Delft University of Technology

Relative Entropy, Gaussian Concentration and Uniqueness of Equilibrium States

Chazottes, Jean-René ; Redig, F.H.J.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Article

Relative Entropy, Gaussian Concentration and Uniqueness of Equilibrium States

Jean-René Chazottes [1,*,†] [ID] **and Frank Redig** [2,†]

1   Centre de Physique Théorique, Centre National de la Recherche Scientifique, Institut Polytechnique de Paris, 91128 Palaiseau, France
2   Institute of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands
*   Correspondence: jeanrene@cpht.polytechnique.fr
†   These authors contributed equally to this work.

Abstract: For a general class of lattice spin systems, we prove that an abstract Gaussian concentration bound implies positivity of the lower relative entropy density. As a consequence, we obtain uniqueness of translation-invariant Gibbs measures from the Gaussian concentration bound in this general setting. This extends earlier results with a different and very short proof.

1. Introduction

1.1. Uniqueness Criteria for Gibbs Measures

In mathematical statistical mechanics, it is important to have good and useful criteria for the absence of phase transition, or equivalently, uniqueness of the Gibbs measure associated with a given potential. Such criteria, also known under the name high-temperature criteria, show that when the interaction is small enough (high temperature), there is no phase transition, and the unique phase has strong mixing properties, i.e., it is close to a product measure (infinite temperature).

The most famous among such criteria is the Dobrushin uniqueness criterion; see, e.g., ([1], Chapter 8). Under the Dobrushin uniqueness criterion, besides uniqueness, one derives strong mixing properties of the unique Gibbs measure, i.e., quantitative bounds on the decay of covariance of local observables, and quantitative bounds on the difference between finite and infinite-volume expectations, i.e., on the influence of the boundary condition on the expectation of a local function. The basic idea behind the Dobrushin uniqueness criterion is that, when it holds, the conditional expectation operator acts as a contraction on the space of probability measures equipped with the Wasserstein distance. Because the Gibbs measure is a fixed point of this contraction, and fixed points of contractions are unique, one obtains uniqueness of Gibbs measures. Later on, the Dobrushin criterion was generalized to the Dobrushin–Shlosman criterion, and a connection has been made between this criterion and an important functional inequality, the log-Sobolev inequality. More precisely, for finite-range Glauber dynamics of Ising spins, in [2] the equivalence between the Dobrushin-Shlosman criterion and the log-Sobolev inequality was proved. This implies e.g., that under the Dobrushin–Shlosman criterion, the reversible Glauber dynamics converges exponentially fast (in L^∞) to its unique stationary measure.

Related to the Dobrushin criterion, there is a general criterion in the context of interacting particle systems under which one obtains uniqueness of the stationary measure and uniform ergodicity, i.e., from any initial measures, in time, the dynamics converges exponentially fast to the unique stationary measures. This criterion, the so-called "$M < \varepsilon$ criterion" ([3], Chapter 1), is based on a similar contraction argument, i.e., when it holds, the semigroup of the interacting particle system acts as a contraction on a suitable space

of functions, equipped with a norm (the so-called triple (semi)-norm) which controls the oscillations of a function, and the semigroup acts as a contraction in this norm. Similar to the setting of the unique measures in the setting of the Dobrushin uniqueness, also under the $M < \varepsilon$ criterion, one obtains strong mixing properties of the unique stationary measure.

In the context of probabilistic cellular automata, as well as in the context of Glauber dynamics, the $M < \varepsilon$ criterion is shown to be equivalent to the Dobrushin uniqueness criterion for the space–time Gibbs measure; see [4,5].

1.2. Concentration Inequalities

Concentration inequalities are inequalities in which one estimates the deviation between a function $f = f(\sigma_1, \ldots, \sigma_n)$ of n random variables $\sigma_1, \ldots, \sigma_n$ and its expectation $\mathbb{E}(f)$. The idea is that whenever the function depends only weakly on individual variables σ_i, and the distribution of $(\sigma_1, \ldots, \sigma_n)$ is a product, or close to a product, then the probability that f deviates from $\mathbb{E}(f)$ becomes very small. To measure the dependence of f of individual coordinates, one considers, e.g., the oscillation

$$\delta_i f = \sup_{\eta_j = \sigma_j, \forall j \neq i} \big(f(\sigma) - f(\eta) \big).$$

An important example of a concentration inequality is the so-called Gaussian concentration bound

$$\mathbb{P}(|f - \mathbb{E}(f)| > \varepsilon) \leq 2 \exp\left(-\frac{C \varepsilon^2}{\sum_{i=1}^{n} (\delta_i f)^2}\right), \ \varepsilon > 0, \tag{1}$$

where $C > 0$ is a constant which does not depend on f, and in particular, it does not depend on n. For instance, if $f(\sigma) = \frac{1}{n} \sum_{i=1}^{n} \sigma_i$, and $\sigma_i \in \{-1, 1\}$, then $\delta_i f = 2/n$, and we find the upper bound $2\,e^{-Cn\varepsilon^2/4}$ for all $n \geq 1$. The power of concentration inequalities of the type (1) is that they hold for general f, i.e., far beyond empirical averages.

Concentration inequalities in the context of Gibbs measures for lattice spin systems have been studied in several works. In particular, in [6] the author proves an inequality of the type (1) under the Dobrushin uniqueness criterion. See [7] for a recent overview of concentration inequalities in the context of Gibbs measures.

1.3. Concentration and Uniqueness

The central question of this paper is the following. Assume that a Gibbs measure associated with a given potential satisfies a Gaussian concentration bound, i.e., an inequality of the type (1). Can we then conclude that it is the unique Gibbs measure, i.e., that there is no phase transition?

In this paper, we restrict ourselves to translation-invariant Gibbs measures (i.e., so-called equilibrium states), but in a very general setting. Following [8], we started in [9] the study of the relation between the Gaussian concentration bound and the uniqueness of equilibrium states in the context of spin systems on the lattice $\mathbb{Z}^d$, where the spin at each lattice site takes a finite number of values. Examples there include the Ising model at high temperature. Notice that for this model, at low temperature in $d \geq 2$, there is a phase transition, and the large deviation probabilities of the magnetization are surface-like, rather than volume-like. This manifestation of a phase transition excludes the Gaussian concentration bound, under which all ergodic averages have volume-like large-deviation probabilities.

Here, we show uniqueness of equilibrium states under an inequality of the type (1), and next, we generalize both the context of the concentration inequality, as well as the context of Gibbs measures, showing uniqueness in the context of so-called zero-information sets. An important result in the context of equilibrium states is the variational principle, which implies that the relative entropy density between two equilibrium states is zero. Therefore, if one can show a strictly positive lower bound for relative entropy density, one obtains uniqueness of the equilibrium states. The set of equilibrium states associated with a given translation-invariant potential is a special case of a set in which the relative

entropy density between two elements of the set is always zero. We call such a set a zero-information set, and generalize our results of uniqueness to this context, which includes, e.g., transformations of Gibbs measures, and stationary measures of certain interacting particle systems.

1.4. Content and Organization of the Paper

As sketched above, we obtain a lower bound for the lower relative entropy density in terms of a natural distance between translation-invariant probability measures, reminiscent and in the spirit of the results of Bobkov and Götze [10], who proved (in a different setting) a lower bound for the relative entropy in terms of the square of the Wasserstein distance. Because we work in the thermodynamic limit on a product space and are interested in translation-invariant probability measures, there is no translation-invariant distance on the configuration space for which we can apply the Bobkov-Götze theorem. We can avoid this problem by introducing a suitable distance on the translation-invariant probability measures (rather than on configurations).

We start by proving the lower bound on the lower relative entropy density in the context of general lattice spin systems with state space $\Omega = S^{\mathbb{Z}^d}$, where the single spins take values in a metric space S of bounded diameter. The bounded diameter property allows us to associate a quasi-local function f with a natural sequence of oscillations $\delta_i f, i \in \mathbb{Z}^d$, where $\delta_i f$ represents the maximal influence on the function f of a change in the spin at site i. In the final section of this paper, we provide a generalization of this by allowing more abstract single-spin spaces, and more general associated sequences of oscillations.

The rest of our paper is organized as follows. In Section 2, we introduce the basic setting of lattice spin systems and important function spaces. In Section 3, we introduce the Gaussian concentration bound, the relative entropy (density), and formulate and prove our main result in the context of a single-spin space with finite diameter. In Section 4, we discuss applications of our result to zero information distance sets, including, e.g., the set of equilibrium states with regard to absolutely summable translation-invariant potentials. In Section 5, we consider a generalization by introducing an abstract sequence of oscillations, the associated Gaussian concentration bound and state, and prove the analogue of our main result in this generalized context.

2. Setting

2.1. Configuration Space and the Translation Operator

We start from a standard Borel space $(S, \mathfrak{b})$ with metric d_S, and we let $\text{diam}(S) = \sup_{s,s' \in S} d_S(s, s')$. (A measurable space $(S, \mathfrak{b})$ is said to be standard Borel if there exists a metric on S which makes it a complete separable metric space, and $\mathfrak{b}$ then denotes the associated Borel σ-algebra.) In the sequel, for notational convenience, we omit the symbol $\mathfrak{b}$ and call S a standard Borel space, where we always assume that the associated σ-algebra is the Borel σ-algebra $\mathfrak{b}$.

We assume that $\text{diam}(S) < \infty$. Later on, in Section 5, we will show how to weaken this assumption.

This space S represents the "single-spin space", i.e., we will consider lattice spin configurations in which individual "spins" take values in S. We denote by $(\Omega_\Lambda, \mathfrak{b}_\Lambda)$ the product space $(S^\Lambda, \prod_{i \in \Lambda} \mathfrak{b}_i)$, and $(\Omega, \mathfrak{B})$ stands for the lattice spin configuration space $(\Omega_{\mathbb{Z}^d}, \mathfrak{b}_{\mathbb{Z}^d})$. We equip this space with the product topology. Elements of Ω are called configurations. For $\eta \in \Omega$, we denote by $\eta_i \in S$ its evaluation at site $i \in \mathbb{Z}^d$. By σ_Λ we mean an element of Ω_Λ, and by $\eta_\Lambda \xi_{\Lambda^c}$, a configuration coinciding with η on Λ and with ξ on Λ^c. We denote by $\mathbb{S}$ the set of finite subsets of $\mathbb{Z}^d$.

We denote by $\tau_i : \mathbb{Z}^d \to \mathbb{Z}^d, i \in \mathbb{Z}^d$, the map which shifts, or translates, by i; that is, $\tau_i(j) = j + i, j \in \mathbb{Z}^d$, and for $\Lambda \subset \mathbb{Z}^d$, we write $\tau_i(\Lambda) = \Lambda + i = \{j \in \mathbb{Z}^d : j = k + i, k \in \Lambda\}$, $\Lambda \subset \mathbb{Z}^d$. We define the translation operator acting on configurations as follows (and use the same symbol). For each $i \in \mathbb{Z}^d, (\tau_i \sigma)_j = \sigma_{j-i}$, for all $j \in \mathbb{Z}^d$. This corresponds to translating σ forward by i. We denote by the same symbol the translation operator acting on a function

$f : \Omega \to \mathbb{R}$. For each $i \in \mathbb{Z}^d$, $\tau_i f$ is the function defined as $\tau_i f(\sigma) = f(\tau_i \sigma)$. A (Borel) probability measure on Ω is translation invariant if, for all $B \in \mathfrak{B}$ and for all $i \in \mathbb{Z}^d$, we have $\mu(\tau_i B) = \mu(B)$.

We denote by $\mathcal{P}_\tau(\Omega)$ the set of translation-invariant probability measures on Ω. We denote by $\mathscr{C}(\Omega), \mathscr{C}_b(\Omega)$ the space of continuous, respectively bounded continuous, real-valued functions on Ω.

2.2. Local Oscillations and Function Spaces

To a continuous function $f : \Omega \to \mathbb{R}$ we associate a "sequence" of "local oscillations", $\delta f := (\delta_i f)_{i \in \mathbb{Z}^d}$, defined via

$$\delta_i f = \sup_{\substack{\sigma, \eta \in \Omega: \\ \sigma_j = \eta_j, \forall j \neq i}} \big(f(\sigma) - f(\eta) \big). \tag{2}$$

Later on, in Section 5, in which we consider the case where S is allowed to have infinite diameter, we will consider a more abstract definition of δf. In the case where S has finite diameter, (2) is the most natural choice.

For integer $p \geq 1$, we define the usual ℓ^p-norm of δf, $\|\delta f\|_p^p = \sum_{i \in \mathbb{Z}^d} (\delta_i f)^p$, and $\|\delta f\|_\infty = \sup_{i \in \mathbb{Z}^d} \delta_i f$.

We call a continuous function local if $\delta_i f \neq 0$ for finitely many $i \in D_f \subset \mathbb{Z}^d$. The set D_f is then called the dependence set of f. We denote by $\mathcal{L}(\Omega)$ the set of local continuous functions on Ω.

We call a continuous function quasi-local if it is the uniform limit of a sequence of local continuous functions. If S is compact, then, according to the Stone–Weierstrass theorem, local continuous functions are uniformly dense in $\mathscr{C}(\Omega)$.

We denote by $\mathcal{QL}(\Omega)$ the space of all continuous quasi-local functions on Ω.

For $1 \leq p \leq \infty$, we introduce the spaces

$$\Delta^p(\Omega) = \{ f \in \mathscr{C}(\Omega) : \|\delta f\|_p < \infty \}.$$

Lemma 1. *If $f \in \mathcal{QL}(\Omega) \cap \Delta^1(\Omega)$, then f is bounded. If $f \in \mathcal{L}(\Omega) \cap \Delta^p(\Omega)$ for some $p \geq 1$, then f is bounded.*

Proof. Choose $\eta, \sigma \in \Omega$. We have for every $\Lambda \in \mathcal{S}$

$$|f(\eta) - f(\sigma)| \leq |f(\eta) - f(\eta_\Lambda \sigma_{\Lambda^c})| + \sum_{i \in \Lambda} \delta_i f \tag{3}$$

which, upon taking the limit $\Lambda \uparrow \mathbb{Z}^d$, using the assumed quasilocality of f gives

$$|f(\eta) - f(\sigma)| \leq \sum_{i \in \mathbb{Z}^d} \delta_i f = \|\delta f\|_1 < \infty. \tag{4}$$

If f is local, then we still have the inequality (3) for Λ containing the dependence set of f. Because, by assumption, $\sum_i (\delta_i f)^p$ is finite, it follows that $\delta_i f < \infty$ for all $i \in \Lambda$, and therefore $\|\delta f\|_1 < \infty$. Then, we obtain (4), which implies that f is bounded. $\square$

We say that $\mu_n \to \mu$ if, for all bounded continuous local functions, we have $\int f \, d\mu_n \to \int f \, d\mu$ (then by definition of quasilocality, the same holds for bounded continuous quasi-local functions). This induces the so-called weak quasi-local topology on probability measures. Notice that in our setting, where by assumption the single-spin space S is a complete separable metric space, this topology coincides with the ordinary weak topology; see [11] (p. 898).

In our setting, the set of bounded quasi-local continuous functions is measure separating, i.e., for two probability measures $\mu \neq \mu'$, there exists a bounded quasi-local continuous f, such that

$$\int f \, \mathrm{d}\mu \neq \int f \, \mathrm{d}\mu'.$$

Because, by definition, bounded continuous quasi-local functions can be uniformly approximated by bounded continuous local functions, if the set of bounded quasi-local functions is measure separating, then the set of bounded continuous local functions is also measure separating. Therefore, in our setting, for two probability measures $\mu \neq \mu'$, there exists a bounded local f (which is not constant), such that

$$\int f \, \mathrm{d}\mu \neq \int f \, \mathrm{d}\mu'.$$

This can be seen as follows. If $\mu \neq \mu'$, then there exists a bounded closed cylindrical set $A \subset \Omega$, such that $\mu(A) \neq \mu'(A)$, because the Borel σ-algebra on Ω is generated by such sets. The indicator of this set can be approximated by bounded local continuous functions in both $L^1(\mu)$ and $L^1(\mu')$.

3. Gaussian Concentration Bound and Relative Entropy

3.1. Abstract Gaussian Concentration Bound

We can now give the definition of the Gaussian concentration bound in our setting.

Definition 1. *Let $\Omega = S^{\mathbb{Z}^d}$, where S is a standard Borel space with a finite diameter. Let μ be a probability measure on Ω. We say that μ satisfies the Gaussian concentration bound with constant $C > 0$, abbreviated $\mathrm{GCB}(C)$, if for all bounded local functions f we have*

$$\int \mathrm{e}^{f - \int f \, \mathrm{d}\mu} \, \mathrm{d}\mu \leq \mathrm{e}^{\frac{C}{2}\|\delta f\|_2^2}. \tag{5}$$

Remark 1.
(a) *Observe that the bound (5) does not change if f is replaced by $f + c$, where $c \in \mathbb{R}$ is arbitrary, since $\delta_i(f + c) = \delta_i(f)$ for any $i \in \mathbb{Z}^d$. This "insensitivity" to constant offsets on the left-hand side is ensured by the fact that we center f around its expected value. We also observe that (5) is trivially true for functions which are constant.*
(b) *We have $\delta_i(\beta f) = |\beta| \, \delta_i f$, for all $i \in \mathbb{Z}^d$ and $\beta \in \mathbb{R}$; we thus have*

$$\log \int \mathrm{e}^{\beta(f - \int f \, \mathrm{d}\mu)} \, \mathrm{d}\mu \leq \frac{C\beta^2}{2}\|\delta f\|_2^2, \ \forall \beta \in \mathbb{R}.$$

This quadratic upper bound will be crucial in the sequel.
(c) *The quadratic nature of the upperbound in (5) resembles the quadratic upperbound for the pressure in [12], Theorem 1.1, Equation (2.7), in terms of the Dobrushin norm. This suggests that in the Dobrushin uniqueness regime, the quadratic bound which is obtained from (5) might also be obtainable from this result. However, the Gaussian concentration inequality does not require the Dobrushin uniqueness condition; the latter is sufficient, but not necessary.*

The following proposition asserts that (5) automatically extends to a wider class of functions.

Proposition 1 (Self-enhancement of GCB). *Suppose that (5) holds for all bounded local f. Then, it holds for all $f \in \mathcal{QL}(\Omega) \cap \Delta^2(\Omega)$.*

Proof. By assumption, for a fixed $\xi \in \Omega$, $f \in \mathcal{QL}(\Omega) \cap \Delta^2(\Omega)$ can be uniformly approximated by the local functions

$$f_{\Lambda, \xi}(\eta) = f(\eta_\Lambda \xi_{\Lambda^c}).$$

By definition (2), $\delta_i f_{\Lambda,\xi}$ is non-decreasing when Λ grows, and is bounded by $\delta_i f$. According to Lemma 1, it follows that $f_{\Lambda,\xi}$ is bounded. Therefore $\|\delta f_{\Lambda,\xi}\|_2$ is bounded by $\|\delta f\|_2$, which is finite because $f \in \Delta^2(\Omega)$. Therefore, using the assumed uniform convergence of $f_{\Lambda,\xi}$ to f, and the assumed bound (5) for bounded local functions in $\Delta^2(\Omega)$, we obtain

$$\int e^{f - \int f \, d\mu} \, d\mu = \lim_{\Lambda \uparrow \mathbb{Z}^d} \int e^{f_{\Lambda,\xi} - \int f_{\Lambda,\xi} \, d\mu} \, d\mu \leq \lim_{\Lambda \uparrow \mathbb{Z}^d} e^{\frac{C}{2} \|\delta f_{\Lambda,\xi}\|_2^2} = e^{\frac{C}{2} \|\delta f\|_2^2}.$$

Here, in the first equality, we used the uniform convergence of $f_{\Lambda,\xi}$ to f. In the last equality, we used $\delta_i f_{\Lambda,\xi} \leq \delta_i f$, and by assumption $\sum_i (\delta_i f)^2 < \infty$, so by dominated convergence applied to the counting measure on $\mathbb{Z}^d$, we have

$$\lim_{\Lambda \uparrow \mathbb{Z}^d} \sum_{i \in \mathbb{Z}^d} (\delta_i f_{\Lambda,\xi})^2 = \|\delta f\|_2^2.$$

$\square$

3.2. Relative Entropy

For a probability measure μ, we denote by μ_Λ its restriction to the sub-σ-algebra $\mathfrak{B}_\Lambda = \sigma\{\eta_i, i \in \Lambda\}$, generated by the projection $p_\Lambda : \Omega \to \Omega_\Lambda$. We also denote by $\mathcal{B}_\Lambda$ the set of bounded $\mathfrak{B}_\Lambda$-measurable functions from Ω to $\mathbb{R}$.

For two probability measures μ, μ' on Ω and $\Lambda \in \mathcal{S}$, we define the relative entropy of μ' with respect to μ by

$$s_\Lambda(\mu'|\mu) = \begin{cases} \int d\mu'_\Lambda \log \dfrac{d\mu'_\Lambda}{d\mu_\Lambda} & \text{if} \quad \mu'_\Lambda \ll \mu_\Lambda \\ +\infty & \text{otherwise.} \end{cases}$$

We further denote by $(\Lambda_n)_{n \in \mathbb{N}}$ the sequence of "cubes" $\Lambda_n = [-n, n]^d \cap \mathbb{Z}^d$, $n \geq 1$.

Definition 2 (Lower relative entropy density). *For two probability measures μ, μ' on Ω, we define the* lower relative entropy density *by*

$$s_*(\mu'|\mu) = \liminf_{n \to +\infty} \frac{s_{\Lambda_n}(\mu'|\mu)}{|\Lambda_n|}.$$

We have the following variational characterization of relative entropy (for proof, see, for instance, [13] (p. 100))

$$s_\Lambda(\mu'|\mu) = \sup_f \left(\int f \, d\mu'_\Lambda - \log \int e^f \, d\mu_\Lambda \right) \tag{6}$$

where the supremum is taken over all $\mathfrak{B}_\Lambda$-measurable functions, such that $\int e^f \, d\mu_\Lambda < \infty$.

3.3. Main Result

In the main theorem below, we prove that the Gaussian concentration bound implies strict positivity of the lower relative entropy density. Introducing an appropriate metric on the set of probability measures, we show that the lower relative entropy density is lower bounded by a constant multiplied by the square of this distance. This result substantially generalizes the corresponding result from [9], where it is essential that the single-spin space is finite. Moreover, the proof is simpler and based on the variational formula for the relative entropy, combined with a quadratic estimate for the log-moment-generating function coming from the assumed Gaussian concentration bound.

Definition 3. *Define the following distance between probability measures*

$$d(\mu, \mu') = \sup \left\{ \int f \, d\mu - \int f \, d\mu' : f \in \mathcal{L}(\Omega), \|\delta f\|_1 \le 1 \right\}. \tag{7}$$

The metric defined above generates the quasi-local topology, and therefore convergence in this metric implies weak convergence. Indeed, convergence $\mu_n \to \mu$ in the metric d clearly implies $\int f \, d\mu_n \to \int f \, d\mu$ for all local continuous f, and hence also for all quasi-local continuous f. The latter implies $\mu_n \to \mu$ in the quasi-local topology, which coincides with the weak topology.

We can then formulate our main result.

Theorem 1. *If μ is translation invariant and satisfies* GCB(C), *then for all μ' translation invariant, and $\mu' \ne \mu$, we have*

$$s_*(\mu'|\mu) > 0.$$

More precisely, we have

$$s_*(\mu'|\mu) \ge \frac{d(\mu', \mu)^2}{2C}, \tag{8}$$

where d is the distance (7).

We start with a lemma from [7]. For the reader's convenience, we repeat the short proof here.

Lemma 2. *For f, such that $\|\delta f\|_1 < +\infty$ and $\Lambda \in \mathcal{S}$, we have*

$$\left\| \delta \left(\sum_{i \in \Lambda} \tau_i f \right) \right\|_2^2 \le |\Lambda| \|\delta f\|_1^2.$$

Proof. For $\Lambda \subset \mathbb{Z}^d$, let $\mathbb{1}_\Lambda$ denote the indicator function of Λ (that is, $\mathbb{1}_\Lambda(i) = 1$ if $i \in \Lambda$ and $\mathbb{1}_\Lambda(i) = 0$ otherwise). Then, for every $j \in \mathbb{Z}^d$ we have

$$0 \le \delta_j \left(\sum_{i \in \Lambda} \tau_i f \right) \le \sum_{i \in \mathbb{Z}^d} (\delta_{i+j} f) \mathbb{1}_\Lambda(i) = (\delta f * \mathbb{1}_\Lambda)_j.$$

As a consequence, using Young's inequality for convolutions, we obtain

$$\left\| \delta \left(\sum_{i \in \Lambda} \tau_i f \right) \right\|_2^2 \le \|\delta f * \mathbb{1}_\Lambda\|_2^2 \le \|\mathbb{1}_\Lambda\|_2^2 \|\delta f\|_1^2 = |\Lambda| \|\delta f\|_1^2.$$

$\square$

Proof of Theorem 1. For the cube Λ_n and a bounded local function f whose dependence set is included in the cube Λ_r, for some r, it follows from (6) that

$$\frac{s_{\Lambda_{n+r}}(\mu'|\mu)}{|\Lambda_n|} \ge \frac{1}{|\Lambda_n|} \left(\int \sum_{i \in \Lambda_n} \tau_i f \, d\mu' - \log \int e^{\sum_{i \in \Lambda_n} \tau_i f} \, d\mu \right)$$

where we used that $\sum_{i \in \Lambda_n} \tau_i f$ is measurable with respect to $\mathfrak{B}_{\Lambda_{n+r}}$. Now, if μ satisfies GCB(C) and both μ' and μ are translation invariant, then we can estimate further as follows. Start by noticing that, through combination of the assumed GCB(C) and Lemma 2, we have

$$\log \int e^{\sum_{i \in \Lambda_n} (\tau_i f - \int f \, d\mu)} \, d\mu \le \frac{C}{2} |\Lambda_n| \|\delta f\|_1^2.$$

As a consequence, using translation invariance of both μ and μ', we obtain

$$
\begin{aligned}
&\frac{s_{\Lambda_{n+r}}(\mu'|\mu)}{|\Lambda_n|} \\
&\geq \frac{1}{|\Lambda_n|}\left(\int \sum_{i\in\Lambda_n} \tau_i f \,\mathrm{d}\mu' - \log \int \mathrm{e}^{\sum_{i\in\Lambda_n} \tau_i f} \,\mathrm{d}\mu\right) \\
&= \frac{1}{|\Lambda_n|}\left(\int \sum_{i\in\Lambda_n} \tau_i f \,\mathrm{d}\mu' - \int \sum_{i\in\Lambda_n} \tau_i f \,\mathrm{d}\mu - \log \int \mathrm{e}^{\sum_{i\in\Lambda_n}(\tau_i f - \int \tau_i f \,\mathrm{d}\mu)} \,\mathrm{d}\mu\right) \\
&\geq \int f \,\mathrm{d}\mu' - \int f \,\mathrm{d}\mu - \frac{C}{2}\|\delta f\|_1^2.
\end{aligned} \tag{9}
$$

Consider a bounded local function f, such that $\int f \,\mathrm{d}\mu' - \int f \,\mathrm{d}\mu \geq u > 0$ (this function exists by the assumption that bounded local functions are measure separating). Put $\|\delta f\|_1^2 =: \varrho$. (Observe that $\varrho < \infty$ by assumption, and $\varrho \neq 0$ since f cannot be a constant.) Assume that the dependence set of f is included in the cube Λ_r. Replace f by βf in the inequality (9), and optimize over β. Then, we obtain, for all $n \in \mathbb{N}$, the inequality

$$
\begin{aligned}
\frac{s_{\Lambda_{n+r}}(\mu'|\mu)}{|\Lambda_n|} &\geq \sup_{\beta\geq 0}\left\{\beta\left(\int f \,\mathrm{d}\mu' - \int f \,\mathrm{d}\mu\right) - \frac{C}{2}\beta^2\|\delta f\|_1^2\right\} \\
&\geq \sup_{\beta\geq 0}\left\{\beta u - \frac{C}{2}\beta^2\varrho\right\} = \frac{u^2}{2C\varrho} > 0.
\end{aligned} \tag{10}
$$

Since r is fixed, we can take the limit inferior in n, and using $|\Lambda_n|/|\Lambda_{n+r}| \to 1$ as $n \to \infty$, we obtain

$$
\liminf_{n\to+\infty} \frac{s_{\Lambda_n}(\mu'|\mu)}{|\Lambda_n|} > 0.
$$

From (10), we infer that for f, such that

$$
\int f \,\mathrm{d}\mu' - \int f \,\mathrm{d}\mu \geq u, \text{ and } \|\delta f\|_1^2 \leq \varrho
$$

we have

$$
s_*(\mu'|\mu) \geq \frac{1}{2C}\left(\frac{u}{\sqrt{\varrho}}\right)^2.
$$

Therefore, for

$$
\int f \,\mathrm{d}\mu' - \int f \,\mathrm{d}\mu \geq \varepsilon \quad \text{and} \quad \|\delta f\|_1^2 \leq 1
$$

we have

$$
s_*(\mu'|\mu) \geq \frac{\varepsilon^2}{2C}. \tag{11}
$$

By definition of the distance (7), this is equivalent with the statement that $d(\mu',\mu) \geq \varepsilon$ implies (11). This implies (8). $\quad\square$

The following corollary shows that convergence in relative entropy density implies convergence in the distance d. This can be used for stochastic dynamics, provided one can show that the relative entropy density converges. See the application section below for some examples.

Corollary 1. *Let μ be a translation-invariant probability measure which satisfies* $\mathrm{GCB}(C)$. *Assume* $(\mu_n)_{n\geq 1}$ *is a sequence of translation-invariant probability measures, such that*

$$
\lim_{n\to+\infty} s_*(\mu_n|\mu) = 0. \tag{12}
$$

Then, $\mu_n \to \mu$ in the sense of the distance (7), and therefore also $\mu_n \to \mu$ weakly.

Proof. By (8), (12) implies

$$\lim_{n \to +\infty} d(\mu_n, \mu)^2 \leq 2C \lim_{n \to +\infty} s_*(\mu_n | \mu) = 0.$$

Therefore, we have convergence in the metric d, which, as we remarked before, implies weak convergence. □

Remark 2. *As an example of application of Corollary (1), we mention the iteration of renormalization group transformations in the high-temperature regime [14], where convergence of the renormalized potentials can be established, and as a consequence, we obtain convergence of the relative entropy density. Then, Corollary (1) implies that the renormalized measures converge in the metric d as least as fast as the potentials. In the context of stochastic dynamics, i.e., where μ_n is a time-evolved measure (at time n), it is usually not simple to obtain the convergence $s_*(\mu_n | \mu) \to 0$. In the high-temperature setting (high-temperature dynamics, high-temperature initial measure) this can be obtained with similar means as in [14].*

We conclude this section with two further remarks relating our result to the Bobkov-Götze criterion.

Remark 3. *Our distance $d(\mu, \mu')$ between probability measures resembles the so-called Dobrushin distance, denoted by $D(\mu, \mu')$, which consists of taking the supremum of $\int f \, d\mu - \int f \, d\mu'$ over a wider set of functions. Namely, f is required to be measurable, and such that $\|\delta f\|_1 \leq 1$. Hence $d(\mu, \mu') \leq D(\mu, \mu')$ for a general pair μ, μ' of probability measures. In the special case of finite S, one has $d = D$. In [15], it is proved that D is equal to what the authors called the Steiff distance $\bar{d}$, which is defined in terms of couplings, and which generalizes the Ornstein distance. The equality between D and $\bar{d}$ is reminiscent of the Kantorovich–Rubinstein duality theorem.*

Remark 4. *Inequality (8) is reminiscent of a well-known abstract inequality relating the relative entropy and the Wasserstein distance due to Bobkov and Götze [10]. However, our context is different, because we consider the thermodynamic limit and the relative entropy density. Nevertheless, as shown in [7], we can exploit the Bobkov-Götze theorem in the special case of finite S, putting the Hamming distance on S^{Λ_n}, to get*

$$s_*(\nu | \mu) \geq \frac{\bar{d}^2(\mu, \nu)}{2C},$$

where

$$\bar{d}(\mu, \nu) := \lim_{n \to +\infty} \frac{W_1(\mu_{\Lambda_n}, \nu_{\Lambda_n}; \bar{d}_n)}{|\Lambda_n|}$$

and where $\bar{d}_n(\omega, \eta) = \sum_{i \in \Lambda_n} \mathbb{1}_{\{\omega_i \neq \eta_i\}}$, $W_1(\mu_{\Lambda_n}, \nu_{\Lambda_n}; \bar{d}_n)$ is the Wassertein distance between μ_{Λ_n} and ν_{Λ_n}.

4. Applications: Uniqueness of Equilibrium States and Beyond

In this subsection, we provide some settings where we can conclude uniqueness of a set of "(generalized) translation-invariant Gibbs measures" via Theorem 1. We start with the set of translation-invariant Gibbs measures associated with an absolutely summable potential. Then, we consider generalizations and modifications of such sets.

4.1. Uniqueness of Equilibrium States

In this subsection, we briefly introduce the necessary basics of Gibbs measures. The reader familiar with the theory of Gibbs measure can skip this subsection. The reader is referred to [1] (especially Chapter 16) or [11] (Chapter 2) for more background on the Gibbs formalism.

Let λ be a probability measure on S, and for $\lambda_\Lambda(d\sigma_\Lambda) = \otimes_{i\in\Lambda}\lambda(d\sigma_i)$ the corresponding product measure on S^Λ. The measure λ is called the "a priori" measure on S, with associated a priori measure $\otimes_{i\in\mathbb{Z}^d}\lambda(d\sigma_i)$ on Ω.

We call a uniformly absolutely summable translation-invariant potential a function

$$U : \mathcal{S} \times \Omega \to \mathbb{R}$$

with the following properties:

(a) Locality: for all $A \in \mathcal{S}$, $U(A, \cdot)$ is $\mathfrak{B}_A$-measurable and continuous.
(b) Absolute summability: $\sum_{A \ni 0} \|U(A, \cdot)\|_\infty < \infty$.
(c) Translation invariance: $U(A + i, \tau_i\sigma) = U(A, \sigma)$ for all $\sigma \in \Omega$, $A \in \mathcal{S}$.

Let us call $\mathcal{U}$ the set of uniformly absolutely summable translation-invariant potentials. Then, we build the local Gibbs measures with boundary condition $\xi \in \Omega$. For a finite subset $\Lambda \in \mathcal{S}$, the Gibbs measure in volume Λ with boundary condition ξ outside Λ is defined via

$$\gamma_\Lambda(d\sigma_\Lambda|\xi) = \frac{e^{-H_\Lambda^\xi(\sigma_\Lambda)}}{Z_\Lambda^\xi}\lambda_\Lambda(d\sigma_\Lambda)$$

where H_Λ^ξ is the Hamiltonian in volume Λ with boundary condition ξ:

$$H_\Lambda^\xi(\sigma_\Lambda) = \sum_{A\cap\Lambda\neq\varnothing} U(A, \sigma_\Lambda\xi_{\Lambda^c})$$

and where Z_Λ^ξ is the normalization

$$Z_\Lambda^\xi = \int e^{-H_\Lambda^\xi(\sigma_\Lambda)}\lambda_\Lambda(d\sigma_\Lambda).$$

The family $(\gamma_\Lambda(d\sigma_\Lambda|\cdot))_{\Lambda\in\mathcal{S}}$ is called the Gibbsian specification associated with the potential U (with a priori measure λ).

By the uniform absolute summability of U, we automatically have that for all f local and continuous, the function $\xi \mapsto \int f(\sigma_\Lambda\xi_{\Lambda^c})\,\gamma_\Lambda(d\sigma_\Lambda|\xi)$ is quasi-local and continuous. We say that the specification $\gamma_\Lambda(\cdot|\cdot)$ is quasi-local.

We then call a measure Gibbs μ with potential U (and a priori measure λ) if $\gamma_\Lambda(d\sigma_\Lambda|\xi)$ is consistent with the finite-volume Gibbs measures, i.e., if for all $f : \Omega \to \mathbb{R}$ bounded and measurable, and $\Lambda \in \mathcal{S}$, we have

$$\mathbb{E}_\mu(f|\mathfrak{B}_{\Lambda^c})(\xi) = \int f(\sigma_\Lambda\xi_{\Lambda^c})\,\gamma_\Lambda(d\sigma_\Lambda|\xi)$$

for μ-almost every ξ.

We denote by $\mathcal{G}_\tau(U)$ the set of translation-invariant Gibbs measures associated with the potential U. These measures are called the "equilibrium states" associated with U.

The variational principle ([1] Chapter 16) implies that if $\mu, \nu \in \mathcal{G}_\tau(U)$, then $s_*(\mu|\nu) = s_*(\nu|\mu) = 0$, and conversely if $\mu \in \mathcal{G}_\tau(U)$ and $\nu \in \mathcal{P}_\tau$, is such that $s_*(\nu|\mu) = 0$, then $\nu \in \mathcal{G}_\tau(U)$. As a consequence of Theorem 1, we then obtain the following result:

Proposition 2. *Let $U \in \mathcal{U}$. If $\mu \in \mathcal{G}_\tau(U)$ satisfies $\mathrm{GCB}(C)$ for some $C > 0$, then $\mathcal{G}_\tau(U) = \{\mu\}$.*

This substantially extends the implication between $\mathrm{GCB}(C)$ and uniqueness of equilibrium states from [9], where we only considered finite single-spin spaces S.

Remark 5. *Because our techniques are based on relative entropy density, we cannot exclude the existence of non-translation-invariant Gibbs measures, even in the presence of a translation-invariant Gibbs measure satisfying $\mathrm{GCB}(C)$. In other words, even if there exists a unique equilibrium state, there might still be non-translation-invariant Gibbs measures. We believe, however, that the presence*

of a translation-invariant Gibbs measure satisfying GCB(C) *implies a stronger form of uniqueness, which excludes the presence of non-translation-invariant Gibbs measures.*

4.2. Sets of Zero-Information Distance

The example of the set of equilibrium states from the previous subsection leads naturally to the more general notion of "zero-information distance sets" defined below.

Definition 4. *We call a subset $\mathscr{K} \subset \mathcal{P}_\tau(\Omega)$ a zero-information distance set if for all $\mu, \mu' \in \mathscr{K}$, $s_*(\mu|\mu') = s_*(\mu'|\mu) = 0$.*

From Theorem 1, we then immediately obtain the following proposition.

Proposition 3. *Let $\mathscr{K} \subset \mathcal{P}_\tau(\Omega)$ be a zero-information distance set. If there exists $\mu \in \mathscr{K}$, which satisfies* GCB(C) *for some $C \in (0, \infty)$, then $\mathscr{K}$ is a singleton.*

We provide four further examples (beyond equilibrium states) of such zero-information distance sets, illustrating Proposition 3.

(a) **Asymptotically decoupled measures and Π^f-compatible measures.**
A first generalization of the Gibbsian context is provided in the realm of "asymptotically decoupled measures" via the notion of Π^f-compatible measures, see [16]. This setting goes beyond quasi-local specifications, and therefore includes many relevant examples of non-Gibbsian measures.
In this setting, the set of Π^f-compatible measures (associated with a local function f) is a zero-information set ((see [16] Theorem 4.1), and therefore, if this set contains an element μ satisfying GCB(C), then it coincides with the singleton $\{\mu\}$.

(b) **Renormalization group transformations of Gibbs measures.**
Another important class of examples is the following. We say that a transformation $T : \mathcal{P}_\tau(\Omega) \to \mathcal{P}_\tau(\Omega')$ preserves zero- information distance sets if a zero-information distance set is mapped by T onto a zero-information distance set. Important examples of such transformations T are local and translation-invariant renormalization group transformations studied in [11], Section 3.1, p 960, conditions T1-T2-T3. Examples of such transformations include block-spin averaging, decimation, and stochastic transformations such as the Kadanoff transformation. Because the transformations are "local and translation-invariant probability kernels", one immediately infers the property $s_*(\mu T|\nu T) \leq s_*(\mu|\nu)$.
In this setting, Proposition 3 implies that if $U \in \mathcal{U}$, $\mu \in \mathcal{G}_\tau(U)$ is an associated translation-invariant Gibbs measure, and μT satisfies GCB(C) for some $C \in (0, \infty)$, then $\nu = \mu T$ for all ν, such that $s_*(\nu|\mu T) = 0$. In particular, this implies that $\mu'T = \mu T$ for all $\mu' \in \mathcal{G}_\tau(U)$. Indeed, in that case $s_*(\mu'T|\mu T) \leq s_*(\mu'|\mu) = 0$.
Notice that μT can be non-Gibbs; therefore, the implication $\nu = \mu T$ for all ν such that $s_*(\nu|\mu T) = 0$ cannot be derived from the variational principle.

(c) **Projections of Gibbs measures.**
Let μ be a translation-invariant Gibbs measure on the state space $S^{\mathbb{Z}^d}$ (associated with a translation-invariant potential) which satisfies GCB(C) for some $C \in (0, \infty)$. Let for $d' < d$, $\mu_{d'}$ denote its restriction to the sublattice $L_{d'} := \{(x_1, \ldots, x_{d'}, 0, \ldots, 0) : x_1, \ldots, x_{d'} \in \mathbb{Z}\}$. It is clear that $\mu_{d'}$ satisfies GCB(C) with the same constant $C \in (0, \infty)$. Therefore, any translation-invariant measure on $S^{L_{d'}}$ that differs from $\mu_{d'}$ has strictly positive lower relative entropy density with regard to $\mu_{d'}$.
As a consequence, if $\mu_{d'}$ is a Gibbs measure for a translation-invariant potential $U'_{d'}$, then this potential $U'_{d'}$ has no other translation-invariant Gibbs measures. This gives uniqueness for a set of Gibbs measures where the potential is only implicitly defined, and can be complicated, i.e., uniqueness is not a consequence of a simple criterion.
Projections of Gibbs measures arise naturally in the context of probabilistic cellular automata, where the stationary measures are projections of the space–time Gibbs

measures [4]. In this setting, the result tells us that if the space–time measure satisfies GCB(C) for some $C > 0$, then the unique stationary measure, if Gibbs, has a potential with a unique equilibrium state. Projections of Gibbs measures can fail to be Gibbs, as is shown in [17] for projection of the low-temperature Ising model in $d = 2$ on the X-axis. It is an open and interesting problem to investigate whether this projected measure satisfies the Gaussian concentration bound.

(d) **Stationary measures for Ising spin Glauber dynamics.**
An additional example of a zero-information distance set is the set of stationary and translation-invariant measures for (Ising spin, i.e., S is finite) Glauber dynamics under the condition that this set contains at least one translation-invariant Gibbs measure as a stationary measure; see [18], Section 4. See also [19,20] for earlier results in the setting of reversible Glauber dynamics, and [21] for recent results in this spirit for more general local dynamics. As a consequence of Proposition 3, we then conclude that if there exists a translation-invariant Gibbs measure ν as stationary measure, and there exists a translation-invariant stationary measure μ satisfying GCB(C) for some $C > 0$, then $\mu = \nu$ coincide, and μ is the unique translation-invariant stationary measure. Moreover, if, in this setting, one can show that when starting the dynamics from a translation-invariant initial measure μ and denoting μ_t for the measure at time $t > 0$, we have $s_*(\mu_t | \nu) \to 0$ as $t \to \infty$, then, from Corollary 1, we obtain that $\mu_t \to \nu$ as $t \to \infty$ in the sense of the distance (7).

5. Generalization

In the setting of Section 2.1, without the additional assumption of finiteness of the diameter of S, the definition of the oscillation of f in (2) is no longer appropriate. Indeed, it becomes natural to include unbounded functions, which makes (2) infinite. Consider, e.g., $S = \mathbb{R}$, and $\Omega = S^{\mathbb{Z}^d}$ equipped with a product of Gaussian measures, then the function $f(\eta) = \eta_i$ should be a possible choice. We consider now a general standard Borel S, which is such that for the product space $\Omega = S^{\mathbb{Z}^d}$, quasi-local bounded functions are measure separating.

In order to proceed, we therefore associate with a function $f : \Omega \to \mathbb{R}$ an abstract sequence of oscillations $\delta f = (\delta_i f)_{i \in \mathbb{Z}^d}$ satisfying the following conditions.

Definition 5. *We say that a map $\delta : \mathscr{C}(\Omega) \to [0, \infty]^{\mathbb{Z}^d}$ is an allowed sequence of oscillations if the following four conditions are met.*

1. *Translation invariance: $(\delta(\tau_i f))_j = \delta_{i+j} f$, $i, j \in \mathbb{Z}^d$.*
2. *Non-degeneracy: $\delta_i f$ is zero for a function f if and only if f does not depend on the i-th coordinate, i.e., $\delta_i f = 0$ if and only if for all η, σ such that $\eta_j = \sigma_j$ for all $j \neq i$, $f(\eta) = f(\sigma)$.*
3. *Monotonicity: for $\xi \in \Omega$ and f a bounded quasi-local function, we consider the local approximation of f given by*
$$f_{\Lambda, \xi}(\eta) = f(\eta_\Lambda \xi_{\Lambda^c}).$$
   Then, we require that for all ξ, for all Λ and for all $i \in \mathbb{Z}^d$, $\delta_i f_{\Lambda, \xi} \leq \delta_i f$.
4. *Degree one homogeneity: $\delta_i(\beta f) = |\beta| \delta_i f$ for all $\beta \in \mathbb{R}$ and for all $i \in \mathbb{Z}^d$.*

Notice that Condition 3 implies that for given f, ξ, $\Lambda \subset \Lambda'$, and $i \in \mathbb{Z}^d$, we have $\delta_i f_{\Lambda, \xi} \leq \delta_i f_{\Lambda', \xi}$. Indeed, notice that $(f_{\Lambda', \xi})_{\Lambda, \xi} = f_{\Lambda, \xi}$ for $\Lambda \subset \Lambda'$.

The most natural example different from (2) is

$$\delta_i f = \sup_{\substack{\sigma, \eta \in \Omega: \\ \sigma_j = \eta_j, \forall j \neq i}} \frac{f(\sigma) - f(\eta)}{d_S(\sigma_i, \eta_i)}.$$

More generally, one can define

$$\delta_i f = \sup_{\substack{\sigma, \eta \in \Omega: \\ \sigma_j = \eta_j, \forall j \neq i}} \frac{f(\sigma) - f(\eta)}{\psi(\sigma_i, \eta_i)}$$

where $\psi : S \times S \to [0, \infty)$ satisfies $\psi(s, s') = 0$ iff $s = s'$.

For a given sequence of oscillations δ, we call a function δ-Lipschitz if $\sup_{i \in \mathbb{Z}^d} \delta_i f < \infty$. We then introduce

$$\Delta^p(\Omega) = \{ f \in \mathscr{C}(\Omega) : \|\delta f\|_p < \infty \}.$$

Definition 6. *Let $\Omega = S^{\mathbb{Z}^d}$, where S is a standard Borel space. Assume an allowed sequence of oscillations δ is given. Let μ be a probability measure on Ω. We say that μ satisfies the Gaussian concentration bound with regard to δ with constant $C > 0$ (still abbreviated $\mathrm{GCB}(C)$), if for all bounded local functions we have*

$$\int \mathrm{e}^{f - \int f \, \mathrm{d}\mu} \, \mathrm{d}\mu \leq \mathrm{e}^{\frac{C}{2} \|\delta f\|_2^2}.$$

We then have the following analogue of Theorem 1. Because the proof follows exactly the same steps as the proof of Theorem 1, we leave it to the reader.

Theorem 2. *Assume δ is an allowed vector of oscillations. Assume that the set of bounded local δ-Lipschitz functions is measure separating. If μ is translation invariant and satisfies $\mathrm{GCB}(C)$, then for all μ' translation invariant, and $\mu' \neq \mu$ we have*

$$s_*(\mu'|\mu) > 0.$$

As a final comment, we remark that the fact that we have chosen the group $\mathbb{Z}^d$ is for the sake of simplicity. We can work with more general amenable groups, as in [22].

References

1. Georgii, H.-O. *Gibbs Measures and Phase Transitions*, 2nd ed.; De Gruyter: Berlin, Germany, 2011.
2. Stroock, D.W.; Zegarlinski, B. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Commun. Math. Phys.* **1992**, *149*, 175–193.
3. Liggett, T.M. *Interacting Particle Systems*, 2nd ed.; Springer: Berlin, Germany, 2005.
4. Lebowitz, J.L.; Maes, C.; Speer, E.R. Statistical mechanics of probabilistic cellular automata. *J. Stat. Phys.* **1990**, *59*, 117–170.
5. Maes, C.; Shlosman, S.B. Ergodicity of probabilistic cellular automata: A constructive criterion. *Comm. Math. Phys.* **1991**, *135*, 233–251.
6. Külske, C. Concentration inequalities for functions of Gibbs fields with application to diffraction and random Gibbs measures. *Comm. Math. Phys.* **2003**, *239*, 29–51.
7. Chazottes, J.-R.; Collet, P.; Redig, F. On concentration inequalities and their applications for Gibbs measures in lattice systems. *J. Stat. Phys.* **2017**, *169*, 504–546.
8. Chazottes, J.-R.; Gallo, S.; Takahashi, D. Gaussian concentration bounds for stochastic chains of unbounded memory. *Ann. Appl. Probab.* **2022**, in press.
9. Chazottes, J.-R.; Moles, J.; Redig, F.; Ugalde, E. Gaussian concentration and uniqueness of equilibrium states in lattice systems. *J. Stat. Phys.* **2020**, *181*, 2131–2149.
10. Bobkov, S.G.; Götze, F. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **1999**, *163*, 1–28.

11. van Enter, A.C.D.; Fernández, R.; Sokal, A.D. Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory. *J. Statist. Phys.* **1993**, *72*, 879–1167.
12. Gross, L. Absence of second-order phase transitions in the Dobrushin uniqueness region. *J. Stat. Phys.* **1981**, *25*, 57–72.
13. Boucheron, S.; Lugosi, G.; Massart, P. *Concentration Inequalities: A Nonasymptotic Theory of Independence*; Oxford University Press: Oxford, UK, 2013.
14. Bertini, L.; Cirillo, E.N.M.; Olivieri, E. Renormalization-group transformations under strong mixing conditions: Gibbsianness and convergence of renormalized interactions. *J. Stat. Phys.* **1999**, *97*, 831–915.
15. Armstrong-Goodall, J.; MacKay, R.S. Dobrushin and Steif metrics are equal. *arxiv preprint* **2021**, *arXiv:2104.08365.*
16. Pfister, C.-E. Thermodynamical aspects of classical lattice systems. In *In and Out of Equilibrium*; Progress in Probability; Birkhäuser: Boston, MA, USA, 2002; Volume 51, pp. 393–472.
17. Schonmann, R.H. Projections of Gibbs measures may be non-Gibbsian. *Comm. Math. Phys.* **1989**, *124*, 1–7.
18. H. Künsch. Nonreversible stationary measures for infinite interacting particle systems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **1984**, *66*, 407–424.
19. Holley, R.A.; Stroock, D.W. In one and two dimensions, every stationary measure for a stochastic Ising model is a Gibbs state. *Comm. Math. Phys.* **1977**, *55*, 37–45.
20. Higuchi, Y.; Shiga, T. Some results on Markov processes of infinite lattice spin systems. *J. Math. Kyoto Univ.* **1975**, *15*, 211–229.
21. Jahnel, B.; Külske, C. Attractor properties for reversible and irreversible particle systems. *Comm. Math. Phys.* **2019**, *366*, 139–172.
22. Tempelman, A. *Ergodic Theorems for Group Actions. Informational and Thermodynamical Aspects. Mathematics and Its Applications*; Springer: Berlin, Germany, 1992; Volume 78.