43) we obtain the estimate

$$\left| \frac{d^3}{d\hat{\xi}^3} \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle e \rangle \right| \lesssim 1 + \hat{\xi}^2.$$

In combination with (66), this estimate yields (51). □

*Proof of Proposition 3.1.* We focus on (36) and (38). The intermediate (37) can be established as (38).

We start with (36). Fix a  $\delta > 0$  so small such that the expansion (46) of  $\langle \exp(i\hat{x}\hat{\xi}) \rangle$  holds for  $|\hat{\xi}| \leq \delta$ . We split the integral representation (44) accordingly:

$$\begin{aligned} 2\pi g_{K,\sigma}(0) &= \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta\}} \langle \exp(i\hat{x}\frac{1}{\sqrt{K}}\hat{\xi}) \rangle^K d\hat{\xi} \\ &+ \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| > \delta\}} \langle \exp(i\hat{x}\frac{1}{\sqrt{K}}\hat{\xi}) \rangle^K d\hat{\xi}. \end{aligned} \quad (67)$$

We consider the first term  $I$  on the r.h.s. of (67), which will turn out to be of leading order. Since  $\delta$  is so small that (46) holds, we may rewrite it as

$$I := \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta\}} \langle \exp(i\hat{x}\frac{1}{\sqrt{K}}\hat{\xi}) \rangle^K d\hat{\xi} = \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta\}} \exp(-Kh(\frac{1}{\sqrt{K}}\hat{\xi})) d\hat{\xi}. \quad (68)$$

We note that for  $|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta$  we have by (46),

$$|Kh(\frac{1}{\sqrt{K}}\hat{\xi}) - \frac{1}{2}\hat{\xi}^2| \lesssim \frac{1}{\sqrt{K}}|\hat{\xi}|^3, \quad (69)$$

in particular for  $\delta$  small enough

$$\operatorname{Re} \left( Kh(\frac{1}{\sqrt{K}}\hat{\xi}) \right) \geq \frac{1}{4}\hat{\xi}^2, \quad (70)$$

so that (69) implies by the Lipschitz continuity of  $\mathbb{C} \ni y \mapsto \exp(y) \in \mathbb{C}$  on  $\operatorname{Re} y \leq -\frac{1}{4}\hat{\xi}^2$  with constant  $\exp(-\frac{1}{4}\hat{\xi}^2)$ :

$$|\exp(-Kh(\frac{1}{\sqrt{K}}\hat{\xi})) - \exp(-\frac{1}{2}\hat{\xi}^2)| \lesssim \frac{1}{\sqrt{K}}|\hat{\xi}|^3 \exp(-\frac{1}{4}\hat{\xi}^2).$$

Inserting this estimate into (68) we obtain

$$\begin{aligned} |I - \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta\}} \exp(-\frac{1}{2}\hat{\xi}^2) d\hat{\xi}| &\lesssim \frac{1}{\sqrt{K}} \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| \leq \delta\}} |\hat{\xi}|^3 \exp(-\frac{1}{4}\hat{\xi}^2) d\hat{\xi} \\ &\lesssim \frac{1}{\sqrt{K}} \int |\hat{\xi}|^3 \exp(-\frac{1}{4}\hat{\xi}^2) d\hat{\xi} \\ &\lesssim \frac{1}{\sqrt{K}}. \end{aligned}$$

The latter turns as desired into

$$\begin{aligned} |I - \sqrt{2\pi}| &= |I - \int \exp(-\frac{1}{2}\hat{\xi}^2) d\hat{\xi}| \\ &\lesssim \frac{1}{\sqrt{K}} + \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}| > \delta\}} \exp(-\frac{1}{2}\hat{\xi}^2) d\hat{\xi} \\ &\lesssim \frac{1}{\sqrt{K}}, \end{aligned}$$

since  $\int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}|>\delta\}} \exp(-\frac{1}{2}\hat{\xi}^2) d\hat{\xi}$  is exponentially small in  $K$ .

We now address the second term  $II$  on the r.h.s. of (67); On the integrand we use Lemma 3.4 (on  $K - 2$  of the  $K$  factors) and (47) (on the remaining 2 factors):

$$\begin{aligned} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^K &\lesssim \lambda^{K-2} \left( \frac{1}{1 + \frac{1}{\sqrt{K}} |\hat{\xi}|} \right)^2 \\ &\lesssim K \lambda^{K-2} \frac{1}{K + \hat{\xi}^2} \lesssim K \lambda^{K-2} \frac{1}{1 + \hat{\xi}^2}. \end{aligned}$$

It follows that the second term  $II$  on the r.h.s. of (67) is exponentially small and thus higher order:

$$\begin{aligned} \left| \int_{\{|\frac{1}{\sqrt{K}}\hat{\xi}|>\delta\}} \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle^K d\hat{\xi} \right| &\lesssim K \lambda^{K-2} \int \frac{1}{1 + \hat{\xi}^2} d\hat{\xi} \\ &\lesssim K \lambda^{K-2} \stackrel{\lambda \ll 1}{\ll} \frac{1}{\sqrt{K}}. \end{aligned}$$

We now turn to (38). We take the second  $\sigma$ -derivative of the integral representation (44):

$$\begin{aligned} &2\pi \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 g_{K,\sigma}(0) \\ &= \int \left( K(K-1) \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle^{K-2} \left( \frac{1}{s} \frac{d}{d\sigma} \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle \right)^2 \right. \\ &\quad \left. + K \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle^{K-1} \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 \langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle \right) d\hat{\xi} \end{aligned} \quad (71)$$

and use Lemma 3.5:

$$\begin{aligned} &\left| \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 g_{K,\sigma}(0) \right| \\ &\lesssim \int \left( K^2 |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (1 + |\frac{1}{\sqrt{K}} \hat{\xi}|^2) |\frac{1}{\sqrt{K}} \hat{\xi}|^6 \right. \\ &\quad \left. + K |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-1} (1 + |\frac{1}{\sqrt{K}} \hat{\xi}|^2) |\frac{1}{\sqrt{K}} \hat{\xi}|^3 \right) d\hat{\xi} \\ &\lesssim \frac{1}{\sqrt{K}} \int |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (1 + |\frac{1}{\sqrt{K}} \hat{\xi}|^2) (|\hat{\xi}|^6 + 1) d\hat{\xi}. \end{aligned} \quad (72)$$

As for (36), we split the integral representation (72) according to  $\delta$ :

$$\begin{aligned}
& \left| \left( \frac{1}{s} \frac{d}{d\sigma} \right)^2 g_{K,\sigma}(0) \right| \\
& \lesssim \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| \leq \delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (1 + |\frac{1}{\sqrt{K}} \hat{\xi}|^2) (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \quad + \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| > \delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (1 + |\frac{1}{\sqrt{K}} \hat{\xi}|^2) (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \lesssim \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| \leq \delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \quad + \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| > \delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (\hat{\xi}^8 + 1) d\hat{\xi}. \tag{73}
\end{aligned}$$

On the first r.h.s. term we use (70):

$$\begin{aligned}
& \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| \leq \delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \lesssim \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}| \leq \delta\}} \exp\left(- (K-2) \frac{1}{4} \left(\frac{1}{\sqrt{K}} \hat{\xi}\right)^2\right) (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \stackrel{K \gg 1}{\lesssim} \frac{1}{\sqrt{K}} \int \exp\left(-\frac{1}{8} \hat{\xi}^2\right) (\hat{\xi}^6 + 1) d\hat{\xi} \\
& \lesssim \frac{1}{\sqrt{K}}. \tag{74}
\end{aligned}$$

On the integrand of the second r.h.s. term in (73) we use Lemma 3.4 (on  $K - 12$  of the  $K - 2$  factors) and (47) (on the remaining 10 factors):

$$\begin{aligned}
|\langle \exp(i\hat{x} \frac{1}{\sqrt{K}} \hat{\xi}) \rangle|^{K-2} (\hat{\xi}^8 + 1) & \lesssim \lambda^{K-12} \left( \frac{1}{1 + \frac{1}{\sqrt{K}} |\hat{\xi}|} \right)^{10} (\hat{\xi}^8 + 1) \\
& \lesssim K^5 \lambda^{K-12} \frac{1}{K^5 + \hat{\xi}^{10}} (\hat{\xi}^8 + 1) \\
& \lesssim K^5 \lambda^{K-12} \frac{1}{1 + \hat{\xi}^2}.
\end{aligned}$$

Hence, we see that this second term in (73) is exponentially small and thus higher order:

$$\begin{aligned}
& \frac{1}{\sqrt{K}} \int_{\{\frac{1}{\sqrt{K}}|\hat{\xi}|>\delta\}} |\langle \exp(i\hat{x} \frac{1}{\sqrt{K}}\hat{\xi}) \rangle|^{K-2} (|\hat{\xi}|^8 + 1) d\hat{\xi} \\
& \lesssim K^{9/2} \lambda^{K-12} \int \frac{1}{1 + \hat{\xi}^2} d\hat{\xi} \\
& \lesssim K^{9/2} \lambda^{K-12} \stackrel{\lambda \leq 1}{\ll} \frac{1}{\sqrt{K}}.
\end{aligned}$$

□

For the proof of Lemma 3.2 we need the following auxiliary statement, based on elementary calculus.

**Lemma 3.6.** *Assume that the single-site potential  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  is convex. We consider the corresponding Gibbs measure*

$$\nu(dx) = \frac{1}{Z} \exp(-\psi(x)) dx.$$

Let  $M$  denote the maximum of the density of  $\nu$  i.e.

$$M := \max_x \frac{1}{Z} \exp(-\psi(x)).$$

Then we have for all  $k \in \mathbb{N}$

$$\int |x|^k \nu(dx) \lesssim \frac{1}{M^k}$$

for some constant only depending on  $k$ .

*Proof of Lemma 3.6.* We may assume w.l.o.g. that

$$Z = \int \exp(-\psi(x)) dx = 1 \tag{75}$$

and  $M := \sup_x \exp(-\psi(x))$  is attained at  $x = 0$ , which means

$$M = \exp(-\psi(0)). \tag{76}$$

It follows from convexity of  $\psi$  that

$$\psi'(x) \leq 0 \text{ for } x \leq 0 \quad \text{and} \quad \psi'(x) \geq 0 \text{ for } x \geq 0. \tag{77}$$

We start with an analysis of the convex single-site potential  $\psi$ . We first argue that

$$\psi\left(\pm\frac{e}{M}\right) \geq -\log M + \log e. \quad (78)$$

Indeed in view of the monotonicity (77) we have

$$1 \stackrel{(75)}{\geq} \int_0^{\frac{e}{M}} \exp(-\psi(y)) dy \stackrel{(77)}{\geq} \frac{e}{M} \exp\left(-\psi\left(\frac{e}{M}\right)\right)$$

and

$$1 \stackrel{(75)}{\geq} \int_{-\frac{e}{M}}^0 \exp(-\psi(y)) dy \stackrel{(77)}{\geq} \frac{e}{M} \exp\left(-\psi\left(-\frac{e}{M}\right)\right).$$

We now argue that for  $|x| \geq \frac{e}{M}$

$$\psi(x) \geq \frac{M}{e} \left(|x| - \frac{e}{M}\right) - \log M. \quad (79)$$

W.l.o.g. we may restrict ourselves to  $x \geq \frac{e}{M}$ . By the mean-value theorem there is  $0 \leq \xi \leq \frac{e}{M}$  such that

$$\psi'(\xi) = \frac{\psi\left(\frac{e}{M}\right) - \psi(0)}{\frac{e}{M}}.$$

Using once again the monotonicity of  $\psi'$ , (76), and (78) yields the estimate

$$\psi'\left(\frac{e}{M}\right) \geq \psi'(\xi) \stackrel{(76)}{=} \frac{\psi\left(\frac{e}{M}\right) + \log M}{\frac{e}{M}} \stackrel{(78)}{\geq} \frac{M}{e}.$$

The convexity of  $\psi$ , the last estimate, and (78) yield for  $x \geq \frac{e}{M}$  as desired

$$\begin{aligned} \psi(x) &\geq \psi'\left(\frac{e}{M}\right) \left(x - \frac{e}{M}\right) + \psi\left(\frac{e}{M}\right) \\ &\geq \frac{M}{e} \left(x - \frac{e}{M}\right) - \log M. \end{aligned}$$

We finished the analysis on  $\psi$  and turn to the verification of the estimate of Lemma 3.6. We split the integral according to

$$\int |x|^k \exp(-\psi(x)) dx = \int_{-\infty}^0 |x|^k \exp(-\psi(x)) dx + \int_0^{\infty} |x|^k \exp(-\psi(x)) dx.$$

We will now deduce the estimate

$$\int_0^\infty |x|^k \exp(-\psi(x)) dx \lesssim \frac{1}{M^k}.$$

A similar estimate for the integral  $\int_{-\infty}^0 |x|^k \exp(-\psi(x)) dx$  follows from the same argument by symmetry. We split the integral:

$$\int_0^\infty |x|^k \exp(-\psi(x)) dx = \int_0^{\frac{e}{M}} |x|^k \exp(-\psi(x)) dx + \int_{\frac{e}{M}}^\infty |x|^k \exp(-\psi(x)) dx.$$

The first integral on the r.h.s. can be estimated as

$$\int_0^{\frac{e}{M}} |x|^k \exp(-\psi(x)) dx \leq \frac{e^k}{M^k} \int \exp(-\psi(x)) dx \stackrel{(75)}{=} \frac{e^k}{M^k}.$$

For the estimation of the second integral we apply (79), which yields by the change of variables  $\frac{M}{e} \left(x - \frac{e}{M}\right) = \hat{x}$

$$\begin{aligned} \int_{\frac{e}{M}}^\infty |x|^k \exp(-\psi(x)) dx &\leq \int_{\frac{e}{M}}^\infty |x|^k \exp\left(-\frac{M}{e} \left(x - \frac{e}{M}\right) + \log M\right) dx \\ &= M \frac{e}{M} \int_0^\infty \left|\frac{e}{M} \hat{x} + \frac{e}{M}\right|^k \exp(-\hat{x}) d\hat{x} \\ &= e \left(\frac{e}{M}\right)^k \int_0^\infty |\hat{x} + 1|^k \exp(-\hat{x}) d\hat{x} \\ &\lesssim \frac{1}{M^k}. \end{aligned}$$

□

Equipped with Lemma 3.6 we are able to give an elementary proof of Lemma 3.2:

*Proof of Lemma 3.2.* We argue that  $s \lesssim 1$ . Because  $\psi$  is a bounded perturbation of a uniformly strictly convex function, the measure  $\mu^\sigma$  given by (32) satisfies the SG uniformly in  $\sigma$ . This implies in particular

$$s^2 = \text{var}_{\mu^\sigma}(x) \lesssim \int \left(\frac{d}{dx} x\right)^2 d\mu^\sigma \lesssim 1 \quad (80)$$

uniformly in  $\sigma$  and thus in  $m$ .

Now, we verify (34). Using  $|\delta\psi| \lesssim 1$  to pass from  $\psi$  to  $\psi_e$ , we may assume that  $\psi$  is strictly convex. In fact, we can give up *strict* convexity of  $\psi$  and may only assume that  $\psi$  is convex. By the change of variables  $\hat{x} = \frac{x-m}{s}$  we have for any  $k \in \mathbb{N}$

$$\frac{\int |x - m|^k d\mu}{s^k} = \int |\hat{x}|^k \exp(-\hat{\psi}(\hat{x})) d\hat{x}$$

for some convex function  $\hat{\psi}$ , which is normalized in the sense that

$$\int \exp(-\hat{\psi}(\hat{x}))d\hat{x} = 1 \quad \text{and} \quad \int \hat{x}^2 \exp(-\hat{\psi}(\hat{x}))d\hat{x} = 1. \quad (81)$$

An application of Lemma 3.6 yields the estimate

$$\frac{\int |x - m|^k d\mu}{s^k} \leq \int |\hat{x}|^k \exp(-\hat{\psi}(\hat{x}))d\hat{x} \lesssim \frac{1}{M^k},$$

where  $M$  is given by  $M := \max_{\hat{x}} \exp(-\hat{\psi}(\hat{x}))$ . Now, we argue that due to the normalization of  $\hat{\psi}$  we have

$$M \geq C$$

for some universal constant  $C > 0$ , which verifies the desired estimate (34). Indeed the normalization (81) implies

$$\begin{aligned} \int_{(-2,2)} \exp(-\psi(x))dx &\stackrel{(81)}{=} 1 - \int_{\mathbb{R} \setminus (-2,2)} \exp(-\psi(x))dx \\ &\geq 1 - \frac{1}{4} \int x^2 \exp(-\psi(x))dx \stackrel{(81)}{\geq} \frac{3}{4}. \end{aligned}$$

Hence, there exists an  $x_0 \in (-2, 2)$  such that  $\exp(-\psi(x_0)) \geq \frac{3}{8}$ , which yields

$$M = \max_{\hat{x}} \exp(-\hat{\psi}(\hat{x})) \geq \exp(-\psi(x_0)) \geq \frac{3}{8}.$$

Let us turn to the statement (35) of Proposition 3.1. Writing

$$\exp(ix\xi) = \frac{d}{dx} \left( -i \frac{1}{\xi} \exp(ix\xi) \right)$$

we obtain by integration by parts that

$$\begin{aligned} \langle \exp(ix\xi) \rangle &= i \frac{1}{\xi} \int \exp(ix\xi) \frac{d}{dx} (\exp(-\varphi^*(\sigma) + \sigma x - \psi(x))) dx \\ &= i \frac{1}{\xi} \int \exp(ix\xi) (\sigma - \psi'(x)) \exp(-\varphi^*(\sigma) + \sigma x - \psi(x)) dx. \end{aligned}$$

The splitting  $\psi = \psi_c + \delta\psi$  with  $|\delta\psi|, |\delta\psi'| \lesssim 1$  and definition (28) of  $\varphi^*$  yield the estimate

$$|\langle \exp(ix\xi) \rangle| \lesssim \frac{1}{s|\xi|} \frac{s \int |\sigma - \psi'_c(x)| \exp(\sigma x - \psi_c(x)) dx}{\int \exp(\sigma x - \psi_c(x)) dx} + \frac{1}{s|\xi|} s,$$

where  $s$  is defined as in Proposition 3.1. Because  $s \lesssim 1$  by (80), we only have to consider the first term of the r.h.s. of the last inequality. We argue that for

$$M := \max_x \frac{\exp(\sigma x - \psi_c(x))}{\int \exp(\sigma x - \psi_c(x)) dx}$$

it holds

$$2M = \frac{\int |\sigma - \psi'_c(x)| \exp(\sigma x - \psi_c(x)) dx}{\int \exp(\sigma x - \psi_c(x)) dx}. \quad (82)$$

For the proof of the last statement, we only need the fact that the function  $H(x) = -\sigma x + \psi_c(x)$  is convex. W.l.o.g. we may assume that  $\int \exp(-H(x)) dx = 1$  and that  $M$  is attained at  $x = 0$ , which means

$$M = \exp(-H(0)).$$

It follows from convexity of  $H$  that

$$H'(x) \leq 0 \text{ for } x \leq 0 \quad \text{and} \quad H'(x) \geq 0 \text{ for } x \geq 0.$$

Therefore, we get

$$\begin{aligned} \int |H'(x)| \exp(-H(x)) dx &= - \int_{-\infty}^0 H'(x) \exp(-H(x)) dx + \int_0^{\infty} H'(x) \exp(H(x)) dx \\ &= \int_{-\infty}^0 \exp(-H(x))' dx - \int_0^{\infty} \exp(-H(x))' dx \\ &= 2 \exp(-H(0)) = 2M. \end{aligned}$$

Because the mean of a measure  $\mu$  is optimal in the sense that for all  $c \in \mathbb{R}$

$$\begin{aligned} \int (x - c)^2 \mu(dx) &= \int x^2 \mu(dx) - 2c \int x \mu(dx) + c^2 \\ &\geq \int x^2 \mu(dx) - \left( \int x \mu(dx) \right)^2 \\ &= \int \left( x - \int y \mu(dy) \right)^2 \mu(dx), \end{aligned} \quad (83)$$

we can estimate

$$s^2 \leq \frac{\int x^2 \exp(\sigma x - \psi(x)) dx}{\int \exp(\sigma x - \psi(x)) dx} \stackrel{|\delta\psi| \lesssim 1}{\lesssim} \frac{\int x^2 \exp(\sigma x - \psi_c(x)) dx}{\int \exp(\sigma x - \psi_c(x)) dx}. \quad (84)$$

Therefore, Lemma 3.6 applied to  $k = 2$  and  $\psi$  replaced by  $-\sigma x + \psi_c$  yields

$$\frac{s \int |\sigma - \psi'_c(x)| \exp(\sigma x - \psi_c(x)) dx}{\int \exp(\sigma x - \psi_c(x)) dx} \stackrel{(82),(84)}{\lesssim} \left( \frac{\int x^2 \exp(\sigma x - \psi_c(x)) dx}{\int \exp(\sigma x - \psi_c(x)) dx} \right)^{\frac{1}{2}} M \lesssim 1,$$

which verifies (35) of Proposition 3.1.  $\square$

Before we turn to the proof of Lemma 3.3 we will deduce the following auxiliary result.

**Lemma 3.7.** *Assume that (34) of Proposition 3.1 is satisfied. Then, using the notation of Proposition 3.1, it holds:*

$$(i) \quad \left| \frac{d}{dm} s \right| \lesssim 1 \quad \text{and} \quad (ii) \quad \left| \frac{d^2}{dm^2} s \right| \lesssim \frac{1}{s}.$$

*Proof of Lemma 3.7.* We start with restating some basic identities (cf. (61) and (62)): It holds that

$$\frac{d}{d\sigma} m = s^2, \tag{85}$$

$$\frac{d^2}{d\sigma^2} m = \frac{d}{d\sigma} s^2 = \int (x - m)^3 \mu^\sigma(dx), \tag{86}$$

$$\frac{d^3}{d\sigma^3} m = \int (x - m)^4 \mu^\sigma(dx). \tag{87}$$

Let us consider (i): It follows from (85) and (86) that

$$\begin{aligned} \frac{d}{dm} s^2 &= \frac{d}{d\sigma} s^2 \frac{d}{dm} \sigma \\ &= \int (x - m)^3 \mu^\sigma(dx) \left( \frac{d}{d\sigma} m \right)^{-1} \\ &= \frac{\int (x - m)^3 \mu^\sigma(dx)}{s^3} s, \end{aligned}$$

which yields by assumption (34) of Proposition 3.1 the estimate

$$\left| \frac{d}{dm} s^2 \right| \lesssim s.$$

The statement of (i) is a direct consequence of the last estimate and the identity

$$\frac{d}{dm} s = \frac{1}{2s} \frac{d}{dm} s^2.$$

We turn to the statement (ii): Differentiating the last identity yields

$$\frac{d^2}{dm^2} s = -\frac{1}{2} \frac{1}{s^2} \frac{d}{dm} s \frac{d}{dm} s^2 + \frac{1}{2s} \frac{d^2}{dm^2} s^2.$$

The estimation of the first term on the r.h.s. follows from the estimates

$$\left| \frac{d}{dm} s^2 \right| \lesssim s \quad \text{and} \quad \left| \frac{d}{dm} s \right| \lesssim 1,$$

which we have deduced in the first step of the proof. We turn to the estimation of the second term. A direct calculation using (85) yields the identity

$$\frac{d^2}{dm^2}s^2 = \frac{d^2}{dm^2}\frac{d}{d\sigma}m = \frac{d}{dm}\left(\frac{d^2}{d\sigma^2}m\frac{d}{dm}\sigma\right) = \frac{d^3}{d\sigma^3}m\left(\frac{d}{dm}\sigma\right)^2 + \frac{d^2}{d\sigma^2}m\frac{d^2}{dm^2}\sigma. \quad (88)$$

Considering the first term on the r.h.s. we get from the identities (85) and (87), and the assumption (34) of Proposition 3.1 that

$$\left|\frac{d^3}{d\sigma^3}m\left(\frac{d}{dm}\sigma\right)^2\right| = \frac{\int (x-m)^4 \mu^\sigma(dx)}{s^4} \lesssim 1.$$

Before we consider the second term of the r.h.s. of (88) we establish the following estimate:

$$\left|\frac{d^2}{dm^2}\sigma\right| \lesssim \frac{1}{s^3}. \quad (89)$$

Indeed, direct calculation using (85) and (86) yields

$$\begin{aligned} \frac{d^2}{dm^2}\sigma &= \left(\frac{d}{d\sigma}\frac{d}{dm}\sigma\right)\frac{d}{dm}\sigma \\ &= \left(\frac{d}{d\sigma}\left(\frac{d}{d\sigma}m\right)^{-1}\right)\left(\frac{d}{d\sigma}m\right)^{-1} \\ &= -\left(\frac{d}{d\sigma}m\right)^{-3}\frac{d^2}{d\sigma^2}m \\ &= -\frac{1}{s^3}\frac{\int (x-m)^3 \mu^\sigma(dx)}{s^3}. \end{aligned}$$

The last identity yields (89) using the assumption (34) of Proposition 3.1. Using (89) and (86) we can estimate the second term of the r.h.s. of (88) as

$$\left|\frac{d^2}{d\sigma^2}m\frac{d^2}{dm^2}\sigma\right| \lesssim \frac{1}{s^3}\left|\int (x-m)^3 \mu^\sigma(dx)\right|.$$

By applying the assumption (34) of Proposition 3.1 this yields

$$\left|\frac{d^2}{d\sigma^2}m\frac{d^2}{dm^2}\sigma\right| \lesssim 1,$$

which concludes the argument for (ii).  $\square$

*Proof of Lemma 3.3.* Recall the representation (31) i.e.

$$\tilde{g}_{K,m}(0) = \exp \left( K\varphi(m) - K\bar{H}_K(m) \right).$$

Here  $\tilde{g}_{K,m}(\xi)$  denotes the Lebesgue density of the random variable  $\frac{1}{\sqrt{K}} \sum_{i=1}^K (X_i - m)$ , where  $X_i$  are real-valued independent random variables identically distributed according to  $\mu^\sigma$  (cf. (32)). Let  $g_{K,\sigma}$  denote the density of the normalized random variable  $\frac{X}{s}$ , where  $s$  is given by (33). Then the densities are related by

$$\frac{1}{s} g_{K,\sigma} \left( \frac{x}{s} \right) = \tilde{g}_{K,m}(x).$$

It follows from (31) that

$$K\varphi(m) - K\bar{H}_K(m) = \log g_{K,\sigma}(0) - \log s.$$

In order to deduce the desired estimate it thus suffices to show

$$\left| \frac{d^2}{dm^2} \log s \right| \lesssim \frac{1}{s^2} \quad (90)$$

and

$$\left| \frac{d^2}{dm^2} \log g_{K,\sigma}(0) \right| \lesssim \frac{1}{s^2}. \quad (91)$$

The first estimate follows directly from the identity

$$\frac{d^2}{dm^2} \log s = \frac{d}{dm} \left( \frac{1}{s} \frac{d}{dm} s \right) = -\frac{1}{s^2} \left( \frac{d}{dm} s \right)^2 + \frac{1}{s} \frac{d^2}{dm^2} s$$

and the estimates provided by Lemma 3.7. We turn to the second estimate. The identity

$$\frac{d^2}{dm^2} \log g_{K,\sigma} = -\frac{1}{g_{K,\sigma}^2} \left( \frac{d}{dm} g_{K,\sigma} \right)^2 + \frac{1}{g_{K,\sigma}} \frac{d^2}{dm^2} g_{K,\sigma}$$

and (36) yield for large  $K$  the estimate

$$\left| \frac{d^2}{dm^2} \log g_{K,\sigma}(0) \right| \lesssim \left( \frac{d}{dm} g_{K,\sigma}(0) \right)^2 + \left| \frac{d^2}{dm^2} g_{K,\sigma}(0) \right|.$$

The estimation of the first term on the r.h.s. follows from the estimate (37) of Proposition 3.1 and the identity

$$\frac{1}{s} \frac{d}{d\sigma} = s \frac{d}{dm}, \quad (92)$$

which is a direct consequence of (61). Let us consider the second term. The identity

$$\left(\frac{1}{s} \frac{d}{d\sigma}\right)^2 \stackrel{(92)}{=} \left(s \frac{d}{dm}\right) \left(s \frac{d}{dm}\right) = s^2 \frac{d^2}{dm^2} + s \left(\frac{d}{dm} s\right) \frac{d}{dm},$$

which we rewrite as

$$s^2 \frac{d^2}{dm^2} = \left(\frac{1}{s} \frac{d}{d\sigma}\right)^2 - \left(\frac{d}{dm} s\right) \frac{1}{s} \frac{d}{d\sigma}$$

yields

$$\frac{d^2}{dm^2} g_{K,\sigma}(0) = \frac{1}{s^2} \left( \left(\frac{1}{s} \frac{d}{d\sigma}\right)^2 g_{K,\sigma}(0) - \left(\frac{d}{dm} s\right) \frac{1}{s} \frac{d}{d\sigma} g_{K,\sigma}(0) \right).$$

Now, the estimates (37) and (38) of Proposition 3.1 and Lemma 3.7 yield the desired estimate (91).  $\square$

## A. Standard criteria for the SG and the LSI

In this section we quote some standard criteria for the SG and the LSI. For a general introduction to the SG and the LSI we refer to [Led01, Roy99, GZ03]. Note that even if we only formulate the criteria on the level of the LSI, they also hold on the level of the SG. The first one shows that the LSI is compatible with products (cf. for example [GZ03, Theorem 4.4]).

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

The next criterion shows, how the LSI constant behaves under perturbations (cf. [HS87, p. 1184]). Unfortunately, the criterion is not well suited for high dimensions.

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

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

*If  $\mu$  satisfies the LSI with constant  $\varrho$ , then  $\tilde{\mu}$  satisfies the LSI with constant*

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

Because of its perturbative nature, the Holley-Stroock criterion is not well adapted for high dimensions. For the proof of the last statement, we refer the reader to [Led01, Lemma 1.2]. Now, we state the Bakry-Émery criterion, which connects the convexity of the Hamiltonian to the LSI constant (cf. [BE85, Proposition 3 and Corollary 2] or [Led01, Corollary 1.6]).

**Theorem A.3** (Bakry-Émery criterion). *Let  $d\mu := Z^{-1} \exp(-H(x)) dx$  be a probability measure on a Euclidean spaces  $X$ . If there is a constant  $\varrho > 0$  such that in the sense of quadratic forms*

$$\text{Hess } H(x) \geq \varrho$$

*uniformly in  $x \in X$ , then  $\mu$  satisfies the LSI with constant  $\varrho$ .*

A proof using semigroup methods can be found in [Led01, Corollary 1.6]. There is also a nice heuristic interpretation of the Bakry-Émery criterion on a formal Riemannian structure on the space of probability measures (cf. [OV00, Section 3]).

## Acknowledgement

GM and FO thank Franck Barthe, Michel Ledoux, and Cedric Villani for discussions on this subject. GM wants to thank the DFG for financial support through the Gottfried Wilhelm Leibniz program, the Bonn International Graduate School in Mathematics and the Max Planck Institute for Mathematics in the Sciences in Leipzig for partial support.

## References

- [BE85] D. Bakry and M. Émery. Diffusions hypercontractives. *Sem. Probab. XIX, Lecture Notes in Math., Springer*, 1123:177–206, 1985.
- [BL76] H. J. Brascamp and E. H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.*, 22(4):366–389, 1976.
- [Bob99] S. G. Bobkov. Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.*, 27(4):1903–1921, 1999.
- [BW09] F. Barthe and P. Wolff. Remarks on non-interacting conservative spin systems: the case of gamma distributions. *Stochastic Process. Appl.*, 119(8):2711–2723, 2009.

- [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.
- [Csi67] I. Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.*, 2:299–318, 1967.
- [EG92] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [Fel71] W. Feller. *An introduction to probability theory and its applications. Vol II*. Wiley Series in Probability and Mathematical Statistics. Wiley, 2nd edition, 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. H. 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. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.
- [HS87] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.*, 46:1159–1194, 1987.
- [KL99] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems.*, volume 320 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, 1999.
- [Kos01] E. Kosygina. The behavior of the specific entropy in the hydrodynamic scaling limit. *Ann. Probab.*, 29(3):1086–110, 2001.

- [Kul67] S. Kullback. A lower bound for discrimination information in terms of variation. *IEEE Trans. Inform.*, 4:126 – 127, 1967.
- [Led01] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. *Sem. Probab. XXXV, Lecture Notes in Math.*, Springer, 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.
- [Men10] G. Menz. LSI for Kawasaki dynamics with weak interaction. *MPI-MIS preprint, Leipzig*, 31, 2010.
- [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 Spécialisés, Soc. Math. de France*, 1999.
- [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.
- [Var93] S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. II. In *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283 of *Pitman Res. Notes Math. Ser.*, pages 75–128. Longman Sci. Tech., Harlow, 1993.
- [Yau96] H. T. Yau. Logarithmic Sobolev inequality for lattice gases with mixing conditions. *Comm. Math. Phys.*, 181(2):367–408, 1996.

- [Yos99] N. Yoshida. The log-Sobolev inequality for weakly coupled lattice fields. *Probab. Theory Related Fields*, 115(1):1–40, 1999.
- [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.