CONCENTRATION INEQUALITIES FOR FUNCTIONS OF GIBBS FIELDS WITH APPLICATION TO DIFFRACTION AND RANDOM GIBBS MEASURES

Christof Külske*

Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39, D-10117 Berlin, Germany

Abstract. We derive useful general concentration inequalities for functions of Gibbs fields in the uniqueness regime. We also consider expectations of random Gibbs measures that depend on an additional disorder field, and prove concentration w.r.t the disorder field. Both fields are assumed to be in the uniqueness regime, allowing in particular for non-independent disorder field. The modification of the bounds compared to the case of an independent field can be expressed in terms of constants that resemble the Dobrushin contraction coefficient, and are explicitly computable.

On the basis of these inequalities, we obtain bounds on the deviation of a diffraction pattern created by random scatterers located on a general discrete point set in the Euclidean space, restricted to a finite volume. Here we also allow for thermal dislocations of the scatterers around their equilibrium positions. Extending recent results for independent scatterers, we give a universal upper bound on the probability of a deviation of the random scattering measures applied to an observable from its mean. The bound is exponential in the number of scatterers with an upper bound rate that involves only the minimal distance between points in the point set.

2000 Mathematics Subject Classification. 82B44, 78A45, 60F10, 82B20
1999 Physics and Astronomy Classification Scheme (PACS) 61.10.Dp, 05.50.+q, 02.50.Cw

Key words and phrases. Gibbs measures, disordered systems, diffraction theory, random scatterers, random point sets, quasicrystals, large deviations.

Preprint No. 742 ISSN 0946 - 8633

---

* Work supported by the DFG

Contact: kuelske@wias-berlin.de, http://www.wias-berlin.de/private/kuelske

1. INTRODUCTION

Concentration inequalities for functions of random fields play an important role in various areas of probability theory, with numerous applications ranging from the more abstract to the explicit analysis of given models ([LT91],[Ta96],[Le01]). The aim of our present paper is twofold. First of all, motivated by the study of disordered systems, we derive general concentration inequalities for functions of Gibbs fields in the Dobrushin uniqueness regime that have not appeared before in this simple and useful form. The focus in our approach is on applicability of the estimates and not just existence. In particular we are interested not just in the mere finiteness of the constants appearing in the estimates but in explicit expressions that can be readily evaluated (or estimated) in given models.

Secondly, in parallel to the general treatment, we show in this paper how these estimates can be applied to the analysis of the self-averaging properties of random diffraction measures of general point set Γ in the Euclidean space ([BaaHoe00],[Hof95a],[Hof95b]). Randomness appears here naturally as a probability distribution governing the thermal dislocations of the scatterers around their equilibrium positions. Additionally, we also consider a random distribution for the scattering amplitudes appearing. We stress that these random scattering measures are beautiful objects themselves that are of considerable interest. Extending previous results for independent scatterers of [K01b], we give a universal upper bound on the probability of a deviation of the random scattering measures applied to an observable from its mean that depends on the point set Γ only through the minimal distance between its points (Theorem 4,5).

Being motivated by the study of general disordered systems, the first and basic question is for a useful concentration estimate of a function of a Gibbs field in the uniqueness regime, where no assumptions are made about translational invariance (Theorem 1). In the next more interesting step we will be interested also in expectations of functions w.r.t. to Gibbs measures, when the latter are themselves functions of another random field modelling the disorder (Theorems 2,3). While it is very often assumed in the classical models of disordered systems (like the random field Ising model or the Edwards Anderson spin glass) that this random field is independent, there are interesting physical situations where this is not the case. It is our emphasis that we are able to treat also this dependent situation, again assuming no symmetries at all. We believe that these inequalities can be useful tools in a variety of circumstances to extend results for disordered systems from independent disorder to nonindependent disorder.

The assumption we chose to impose on the random distributions is essentially the Dobrushin uniqueness condition. (See [Geo88] Chapter 8. More precisely we assume even a slightly stronger form of it, but the difference is minor from the point of view of applications.) It turns out that the constants that appear in our estimates can in all cases be expressed by the original Dobrushin contraction coefficient, and constants measuring the dependence of one random field from the other one that are defined in the same spirit. We stress that all of these quantities can be estimated in terms of bounds of the potential of a Gibbsian specification in a very simple way.

Coming back to our main example of diffraction measures we will need to estimate the concentration properties of a function that is not convex. Unfortunately non-convex functions are appearing in a lot of applications, and so very often all elegant methods based on convexity are simply not applicable. Let us mention in this context also the very beautiful result of [SZ92] who proved that the Dobrushin-Shlosman Mixing Condition [DS84] implies a Logarithmic Sobolev inequality and vice versa, at least for certain state spaces. (The Dobrushin-Shlosman condition is less restrictive than the Dobrushin condition we are working with. A simple new proof of the first implication was recently given in [Ce01]). In principle one can obtain exponential concentration as a corollary to a log-Sobolev inequality (see [Le01] Theorem 5.3). Here the problem would be that there are no handy formulas for the constant appearing in the log-Sobolev inequality so that also the resulting concentration estimates would not be explicit. Also, for the purpose of the concentration results we are interested in, log-Sobolev inequalities are a detour, assuming an additional structure (gradient) that is not needed for the present problem.

We conclude this introduction with an outline of the rest of the paper. Chapter 2 is an extended introduction containing an overview of the main results, including the general concentration theorems and a first application to random scattering measures. In Chapter 3 we give more results for random scattering measures along with their proofs. They follow in an elementary but slightly tricky way from the general concentration estimates. In Chapter 4 we describe applications to disordered spin systems and provide details about the estimation of constants. Chapter 5 contains a simple proof of the basic concentration estimate of Theorem 1, where in particular the form of the constants appearing becomes clear. It follows from consequent use of estimates in the Dobrushin uniqueness region on the basis of the classical martingale method. Chapter 6 contains a proof of the concentration estimates for expectations w.r.t random Gibbs measures of Theorems 2 and 3. They use the explicit knowledge of the variation of the Gibbs measure in the Dobrushin uniqueness regime when the local specification is perturbed, in combination with a chain rule argument for variations.

2. MAIN RESULTS

2.1 Basic concentration estimate in the Dobrushin uniqueness regime

Suppose that Γ is a countably infinite or finite set and E is a standard Borel space. In our applications below E will be a finite set or a ball in the finite dimensional Euclidean space.

Suppose we are given a *random field* $X = (X_x)_{x\in\Gamma}$ taking values on E^Γ, with distribution μ. Following standard notation, we denote by

$$C = (C_{x,y})_{x,y\in\Gamma} \quad \text{with} \quad C_{x,y} := \sup_{\substack{\xi,\xi'\in E^\Gamma \\ \xi_{y^c}=\xi'_{y^c}}} \left\| \mu(\,\cdot\,|\xi_{x^c}) - \mu(\,\cdot\,|\xi'_{x^c}) \right\|_x \tag{2.1}$$

the *Dobrushin interdependence matrix.* Here the l.h.s. of (2.1) denotes the *variational distance at the site x.* Given two measures ρ and ρ' on E^Γ it is defined by $\left\|\rho(\,\cdot\,)-\rho'(\,\cdot\,)\right\|_x = \max_f\left|\int \rho(d\xi_x)f(\xi_x)-\int \rho'(d\xi_x)f(\xi_x)\right|/\delta(f)$. The maximum is over non-constant functions f on E. Here and throughout the paper $\delta(f) := \sup_{u,u'}|f(u)-f(u')|$ denotes the *total variation* of a function f where u, u' are taken over the range of definition of this function. If f is vector valued, we the Euclidean norm on the r.h.s. We write $y^c \equiv \Gamma\backslash y$ for the complement of the site y.

One says that the random field X (respectively its distribution μ) satisfies the *Dobrushin uniqueness condition* iff

$$c^X := \sup_{x\in\Gamma}\sum_{y\in\Gamma} C_{x,y} < 1 \tag{2.2}$$

We need to introduce a new notion. Let us say that the random field X (resp. μ) satisfies the *transposed Dobrushin uniqueness condition* iff

$$c_t^X := \sup_{y\in\Gamma}\sum_{x\in\Gamma} C_{x,y} < 1 \tag{2.3}$$

Obviously c^X and c_s^X vanish if the X_x's are independent. Then we have the following general concentration estimate.

Theorem 1. *Suppose the random field $X = (X_x)_{x\in\Gamma}$ taking values in E^Γ is distributed according to a Gibbs measure μ that obeys the Dobrushin uniqueness condition with Dobrushin constant c^X, and also the transposed Dobrushin uniqueness condition with constant c_t^X.*

Suppose that F is a real function on E^Γ with $\mu\big(\exp(tF(X))\big) < \infty$ for all real t.

Then we have the Gaussian concentration estimate

$$\mu\Big(F(X)-\mu\big(F(X)\big)\geq r\Big) \leq \exp\left(-\frac{r^2}{2}\,\frac{(1-c^X)(1-c_t^X)}{\left\|\underline{\delta}(F)\right\|_{l^2}^2}\right) \quad \forall r\geq 0 \tag{2.4}$$

Here $\underline{\delta}(F) \equiv \big(\delta_x(F)\big)_{x\in\Gamma}$ is the (infinite) *variation vector* of F, where $\delta_x(F) = \sup_{\xi,\xi';\xi_x=\xi'_x}|F(\xi)-F(\xi')|$ denotes *the variation of F at the site x*. Its l^2-norm is denoted by $\left\|\underline{\delta}(F)\right\|_{l^2}^2 \equiv \sum_{x\in\Gamma}(\delta_x(F))^2$. If this norm is infinite, the statement is empty (and thus correct).

Remark. Ready-to-use upper bounds on the Dobrushin constant c are known when the conditional expectations are given in terms of a Gibbsian specification with a defining interaction potential Φ (see Georgii Chapter 8.1)[1]. Let us mention the following general

---

[1] Prescribing a consistent set of conditional probabilities in terms of an interaction potential Φ is of course the standard way of producing a Gibbs measure. Recall the following well-known facts about Dobrushin uniqueness. If μ is an infinite volume measure for which the Dobrushin uniqueness condition (2.2) holds, it is necessarily the unique Gibbs measure for the local specification defined by the system of its conditional expectations. This can be proved by a contraction method where the Dobrushin constant c appears as contraction coefficient (See e.g. Theorem 8.7 of [Georgii]). Existence must be proved seperately but is of course guaranteed e.g. by a compact state space E.

classic bound on c^X that takes care of all high-temperature situations. We point out here that it gives *the same estimate* we would have on c^X also on the constant c_t^X. So, suppose that $\mu(d\xi_x|X_{x^c} = \xi_{x^c}) = \exp\bigl(-\sum_{A\ni x}\Phi_A(\xi_x\xi_{\Gamma\setminus x})\bigr)\lambda(d\xi_x)/Z_x(\xi_{\Gamma\setminus x})$ for a Gibbsian potential $\Phi = (\Phi_A)_{A\subset\Gamma}$ (meaning that Φ_A is a function on E^Γ that depends only on E^A). Here λ is a σ-finite measure on E, which must be the same for all sites $x \in \Gamma$ and $Z_x(\xi_{\Gamma\setminus x})$ is the usual normalization factor. Then we have that

$$c^X, c_t^X \le \frac{1}{2}\sup_{x\in\Gamma}\sum_{A\ni x}(|A|-1)\delta(\Phi_A) \tag{2.5}$$

independently of the single site part $\delta(\Phi_x)$. This is stated as Proposition 8.8 in [Geo] as a bound for c^X, for a brief explanation why it implies the bound for c_t^X too, see Chapter 4. Be aware however that interdependence constants C_{xy} and C_{yx} whose actual values differ significantly could occur for models with very different Φ_x for different sites $x \in \Gamma$.

Remark. Often the theorem will be used in the following situation. Suppose that $F = F(X_\Lambda)$ is a function that depends only on variables in a finite set $\Lambda\subset\Gamma$. Then $\|\delta(F)\|_{l^2}^2 \le |\Lambda|\,\|\delta(F)\|_{l^\infty}^2$. The reader who likes to see an interesting application of this is advised to go directly to 2.3 "First application to random diffraction measures".

2.2 Chain rule concentration estimates for disordered systems with dependent disorder

The concentration inequalities we are going to present now apply to situations where a random fields Y is given whose distribution depends on the realizations of another 'external' random field X. This is precisely the case in the study of *disordered systems* where X models the external randomness and one is given the Gibbs distribution of Y for any fixed configuration of X. We assume here that both fields are in the Dobrushin uniqueness regime in a natural sense, and that the dependence of Y on X is not completely unreasonable. To control these properties quantitatively we will have to introduce constants (in the spirit of the Dobrushin constant) governing the deviation of the fields X (respectively Y) from the case of product distributions, and constants governing the degree of influence from X to Y. Very often in disordered systems the distribution of the external random field X will even be assumed to be a product distribution, but we don't need this for our estimates. Indeed, it is our emphasis to be able to treat the more general case of Dobrushin uniqueness for X. The resulting concentration estimates will depend only on these constants, and thus contain only minimal information about the distribution of (X, Y). We stress that while the definition of the constants might look at little frightening at first sight, they are very easy to control, so the estimates are very explicit. (This is done e.g. by (2.5) and an analogous consideration given below in Chapter 4.)

We call them 'chain rule estimates' because the distribution of the field Y is a (possibly very complicated) function of the field X, so that in order to control expectations of functions of both fields some 'chain rule for variations' will be needed. Let us now formulate our results in a precise manner.

Suppose that Γ_X and Γ_Y are countable (finite or infinite) sets, and E_X and E_Y are standard Borel spaces. Suppose that we are given two random fields $X = (X_x)_{x\in\Gamma_X}$ taking values in $E_X^{\Gamma_X}$ and $Y = (Y_x)_{x\in\Gamma^Y}$ taking values in $E_Y^{\Gamma^Y}$. Suppose that their joint distribution μ that is as follows.

(i) The marginal of μ on the variable X, denoted by μ^X, is a Gibbs measure that obeys the Dobrushin uniqueness condition (2.2) and the transposed condition (2.3). We denote the corresponding *'marginal Dobrushin constant' by* c^X and its transposed version by c_t^X.

(ii) For any realization η of X the *conditional distribution* of Y *given* X, denoted by $\mu(\,\cdot\,|X=\eta)$, is a Gibbs measure that obeys Dobrushin uniqueness and its transposed version. Moreover we demand uniformity in η in the sense that the following *uniform Dobrushin constant* $c^{Y,\infty}$ and its transposed version $c_t^{Y,\infty}$ obey

$$c^{Y,\infty} := \sup_{x\in\Gamma_Y}\sum_{y\in\Gamma_Y}\sup_{\eta} C^Y_{x,y}(\eta) < 1, \qquad c_t^{Y,\infty} := \sup_{y\in\Gamma_Y}\sum_{x\in\Gamma_Y}\sup_{\eta} C^Y_{x,y}(\eta) < 1 \tag{2.6}$$

Here $C^Y_{x,y}(\eta)$ denotes the Dobrushin matrix for the fixed configuration η.

(iii) To control the dependence of the field Y from the field X let us introduce their *dependence matrix* in the following way

$$C^{Y\leftarrow X}_{z,u} := \sup_{\substack{\eta,\eta';\eta_{u^c}=\eta'_{u^c}\\ \omega_{z^c}}} \|\mu(\,\cdot\,|X=\eta, Y_{z^c}=\omega_{z^c}) - \mu(\,\cdot\,|X=\eta', Y_{z^c}=\omega_{z^c})\|_z \tag{2.7}$$

It describes the possible change of the fixed Y-single site conditional distribution at z w.r.t. variation of the X-variables in u. The supremum is taken over the respective spaces, i.e. $\eta,\eta'\in E_X^{\Gamma_X}$ and $\omega\in E_Y^{\Gamma^Y}$. We demand that the following *dependence constant* and its transposed version obeys

$$c^{Y\leftarrow X} := \sup_{z\in\Gamma_Y}\sum_{u\in\Gamma_X} C^{Y\leftarrow X}_{z,u} < \infty, \qquad c_t^{Y\leftarrow X} := \sup_{u\in\Gamma_X}\sum_{z\in\Gamma_Y} C^{Y\leftarrow X}_{z,u} < \infty \tag{2.8}$$

For independent X and Y these constants vanish, obviously.

We need a little more notation. Let us write $\delta_x^X(G) := \sup_{\eta,\eta';\eta_{x^c}=\eta'_{x^c},\omega} |G(\eta,\omega) - G(\eta',\omega)|$ for the *X-variation at the site* $x\in\Gamma^X$ for a function G on the product space. The notation for $\delta_x^Y(G)$ is analogous. Note that the corresponding *partial infinite variation vectors* $\underline{\delta}^X(G) \equiv \big(\delta_x(G)\big)_{x\in\Gamma_X}$ and $\underline{\delta}^Y(G)$ are not in the same space anymore, in general because the index sets Γ_X and Γ_Y are different.

Then the first result concerns the concentration properties of Y-averages w.r.t the field X.

Theorem 2. *Suppose that X and Y are random fields with joint distribution μ satisfying (i), (ii), (iii). Suppose that G is a real function on $E^{\Gamma_X}\times E^{\Gamma_Y}$ with $\mu\big(\exp(tG(X,Y))\big)<\infty$ for all real t. Then we have the Gaussian concentration estimate*

$$\begin{aligned}&\mu^X\Big(\mu\big(G(X,Y)\,|X\big)-\mu\big(G(X,Y)\big)\ge r\Big)\\&\le\exp\left(-\frac{r^2}{2}\frac{(1-c^X)(1-c_t^X)}{\Big(\big\|\underline{\delta}^X(G)\big\|_{l^2}+c^{Y,\mathit{eff}}\big\|\underline{\delta}^Y(G)\big\|_{l^2}\Big)^2}\right)\quad\forall r\ge 0\end{aligned}\tag{2.9}$$

with the 'effective constant' $c^{Y,\mathit{eff}}=\left(\frac{c^{Y\leftarrow X}c_t^{Y\leftarrow X}}{(1-c^{Y,\infty})(1-c_t^{Y,\infty})}\right)^{\frac{1}{2}}<\infty$.

Remark. We can view $c^{Y,\mathrm{eff}}$ as the 'effective strength' of the influence the random field X has on the field Y. The form of the constants will become clear in the proof that combines an application of Theorem 1 for the X-marginal with a chain rules for variations.

Remark. The reader should realize that the dependence constants (and thus $c^{Y,\mathrm{eff}}$) are as easily estimated as the Dobrushin constants if the single site conditional distribution of Y_x is given in a Gibbsian form with a random energy function. This is analogous to the estimate for the Dobrushin constants in (2.5) and is explained in more detail in Proposition 1 of Chapter 4.

Almost automatically we then also have the following 'total concentration result'.

Theorem 3. *Under the hypothesis of Theorem 2 we have the 'total' concentration estimate*

$$\begin{aligned}&\mu\Big(G(X,Y)-\mu\big(G(X,Y)\big)\ge r\Big)\\&\le\exp\left(-\frac{r^2}{2}\left[\left[\frac{(1-c^X)(1-c_t^X)}{\Big(\big\|\underline{\delta}^X(G)\big\|_{l^2}+c^{Y,\mathit{eff}}\big\|\underline{\delta}^Y(G)\big\|_{l^2}\Big)^2}\right]^{-1}+\left[\frac{(1-c^{Y,\infty})(1-c_t^{Y,\infty})}{\big\|\underline{\delta}^Y(G)\big\|_{l^2}^2}\right]^{-1}\right]^{-1}\right)\end{aligned}\tag{2.10}$$

Remark. The form is easy to understand. The term within the inverse of square brackets has the character of a squared variance. It is the sum of the term for the Y-average from Theorem 2 and a uniform version of the term for the conditional Y-distribution from Theorem 1.

2.3 First application to Random diffraction measures

It is our aim now to look at the self-averaging properties of the diffraction pattern created by random scatterers ('atoms') located on a general discrete point set Γ which is a subset of the Euclidean space. The function F whose concentration properties we will be interested in describes the result of a measurement at the random diffraction pattern. We stress that this function is *not* a convex function, so all methods based on convexity simply cannot be applied. To appreciate the charm of this topic the interested reader may take a look at some of the beautiful experimental diffraction patterns of quasicryals (This is how quasicrystals were discovered in 1982). Here is the problem. Let us very briefly describe at first how this function turns up.[2] Consider the scattering image of the complex random measure (*'random Dirac comb'*) given by

$$\rho_\Gamma(\eta,\omega)=\sum_{x\in\Gamma}\eta_x\delta_{x+\omega_x} \tag{2.11}$$

where δ_x denotes the Dirac-measure at the site x. The point set $\Gamma\subset\mathbb{R}^\nu$ is assumed to be countable. The η_x's are complex numbers modelling *scattering amplitudes.* The ω_x's *'dislocations'* are vectors in the underlying Euclidean space $\mathbb{R}^\nu$. Below they will be made random according a random field $X=(X_x)_{x\in\Gamma}$ taking values $\eta=(\eta_x)_{x\in\Gamma}$ and a random field $Y=(Y_x)_{x\in\Gamma}$ taking values ω. So, the point set Γ modelling the locations of the scatterers in Euclidean space has a geometric meaning here, but it also serves just as an index set for the random fields. The classes of distributions we allow for them will be described later.

Fix any finite volume $\Lambda\subset\Gamma$. Then, the object that contains all information about the scattering image of the points in Λ is the *finite volume scattering measure* which by definition is the Fourier-transform of the corresponding *finite volume autocorrelation measure.* The latter is defined as follows

$$\gamma_\Lambda^{\eta,\omega}:=\frac{1}{|\Lambda|}\sum_{x,x'\in\Lambda}\eta_x\eta^*_{x'}\delta_{x-x'+\omega_x-\omega_{x'}} \tag{2.12}$$

Here the star denotes complex conjugate. Since we allow Λ to be *any* finite set, we have chosen the natural normalization by the number of points, as in [K01b]. A measurement on the scattered intensity is described by an *observable* $k\mapsto\varphi(k)$ *in Fourier-space,* modelling the measurement device, which is usually taken as a Schwartz test-function.

The corresponding result of the measurement is then given by $\hat\gamma_\Lambda^{\eta,\omega}(\varphi)\equiv\int\hat\gamma_\Lambda^{\eta,\omega}(k)\varphi(k)dk$. Here the Fourier-transform of a tempered distribution γ is defined by duality, $\hat\gamma(\varphi)=$

---

[2] For a summary of the basic notions of mathematical scattering theory for point scatterers, see e.g. Chapter II of [BaaHoe00] and Appendix A of [K01b]. The reason for the definitions of the diffraction measures can be understood in an elementary way by superposition of the reflections of an incoming beam at the individual scatterers. The results are physically meaningful when one takes measurements at distances far away from the scatterers and there is only single-scattering.

$\gamma(\hat{\varphi})$, where $\hat{\varphi}$ denotes the Fourier-integral of the Schwartz-function φ over $\mathbb{R}^\nu$. So, the function we are interested in is given by

$$(\eta,\omega) \mapsto \hat{\gamma}_\Lambda^{\eta,\omega}(\varphi) = \frac{1}{|\Lambda|} \sum_{x,x'\in\Lambda} \eta_x \eta_{x'}^* \hat{\varphi}(x - x' + \omega_x - \omega_{x'}) \tag{2.13}$$

We assume that the function $\varphi(k)$ is real and view it as a fixed parameter, so that (2.13) is a real function[3] on the random fields modelling the dislocations and random amplitudes.

We can now take averages of this function describing the random scattering image for instance w.r.t the distribution of the dislocations ω to obtain *an ω-averaged scattering image.* This can of course also be done w.r.t the scattering amplitudes η, or w.r.t. to both random fields η and ω. The study of the *large Λ-behavior of the average* is then one part of the story that is essentially reduced to understanding the diffraction pattern of Γ without disorder. The other part of the story which we are going to discuss now is the *control of the self-averaging properties of the diffraction image.* Concentration estimates were looked at for the first time in [K01b], for the cases independent ω_x's and fixed η_x's, and vice versa. Before that there were only few partial results of the SLLN type can be found in the quasicrystal literature for special sets Γ, see however [Hof95a]. (This is because of the different inclinations of probabilistic, statistical mechanics and diffraction communities which we are hoping to bring together at this point.) The emphasis in this study is to understand the influence of the point set Γ and the function $\hat{\varphi}$ for the quality of the concentration estimate. Since scattering experiments are a tool to guess the structure of Γ one is interested in estimates that depend on very little a priori information about Γ. It turned out in [K01b] that for the independent case we could obtain large deviation upper bounds that involve only the minimal distance between points in Γ and hence do not depend on the structure of the set Γ at all. This means in particular that the quality of the large deviation estimate is independent of the nature of the limiting diffraction image when Λ tends to infinity, be it pure point or diffuse. The dependence on the observable φ is expressed then in terms of a suitable Sobolev-norm.

The proof given in [K01b] for the independent case used a cluster expansion for the logarithmic moment generating function of (2.13). At the price of some technical work, it has the advantage to provide also a central limit theorem (for 'non-pathological' Γ, in particular lattices) and shows that the bounds appearing are essentially optimal. On the basis of the general results in Theorems 1,2,3 we can now extend the concentration result in a rather easy and elegant way to the case of dependent fields that obey Dobrushin uniqueness. Let us give here only the result that corresponds to Theorem 1, and provide more discussion later.

---

[3] Write $\hat{\gamma}_\Lambda^{\eta,\omega}(k) = |\sum_{x\in\Lambda} \eta_x e^{ik\cdot(x+\omega_x)}|^2$ for the Lebesgue density of the finite volume scattering measure. So, for real test functions $\varphi(k)$ the function (2.13) is always real, and it is nonnegative if $\varphi \geq 0$. Of course it is not a convex function in ω but of oscillatory nature! It is convex as a function of η though.

Theorem 4. *Assume that* $X = (X_x)_{x\in\Gamma}$ *is a field of complex random variables ('scatterers') indexed by the point set* $\Gamma\subset\mathbb{R}^\nu$*, and that* $Y = (Y_x)_{x\in\Gamma}$ *is a random field of* $\mathbb{R}^\nu$*-valued random variables ('thermal dislocations'). Assume that the field of the joint variables* $Z = (X, Y) = (X_x Y_x)_{x\in\Gamma}$ *is distributed according to a Gibbs measure* μ *that obeys the Dobrushin uniqueness condition (2.2) with a Dobrushin constant c. Assume also the transposed Dobrushin uniqueness condition (2.3) with constant* c_t*.*

Let $\Lambda\subset\Gamma$ *be any finite set. Assume that the random point set* $\{x+\omega_x, x\in\Lambda\}$ *has minimal distance* $b > 0$*, for* μ*-a.e. realization of* ω *of the dislocations. Moreover we assume the following* μ*-a.s. uniform bounds on the single site distributions*

$$|X_x| \le 1, \qquad \delta(X_x) \le \varepsilon_{sc}, \qquad \delta(Y_x) \le \varepsilon_{dl} \tag{2.14}$$

for all $x \in \Lambda$[4]*. Then the corresponding random scattering image* $\hat\gamma_\Lambda^{X,Y}(\varphi)$ *in the finite volume* Λ *obeys the universal large deviation estimate*

$$\begin{aligned}&\mu\left(\left|\hat\gamma_\Lambda^{X,Y}(\varphi) - \mu\big(\hat\gamma_\Lambda^{X,Y}(\varphi)\big)\right| \ge r\right)\\ &\le 2\exp\left(-\frac{|\Lambda|\, r^2}{8}\frac{(1-c)(1-c_t)}{\big(\varepsilon_{sc}\|\hat\varphi\|_{\nu,b} + \varepsilon_{dl}\|d\hat\varphi\|_{\nu,b}\big)^2}\right) \qquad \forall r > 0\end{aligned} \tag{2.15}$$

Here we have introduced the Sobolev-norm involving integrals of derivatives up to the order of the dimension ν where we make explicit also a scaling factor $b/2$. For a function $g : \mathbb{R}^\nu \to \mathbb{C}$ the norm is given by

$$\|g\|_{\nu,b} := \frac{1}{|B_1|}\sum_{k=0}^{\nu}\frac{1}{k!}\frac{1}{(b/2)^{\nu-k}}\int_{\mathbb{R}^\nu}\|d^k g(y)\|dy \tag{2.16}$$

The constant $b/2$ plays the role of fixing a length scale and here it is the *'uniform packing radius'* as defined above. The constant $|B_1|$ denotes the volume of the ν-dimensional unit ball.[5]

Remark. Theorem 4 shows self-averaging of the diffraction measures applied to an observable that is exponential in the number of scatterers $|\Lambda|$, with an explicit estimate on the rate. We regard this estimate as very satisfactory. Indeed, the l.h.s. of (2.15) depends in a complicated way on three complicated objects, the geometry of the point set $\Lambda\subset\Gamma$, the test function φ, and the distribution μ of the random field (ω, η). The upper bound

---

[4] So ε_{dl} bounds the diameter of the supports of the distribution of the dislocation variables Y_x taken in the Euclidean norm for all sites x.

[5] Of course, $d^k g(y) : (\mathbb{R}^\nu)^k \to \mathbb{R}^\nu$ denotes the k-th differential of g at the point y and $\|d^k g(y)\| = \sup_{|v_1|=\dots|v_k|=1}|d^k g(y)[v_1,\dots,v_k]|$ is the usual norm of a k-multilinear mapping, at any fixed point y, where $|v|$ denotes the Euclidean norm. Similarly $\|dg\|_{\nu,b} = \frac{1}{|B_1|}\sum_{k=0}^{\nu}\frac{1}{k!}\frac{1}{(b/2)^{\nu-k}}\int_{\mathbb{R}^\nu}\|d^{k+1}g(y)\|dy$. The advantage of including the factor $b > 0$ inside the definition of the norm is the scale invariance: Rescaling of the measurement function $\varphi_\sigma(k) = \sigma^{-\nu}\varphi_1(k/\sigma)$ where φ_1 is a probability density w.r.t the ν-dimensional Lebesgue measure, leads to $\|\hat\varphi_\sigma\|_{\nu,b} = \|\hat\varphi_1\|_{\nu,b\sigma}$. Similarly $\varepsilon\ \|d\hat\varphi_\sigma\|_{\nu,b} = \varepsilon\sigma\ \|d\hat\varphi_1\|_{\nu,b\sigma}$.

on the r.h.s. of (2.15) is in comparison very simple. The influence of the dependence structure of the random field is entirely factorized into the constant $(1-c)(1-c_t)$, a structure that is inherited from Theorem 1. The dependence on φ is only through the integrals appearing in the Sobolev norm. The dependence on Γ is only through the uniform packing radius $b/2 > 0$ appearing as the scaling factor in this norm. We stress that all quantities appearing in the estimate (2.15) are explicitly computable, and so an experimentalist can produce actual numbers on the r.h.s. of (2.15). Also the assumption of uniform positivity of the packing radius can be given up, leading to somewhat uglier estimates. For more on this see Chapter 3, Addition to Proposition 1.

Remark. Even for the independent case this bound is slightly better than the one given in [K01b]. It would seem possible to prove a result of this type by an extension of the expansion method described in [K01b], at least to certain smaller classes of weakly dependent Gibbs fields. This would be at the price of adding a huge layer of complexity to the expansions, so the concentration estimate method is definitely favorable.

3. FURTHER APPLICATION TO DIFFRACTION - PROOFS

Concentration result for quenched scatterers or quenched dislocations

It is physically important to know what happens when we have a frozen configuration of scattering amplitudes η and we are interested in the concentration of $\hat{\gamma}_{\Lambda}^{\eta,\omega}(\varphi)$ centered at its average over the dislocations ω, for fixed η. So, we have 'quenched' the η-configuration. This describes a disordered material with frozen types of scatterers that are subjected to thermal motions around their equilibrium positions. We mention that we get the valid bound for this case by the formal application of Theorem 4 (although this case is not logically contained in the statement of the Theorem). The corresponding constant in the denominator of the argument of the exponential is obtained by putting the bound on the variation of the amplitudes $\varepsilon_{sc} = 0$. So, it doesn't depend on the Sobolev norm of φ anymore but only on the Sobolev norm of its differential. Next c, c_t have to be taken as constants for the ω-distribution for that particular η.

An equal game can be played by exchanging the roles of η and ω, so that we are fixing the latter ones. Note that, when ω is fixed we are left with a model on a distorted but fixed point set $\{x+\omega_x, x \in \Gamma\}$ (with modified but positive minimal packing radius $b/2$). Thus we can assume without loss of generality that $\omega_x \equiv 0$ for all $x \in \Gamma$.

Concentration result for average over dislocations

It is physically very natural to consider a model for the joint distribution of scatterers η and dislocations ω whose joint distribution $(X,Y) \equiv (\eta,\omega)$ is of the type as described in Section 2.2. A special case for this would be a model of independent scatterers with thermal dislocations that might depend on the type of the scatterer, but we don't need independence for the scatterers.

Theorem 5. *Suppose a distribution for the scatterers X and dislocations Y as described in Section (2.2). Again we assume the uniform bounds on the scatterers and amplitudes as detailed in Theorem 4 (2.14).*

Then, the corresponding fixed-scatterer scattering image that is averaged over the dislocations obeys the universal large deviation estimate

$$\begin{aligned}&\mu^X\left(\left|\mu(\hat\gamma_\Lambda^{X,Y}(\varphi)|X) - \mu(\hat\gamma_\Lambda^{X,Y}(\varphi))\right| \ge r\right)\\ &\le 2\exp\left(-\frac{|\Lambda|\,r^2}{8}\,\frac{(1-c^X)(1-c_t^X)}{\left(\varepsilon_{sc}\|\hat\varphi\|_{\nu,b} + c^{Y,\mathit{eff}}\,\varepsilon_{dl}\|d\hat\varphi\|_{\nu,b}\right)^2}\right) \qquad \forall r \ge 0\end{aligned} \tag{3.1}$$

We have also have the total bound

$$\begin{aligned}&\mu\left(\left|\hat\gamma_\Lambda^{X,Y}(\varphi) - \mu(\hat\gamma_\Lambda^{X,Y}(\varphi))\right| \ge r\right)\\ &\le 2\exp\left(-\frac{|\Lambda|\,r^2}{8}\left[\left[\frac{(1-c^X)(1-c_t^X)}{\left(\varepsilon_{sc}\|\hat\varphi\|_{\nu,b} + c^{Y,\mathit{eff}}\,\varepsilon_{dl}\|d\hat\varphi\|_{\nu,b}\right)^2}\right]^{-1} + \left[\frac{(1-c^{Y,\infty})(1-c_t^{Y,\infty})}{\left(\varepsilon_{dl}\|d\hat\varphi\|_{\nu,b}\right)^2}\right]^{-1}\right]^{-1}\right)\end{aligned} \tag{3.2}$$

Let us now give the estimate on the l^2-norm of the variation of our function w.r.t. the scatterers and the dislocations. From this, Theorem 4 follows immediately from Theorem 1. Similarly Theorem 5 follows from Theorem 2 and Theorem 3.

Proposition 1. *Look at the function $(\eta,\omega) \mapsto \gamma_\Lambda^{\eta,\omega}(\hat\varphi)$ on the set where $|\eta_x| \le 1$ for all sites $x \in \Lambda$ and the minimal distance of the point set $\{x+\omega_x, x\in\Lambda\}$ is bigger than $b > 0$. Then we have*

$$\left\|\underline{\delta}^\eta\left(\gamma_\Lambda^{\eta,\omega}(\hat\varphi)\right)\right\|_{l^2} \le \frac{2\|\hat\varphi\|_{\nu,b}}{|\Lambda|}\left(\sum_{x\in\Lambda}[\delta(\eta_x)]^2\right)^{\frac{1}{2}} \tag{3.3}$$

and

$$\left\|\underline{\delta}^\omega\left(\gamma_\Lambda^{\eta,\omega}(\hat\varphi)\right)\right\|_{l^2} \le \frac{2\|d\hat\varphi\|_{\nu,b}}{|\Lambda|}\left(\sum_{x\in\Lambda}[\delta(\omega_x)]^2\right)^{\frac{1}{2}} \tag{3.4}$$

Proof. For each $x \in \Lambda$ we have for the variation of the non-normalized observable that

$$\delta_x^\eta\left(\sum_{x,x'\in\Lambda}\eta_x\eta_{x'}^*\hat\varphi(x-x'+\omega_x-\omega_{x'})\right) \le 2\delta(\eta_x)\times\sup_\omega\sum_{x'\in\Lambda}\left|\hat\varphi(x-x'+\omega_x-\omega_{x'})\right| \tag{3.5}$$

where we have used that $|\eta_z| \le 1$ for all z, and that $|\hat{\varphi}(x)| = |\hat{\varphi}(-x)|$. This expression is not particularly transparent, but it can be estimated in terms of the much nicer Sobolev norm. To get good estimates it is important to refrain from the temptation to put the sup inside the sum! Now, let us use the following fact that was proved as Proposition 3 in [K01b]: For any point set $\Gamma' \subset \mathbb{R}^\nu$ whose points have a minimal distance of $a > 0$ we have the estimate

$$\sum_{z\in\Gamma'} |g(z)| \le \|g\|_{\nu,a} \tag{3.6}$$

Here the norm on the r.h.s. was introduced in (2.16). This statement is reminiscent of Sobolev embedding theorems. It follows from the fact that for any ν-times differentiable function g on the unit ball B_1 around the origin one has $|g(0)| \le \frac{1}{|B_1|}\sum_{k=0}^{\nu}\frac{1}{k!}\int_{B_1}\|d^k g(y)\|dy$.

We apply this statement for the set $\Gamma'(x,\omega) \equiv \{x - x' + \omega_x - \omega_{x'}, x' \in \Gamma\}$ that includes the arguments the r.h.s. of (3.5) is summed over. It is simple but important to note that its minimal distance is bounded below by $b > 0$, independently of x and ω. So we get

$$\sum_{x'\in\Lambda}\left|\hat{\varphi}(x - x' + \omega_x - \omega_{x'})\right| \le \sum_{z\in\Gamma'(x,\omega)}\left|\hat{\varphi}(z)\right| \le \|\hat{\varphi}\|_{\nu,b} \tag{3.7}$$

This already proves the desired estimate (3.3) on the l^2-norm.

Next we show the result for the ω-variation. It is in the same spirit but there is a small trick involved. We have

$$\begin{aligned} &\delta_x^\omega\left(\sum_{x,x'\in\Lambda}\eta_x\eta_{x'}^*\hat{\varphi}(x - x' + \omega_x - \omega_{x'})\right)\\ &\le 2\sup_{\omega_{x^c}}\sup_{\omega_x,\omega'_x}\sum_{x'\in\Lambda\setminus x}\left|\hat{\varphi}(x - x' + \omega_x - \omega_{x'}) - \hat{\varphi}(x - x' + \omega'_x - \omega_{x'})\right| \end{aligned} \tag{3.8}$$

This time, for each fixed x, ω_{x^c}, and ω'_x let us define the set $\tilde{\Gamma}(x,\omega_{x^c},\omega'_x) := \{x - x' + \omega'_x - \omega_{x'}, x' \in \Gamma\backslash x\}$ including all the arguments of the second $\hat{\varphi}$-term. We note that the minimal distance between the points of any of these sets is bounded below by $b > 0$. Then we can bound the r.h.s. of (3.8) by

$$\begin{aligned} &2\sup_{\omega_x,\omega'_x}\sup_{\omega_{x^c}}\sum_{z\in\tilde{\Gamma}(x,\omega_{x^c},\omega'_x)}\left|\hat{\varphi}(z + \omega_x - \omega'_x) - \hat{\varphi}(z)\right|\\ &\le 2\sup_{|u|\le\delta(\omega_x)}\sup_{\tilde{\Gamma}}\sum_{z\in\tilde{\Gamma}}\left|\hat{\varphi}(z+u) - \hat{\varphi}(z)\right| \end{aligned} \tag{3.9}$$

where $\sup_{\tilde\Gamma}$ is over all $\tilde\Gamma$ with minimal distance $\geq b$. For $u\neq 0$ and any such $\tilde\Gamma$ we write

$$\begin{aligned}&\sum_{z\in\tilde\Gamma}\Bigl|\hat\varphi(z+u)-\hat\varphi(z)\Bigr| = |u|\sum_{z\in\tilde\Gamma}\Bigl|\int_0^1 \frac{d}{ds}\Big|_{s=0}\hat\varphi(z+tu+su/|u|)dt\Bigr|\\ &\leq |u|\int_0^1\sum_{z\in\tilde\Gamma}\Bigl|\frac{d}{ds}\Big|_{s=0}\hat\varphi(z+tu+su/|u|)\Bigr|dt \leq |u|\sup_{0\leq t\leq 1}\sum_{w\in\tilde\Gamma+tu}\Bigl|\frac{d}{ds}\Big|_{s=0}\hat\varphi(w+su/|u|)\Bigr|\end{aligned}\tag{3.10}$$

It is important to note that $\tilde\Gamma+tu$ is still a set with minimal distance $\geq b$, for any fixed t. So we can estimate the sum uniformly in t and get

$$\sum_{w\in\tilde\Gamma+tu}\Bigl|\frac{d}{ds}\Big|_{s=0}\hat\varphi(w+su/|u|)\Bigr| \leq \Bigl\|\frac{d}{ds}\Big|_{s=0}\hat\varphi(\cdot+su/|u|)\Bigr\|_{\nu,b}\leq \|d\hat\varphi\|_{\nu,b}\tag{3.11}$$

The finishes the proof of Proposition 1.

The assumption that $\{x+\omega_x, x\in\Gamma\}$ may have a positive minimal distance, μ-a.s. is not necessary for a similar estimate to hold. We will now briefly discuss what estimates can be made when the a.s. minimal distance assumption is lifted, however still assuming a.s. uniformly bounded dislocations. In fact, the reader will realize that the proof of Proposition 1 shows the a priori sharper statement (i) given below. The resulting estimate is then exploited more explicitly in statement (ii) under the assumption of bounded dislocations.

Addition to Proposition 1.

(i) For a function $g:\mathbb{R}^\nu\to\mathbb{C}$ define the norm $\|g\|_{\Gamma,\mu}$ to be the smallest number such that

$$\sup_{v\in\mathbb{R}^\nu}\sum_{x\in\Gamma+v}|g(x+Y_x)|\leq\|g\|_{\Gamma,\mu}\quad \textit{for } \mu\textit{-a.e. realization of } Y\tag{3.12}$$

A similar definition is made for a linear form dg by replacing the modulus on the l.h.s. by the norm of the linear functional at $x+Y_x$. Then, under the sole condition that $|\eta_x|\leq 1$ without any restrictions on Γ and μ, Proposition 1 holds with $\|\cdot\|_{\Gamma,\mu}$ replacing $\|\cdot\|_{\nu,b}$.

(ii) Denote the minimal distance of the unperturbed set $\Gamma\subset\mathbb{R}^\nu$ by $b_0>0$ and assume that $|Y_x|\leq R$ a.s, for any fixed arbitrarily large $R<\infty$. Then we have the (crude) estimate $\|\cdot\|_{\Gamma,\mu}\leq\bigl(2+2R/b_0\bigr)^\nu\|\cdot\|_{\nu,b_0}$.

Remark. Note that therefore **Theorem 4 and Theorem 5 have obvious extensions** obtained by the application of the Addition to Proposition 1 on the basis of the general concentration Theorems 1,2,3!

Proof of (ii). The idea is to estimate the sum on the l.h.s. of (3.12) in terms of sums of integrals over balls with fixed radii $b_0/2$ that might overlap using the statement given after (3.6). Then simply count the possible number of overlaps. Without loss put $v = 0$. Then

$$\begin{aligned}\sum_{x\in\Gamma} |g(x+Y_x)| &\leq \frac{1}{|B_1|}\sum_{k=0}^{\nu}\frac{1}{k!}\frac{1}{(b_0/2)^{\nu-k}}\int_{\mathbb{R}^\nu}\sum_{x\in\Gamma} 1_{B_{\frac{b_0}{2}}(x+Y_x)}(y)\,\|d^k g(y)\| dy \\ &\leq \Big[\sup_{y\in\mathbb{R}^\nu}\sum_{x\in\Gamma} 1_{z\in B_{\frac{a}{2}}(x+Y_x)}(y)\Big]\|g\|_{\nu,b_0} \leq \Big(2+2R/b_0\Big)^{\nu}\|g\|_{\nu,b_0}\end{aligned} \tag{3.13}$$

To understand the last inequality note that, at any point y the sum in the bracket must be smaller than the number of points in any set with minimal distance b_0 whose distance to y is smaller than $R' = R + b_0/2$. But this number is certainly bounded by the volume of the ball with radius $R' + b_0/2$ devided by the volume of the ball with radius $b_0/2$. It is obvious from this argument that the given factor could be improved by more careful counting.

4. APPLICATION TO RANDOM GIBBS MEASURES

Example: Self-averaging of free energy density for dependent disorder

Let us mention at first an application that shows exponential self-averaging of the free energy for the case of a disordered model with disorder field that obeys Dobrushin uniqueness. Note that we don't assume absence of phase transition for the spin variables of the model itself. It is a straightforward application of the basic concentration Theorem 1 and reads in the abstract setting as follows.

Corollary 1. *Suppose the random field $X = (X_x)_{x\in\Gamma}$ ('disorder field') taking values in $E_X^{\Gamma_X}$ is distributed according to a Gibbs measure μ^X that obeys the Dobrushin uniqueness condition with Dobrushin constant c^X, and also the transposed Dobrushin uniqueness condition with constant c_t^X. Suppose that Ω is a measurable space ('spin space') and ρ is a positive measure on Ω ('a priori measure on the spin-space'). Suppose that H is a real function ('Hamiltonian') on $E_X^{\Gamma_X} \times \Omega$. Define the function (corresponding 'free energy') by*

$$F(X) := -\log\Big(\int \rho(d\omega)\big(e^{-H(X,\omega)}\big)\Big) \tag{4.1}$$

whenever it exists. Then we have the Gaussian concentration estimate

$$\mu^X\Big(F(X) - \mu^X\big(F(X)\big) \geq r\Big) \leq \exp\left(-\frac{r^2(1-c^X)(1-c_t^X)}{2\|\underline{\delta}^X(H)\|_{l^2}^2}\right) \quad \forall r \geq 0 \tag{4.2}$$

This follows from the easy fact that the variation $\delta_x^X(F)$ is bounded by the partial variation $\delta_x^X(H)$. Note that the estimate can be used to prove self-averaging of the finite volume free energy density that is exponentially fast in the volume (for disordered spin systems whose Hamiltonian have bounded local variations w.r.t the disorder field X). This is clear since $\left\|\underline{\delta}^X(H)\right\|_{l^2}^2$ will be of the order Λ when H is any reasonable finite volume random Hamiltonian depending only on spin variables in Λ (while fixing a spin-boundary condition outside). Note that for very general non-local dependence of H on X this fact is still true, the precise constants depending on the specific model, of course.

Example: Pair interactions on general graphs

Let us now discuss the class of models with pair interactions on a general graph to illustrate how the various "Dobrushin-type" constants can be estimated in terms of simpler constants bounding the pair potentials themselves. Suppose that $G_X = (\Gamma_X, B_X)$ is a graph with vertex set Γ_X and set of edges (or 'bonds') B_X. Suppose that its degree is bounded my m_X. Suppose that μ^X is a measure with state-space $E_X^{\Gamma_X}$ obeying Dobrushin uniqueness and its transposed version with formal Boltzmann weight

$$\propto \exp\Big(-\sum_{\{x,y\}\in B_X} U_{x,y}(\omega_x,\omega_y)\Big)\prod_{x\in\Gamma_X}\lambda(d\eta_x) \tag{4.3}$$

with a pair potential satisfying $\sup_{\omega,\omega'}|U_{x,y}(\omega)-U_{x,y}(\omega')|\leq u$ for all $\{x,y\}\in B_X$. Then we have from (2.5) that $c^X, c_t^X \leq m_X u/2$ for the constants appearing in Theorem 1. The same would be true if there were any additional single site potential possibly differing from site to site (as long as all integrals converge).

Let us now consider a disordered (or nested) system whose fields X and Y are both of the pair potential type and see what constants arise in the chain rule estimates of Theorem 2 and Theorem 3. Let us suppose that Y is a variable whose conditional distribution $\mu(\,\cdot\,|X=\eta)$ is a Gibbs measure on a graph $G_Y = (\Gamma_Y, B_Y)$ with vertex set Γ_Y and set of edges B_Y. Suppose that its degree is bounded my m_Y. Suppose uniform Dobrushin uniqueness and its tranpose for the distribution with formal Boltzmann weight of the form

$$\propto \exp\Big(-\sum_{\{x,y\}\in B_Y} W_{x,y}(\omega_x,\omega_y,\eta_{\{x,y\}})\Big)\prod_{x\in\Gamma_Y}\lambda'(d\omega_x) \tag{4.4}$$

with a pair potential W that is a function also of an edge variable $\eta_{x,y}$. So, we assume that $\Gamma_X = B_Y$ equals the set of edges of the inner variable Y. This is the case e.g. for 'nearest neighbor' pair-interacting spin glass models on arbitrary graphs. Suppose that the X-influence on the interaction between Y's is bounded in the sense that $\sup_\omega \sup_{\eta,\eta'}|W_{x,y}(\omega,\eta)-W_{x,y}(\omega,\eta')|\leq q$. Then we have from Proposition 2 given in the section below that $C_{x,y}^{Y\leftarrow X}\leq q/2$ so that the interaction constants are bounded by $c^{Y\leftarrow X}\leq m_X q/2$ and $c_t^{Y\leftarrow X}\leq q$.

Finally, assuming that $\sup_\eta \sup_{\omega,\omega'}|W_{x,y}(\omega,\eta)-W_{x,y}(\omega',\eta)|\leq w$, we get the bound on the uniform Dobrushin constants $c^{Y,\infty}, c_t^{Y,\infty}\leq m_Y w/2$. In this way all constants appearing in the concentration Theorem 1,2,3 have been expressed in the elementary variation parameters q, w, u of the potentials and the degree of the two graphs appearing.

In the simple situation of the graph $\Gamma_Y = \mathbb{Z}^\nu$ with independent X_x's we thus have in particular $c^X = c^X_t = 0$, $c^{Y\leftarrow X} \leq \nu q$, $c^{Y\leftarrow X}_t \leq q$, and $c^{Y,\infty}, c^{Y,\infty}_t \leq \nu w$.

Simple estimates on the dependence constants

For practical use let us mention the following proposition that was already applied in the previous example.

Proposition 2. *Suppose that the conditional distribution of* Y_x *has the Gibbsian form* $\mu(d\omega_x \big| X = \eta, Y_{x^c} = \omega_{x^c}) = \exp\bigl(-H_x(\eta, \omega_x, \omega_{x^c})\bigr)\lambda(d\omega_x)/Z_x(\eta\omega_{\Gamma\setminus x})$ *where* $H_x(\eta,\omega)$ *is a function on the product space and* λ *is a* σ*-finite measure on* E_X*. Then we have that*

$$C^{Y\leftarrow X}_{x,y} \leq \frac{1}{2}\delta^X_y(H_x) \tag{4.5}$$

Proof of Proposition 2. Within the proof of Proposition 8.8 of [Geo88] the following was shown. Suppose that $\lambda^{(i)}_x(d\omega_x) = e^{u^{(i)}(\omega_x)}\lambda(d\omega_x)/\int \lambda(d\tilde{\omega}_x)e^{u^{(i)}(\tilde{\omega}_x)}$, $i = 1, 2$ are two measures on the single site space E, given in terms of the functions $u^{(i)}$. Then their variational distance can be bounded in terms of the variation of the function $u^{(1)} - u^{(2)}$ so that one has $\|\lambda^{(1)}_x - \lambda^{(2)}_x\|_x \leq \frac{1}{4}\sup_{\omega_x, \omega'_x} |u^{(1)}(\omega_x) - u^{(2)}(\omega_x) - u^{(1)}(\omega'_x) + u^{(2)}(\omega'_x)|$. But from here the Proposition is obvious.

Proof of Estimate on Dobrushin constants and transpose given in (2.5). Assuming the inequality above one sees that $C_{x,y} \leq \frac{1}{2}\sum_{A\supset\{x,y\}}\delta(\Phi_A)$ (which is also explicitly pointed out in the proof of Proposition 8.8 in [Geo88]). We point out for our purposes that it is symmetric in x, y. So one gets (2.5) from here, for *both* c^X and c^X_t.

5. PROOF OF THEOREM 1

The proof of Theorem 1 relies on an appropriate extension of the martingale method that is well-known for the case of functions of independent variables to the case of Dobrushin uniqueness. (See e.g. [Ta96] Paragraph 4 for independent variables). It is based on Lemma 1 which is a uniform estimate on the martingale differences, taken w.r.t. the filtration obtained by introducing an arbitrary order of the sites in the index set Γ. The interesting point of the proof is then to understand how the weak dependence of the Gibbs distribution can be handled, in comparison to the case of independent variables. It turns out that this can be done in a very simple and elegant way by the use of estimates of the variational distance of Gibbs-measures in the Dobrushin uniqueness regime w.r.t changes of the local specification. A clear two-pages proof of the result we need for our purposes can be found in [Geo88], we won't repeat it here and just refer to the necessary information we need as 'Fact about Dobrushin uniqueness'. This 'fact' will be exploited again in more generality below in the proof of Theorem 3.

Now, let us start with the proof. In fact we prove the following stronger (but less convenient) statement.

Theorem 1'. *Fix a bijection from the positive integers to* Γ *and denote by* $<$ *the order on* Γ *that is inherited by that bijection. Denote by* $D_< = \left(D_{x,y} 1_{x<y}\right)_{x,y\in\Gamma}$ *the triangular matrix given in terms of the geometric series* $D = \sum_{n=0}^{\infty} C^n$ *of the Dobrushin interdependence matrix.*

Suppose that F *is a real function on* E^Γ *with* $\mu\big(\exp(tF(X))\big) < \infty$ *for all real* t*. Then we have*

$$\mu\Big(F(X) - \mu\big(F(X)\big) \geq r\Big) \leq \exp\left(-\frac{r^2}{2\big\|\big(\underline{1} + [D_<]^t\big)\,\underline{\delta}(F)\big\|^2_{l^2(\Gamma)}}\right) \tag{5.1}$$

D^t denotes the transpose of a matrix D, and $\underline{1}$ is the unit matrix. Assuming this, Theorem 1 is implied for simple reasons: We first use that $\|(\underline{1}+D^t_<)v\|_2 \leq \|D^t v\|_2$ for vectors v with positive entries, because of the positivity of the matrix elements of C. Next use that $\|D^t v\|^2_{l^2} = |\sum_{x,y}(DD^t)_{x,y} v_x v_y| \leq \frac{1}{2}\sum_{x,y}(DD^t)_{x,y}(v_x^2 + v_y^2) \leq \sup_x \sum_y (DD^t)_{x,y} \|v\|^2_{l^2}$. We have that $\sup_x \sum_{y,z} D_{x,z} D_{y,z} \leq \sup_x \sum_u D_{x,u} \sup_z \sum_y D_{y,z} = \|D\|_{l^\infty} \|D\|_{l^1}$ where the last symbols denote the operator norms. Noting that the Dobrushin constant equals $c^X = \|C\|_{l^1}$ and that $c_t^X = \|C\|_{l^\infty}$ we see that the last expression is bounded by $(1-c^X)^{-1}(1-c_t^X)^{-1}$. This proves the form of the estimate given in Theorem 1.

Now let us start with the proof of the uniform bounds on the martingale differences of the function $F(X)$.

Lemma 1. *Define the decreasing sequence of* σ*-algebras by putting* $\mathcal{T}_x := \sigma\Big(X_y; y \geq x\Big)$*, for* $x \in \Gamma$*. Then the Martingale differences of the random variable* $F(X)$ *taken w.r.t this ordering obey the uniform bound*

$$\|\mu(F(X)|\mathcal{T}_x) - \mu(F(X)|\mathcal{T}_{x+1})\|_\infty \leq \delta_x(F) + \sum_{y\in\Gamma, y<x} \delta_y(F) D_{y,x} \tag{5.2}$$

Proof of Lemma 1. This estimate relies on the following piece of information (see [Geo88], Theorem 8.20).

Fact about Dobrushin uniqueness. *Suppose that* Γ *is a countable set, infinite or finite, and the random variables* $(X_x)_{x\in\Gamma}$ *are distributed according to a Gibbs measure* ρ *that obeys the Dobrushin uniqueness condition (see Introduction). Put* $D = \sum_{n=0}^{\infty} C^n$ *where* C *is the interdependence matrix of* ρ. *Suppose that we are given another Gibbs measure* $\tilde\rho$ *such that the variational distance of the single site conditional probabilities is uniformly bounded*

$$\sup_{\xi} \|\rho(\,\cdot\,|\xi) - \tilde\rho(\,\cdot\,|\xi)\|_x \le b_x \tag{5.3}$$

with constants b_x *for* $x \in \Gamma$. *Then the expectations of any function* $f(\xi)$ *on the infinite volume configurations* ξ *don't differ more than*

$$|\rho(f) - \tilde\rho(f)| \le \sum_{y,x\in\Gamma} \delta_y(f) D_{y,x} b_x \tag{5.4}$$

To show Lemma 1 let us use short notations like $\mu(F(X)|\mathcal{T}_x)(\xi) \equiv \mu(F(X)\big|\xi_{\ge x})$ etc. Now, to estimate the martingale differences in (5.2) let us write

$$\begin{aligned}
&\big|\mu(F(X)\big|\xi_{>x}\xi_x) - \mu(F(X)\big|\xi_{>x})\big| \\
&\le \left|\mu(F(X_{<x}\xi_x\xi_{>x})\big|\xi_{>x}\xi_x) - \mu(F(X_{<x}\xi_x\xi_{>x})\big|\xi_{>x})\right| \\
&+ \mu\Big(\Big|F(X_{<x}\xi_x\xi_{>x}) - F(X_{<x}X_x\xi_{>x})\Big|\Big|\xi_{>x}\Big)
\end{aligned} \tag{5.5}$$

The second term is bounded by $\delta_x(F)$. For the first term we apply the 'fact about Dobrushin uniqueness' on the conditional spin-system on the sites y in Γ with $y < x$ that is obtained from the original conditional probabilities by fixing $\xi_{>x}$. Putting $\rho(d\xi_{\le x}) = \mu(d\xi_{\le x}\big|\xi_{>x})$ and $\tilde\rho(d\xi_{\le x}) = \mu(d\xi_{\le x}\big|\xi_{>x}\xi_x)$ we have the estimate (5.3) with $b_y = 0$ for all $y < x$ and $b_x = 1$. This gives in fact that $\sup_x \sup_{\xi_{>x}\xi_x}$ over the first modulus on the r.h.s. of the last inequality is bounded by the second term in (5.4). This finishes the proof of Lemma 1. Note that, in this application, we applied the 'fact about Dobrushin uniqueness' to the finite index set of the sites that are less or equal than x. In this situation the proof of 'fact' becomes even simpler, as is easily seen by going through the short proof of Lemma 8.18 and Theorem 8.20 given in [Geo88]. It is also simple to verify that the statement holds for any, possibly degenerate kernels $\tilde\rho(\,\cdot\,|\xi)$ allowing for e.g. also for Dirac measures on specific configurations.

To complete the proof of Theorem 1' we apply Lemma A.1 (given in Appendix A) on the filtration defined in Lemma 1. To be able to do so we need that μ is trivial on the tail σ-algebra, but this is clear because it is the only Gibbs measure that is compatible with the specification defined by its conditional expectation, using Dobrushin uniqueness again. So the proof of Theorem 1' is finished.

Lemma A.1 itself, at least in the case of a finite filtration, is a simple application of the Martingale method in the context of uniformly bounded Martingale differences. However, we need to treat correctly the presence of the infinite filtration. Infinities in the filtrations are appearing also in a slightly different way in the proof of Theorem 3, so for the sake of clarity we give the results needed along with their complete proofs in Appendix A.

Remark. We remark that a term like $\mu(f(X_z)|\xi_{>x}\xi_x)-\mu(f(X_z)|\xi_{>x})$ is dangerous in the presence of a phase transition for the measure μ. Then we could not exclude that there might be discontinuous behavior, even for arbitrarily distant sites x,z, for certain $\xi_{>x}$. Therefor the proof doesn't generalize to the phase transition region.

6. PROOF OF THEOREMS 2,3

The Proof of Theorem 2 relies on Theorem 1' and another application of the 'Fact about Dobrushin uniqueness' stated in Chapter 5, along with the application of a chain rule for variations. Again, let us give the strongest version of Theorem 2 first.

Theorem 2'. *Fix a bijection from the positive integers to Γ^X and denote by $<$ the order on Γ^X that is inherited by that bijection. Denote by $D^{Y,\infty}=\sum_{n=0}^{\infty}\left(C^{Y,\infty}\right)^n$ the geometric series of the uniform Dobrushin matrix $C^{Y,\infty}_{x,y}=\sup_\eta C^Y_{x,y}(\eta)$.*

Suppose that G is a real function with $\mu\bigl(\exp(tG(X,Y))\bigr)<\infty$ for all real t.

Then we have the Gaussian concentration estimate

$$\mu\Bigl(\mu\bigl(G(X,Y)\,|X\bigr)-\mu\bigl(G(X,Y)\bigr)\geq r\Bigr)\leq\exp\left(-\frac{r^2}{2M^2}\right)\tag{6.1}$$

where

$$M=\left\|\left(\underline{1}+\left[D^X_<\right]^t\right)\underline{\delta}^X(G)+\left(\underline{1}+\left[D^X_<\right]^t\right)\left[C^{Y\leftarrow X}\right]^t\left[D^{Y,\infty}\right]^t\underline{\delta}^Y(G)\right\|_{l^2}\tag{6.2}$$

whenever this quantity is finite.

Of course the definition of $D^X_<$ is the same as $D_<$ in Theorem 1' for the marginal distribution on X.

Proof of Theorem 2'. We denote the function that is appearing in the estimate by $F(\eta):=\mu^{X=\eta}(G(\eta,Y))$ and apply Theorem 1' for that function. We need to estimate its variation $\delta^X_z(F)$. We will show that, in the sense of inequality between coordinates, we have

$$\underline{\delta}(F)\leq\underline{\delta}^X(G)+\left[C^{Y\leftarrow X}\right]^t\left[D^{Y,\infty}\right]^t\underline{\delta}^Y(G)\tag{6.3}$$

From that Theorem 2' follows by Theorem 1'.

Take any η and η' with $\eta_z = \eta'_z$, and put $G^-(\eta_{z^c}, \omega) := \inf_{\eta_z} G(\eta_{z^c}\eta_z, \omega)$. Then we have

$$\begin{aligned}
&\mu^{X=\eta}(G(\eta,Y)) - \mu^{X=\eta'}(G(\eta,Y)) \\
&\le \mu^{X=\eta}(G(\eta,Y)) - \mu^{X=\eta}(G^-(\eta_{z^c},Y)) + \mu^{X=\eta}(G^-(\eta_{z^c},Y)) - \mu^{X=\eta'}(G^-(\eta_{z^c},Y)) \\
&\le \delta_z^X(G) + \sup_{\bar\eta} \delta_z^X \left(\mu^{X=\eta}\big(G(\bar\eta,Y)\big)\right)
\end{aligned} \tag{6.4}$$

To control the variation of the conditional spin system when we change its local specification by changing the X-variable we need to use again the 'Fact about Dobrushin uniqueness'.

Denoting $\rho(d\omega) = \mu^{X=\eta_{\Gamma\setminus z}\eta_z}(d\omega)$ and $\tilde\rho(d\omega) = \mu^{X=\eta_{\Gamma\setminus z}\eta'_z}(d\omega)$ we have to put $b_x \le C^{Y\leftarrow X}_{x,z}$ in the statement of the 'fact' controlling the change in the local specifications caused by a single site variation of X. For fixed $\bar\eta$ we set $f(\omega) := G(\bar\eta, \omega)$ so that we get from the 'fact'

$$\left|\mu^{X=\eta}\big(f(Y)\big) - \mu^{X=\eta'}\big(f(Y)\big)\right| \le \sum_{y\in\Gamma^Y} \delta_y^Y(f) \sum_{x\in\Gamma^Y} D^{Y,\infty}_{y,x} C^{Y\leftarrow X}_{x,z} \tag{6.5}$$

Collecting terms and using vector notation the desired inequality for $\underline{\delta}(F)$ follows.

Assuming this, Theorems 2 is obtained from Theorem 2' by an analogous estimate on M as Theorem 1 is obtained from Theorem 1'. Using the triangle inequality and splitting off the common matrix we are left with the new term

$$\|\big[C^{Y\leftarrow X}\big]^t\big[D^{Y,\infty}\big]^t v\|_{l^2}^2 \le \|C^{Y\leftarrow X}\big[C^{Y\leftarrow X}\big]^t\|_{l^2}^2 \times \|\big[D^{Y,\infty}\big]^t v\|_{l^2}^2 \tag{6.6}$$

The first factor ist bounded by $c^{Y\leftarrow X}c_t^{Y\leftarrow X}$. The second factor has already been seen to be bounded by $(1-c^{Y,\infty})^{-1}(1-c_t^{Y,\infty})^{-1}\|v\|_{l^2}^2$.

Proof of Theorem 3. To prove Theorem 3 we need a double filtration. Define the filtration $\mathcal{T}_x^{(1)} := \sigma\Big(X_y; y \ge x\Big)$ on the probability space E^{Γ_X} and the filtration $\mathcal{T}_x^{(2)} := \sigma\Big(Y_y; y \ge x\Big)$ on the probability space E^{Γ_Y}. Then Lemma A.2 tells us that we can treat them like they were finite filtrations if the function in question has exponential moments and we have bounds on their martingale differences. Now, the martingale differences in the first line of (A.3) are controlled by (5.2) applied to the conditional distribution of Y given any fixed configuration of X. The martingale differences in the second line of (A.3) are controlled in terms of (6.3). Collecting terms Theorem 3 follows.

A different (although less natural) way to prove the 'total concentration result' of Theorem 3 would be to prove that the joint distribution can be represented as a Gibbs-measure for the joint variables $\xi_x = (\eta_x \omega_x)$, estimate its joint constants c, c_t, and then apply Theorem 4. Note in this context that it won't be true in general that the resulting measure is a Gibbs measure, even for independent X_x's, when one allows for conditional Gibbsian distributions of the Y-variables having phase transitions (which is however excluded here). For more on this, see the research in [EMSS], [K99], [K01].

APPENDIX

Lemma A.1. *Suppose that* $(\Omega, \mathcal{T}_0, \mu)$ *is a probability space. Suppose that* $(\mathcal{T}_i)_{i=0,1,2,\dots}$ *is a decreasing sequence of* σ*-algebras such that* μ *is trivial on the tail-*σ*=algebra* $\mathcal{F}_\infty := \bigcap_{i=0,1,2,\dots} \mathcal{T}_i$. *Suppose that* Z *is a real random variable on* Ω *such that* $\mu(\exp(tZ)) < \infty$ *for all real* t. *Assume that* Z *has uniformly bounded martingale differences*

$$\|\mu(Z|\mathcal{T}_i) - \mu(Z|\mathcal{T}_{i+1})\|_\infty \le M_i \tag{A.1}$$

Then we have the exponential concentration estimate

$$\mu\Big(Z - \mu(Z) \ge a\Big) \le \exp\left(-\frac{a^2}{2\sum_{i=0}^\infty M_i^2}\right) \tag{A.2}$$

Remark. Tail triviality is needed! Otherwise $\mu(Z)$ must be replaced by $\mu(Z|\mathcal{T}_\infty)$ in the l.h.s. of the estimate.

Remark. If the sum in the denominator of the argument of the exponential does not converge, the statement is empty, obviously. In the case of a finite filtration $(\mathcal{T}_i)_{i=0,1,\dots,n}$ the statement is applied by putting $\mathcal{T}_i := \mathcal{T}_n$ for $i \ge n$.

Lemma A.2. *Suppose that* $(\Omega^{(1)}, \mathcal{T}_0^{(1)})$ *and* $(\Omega^{(2)}, \mathcal{T}_0^{(2)})$ *are measurable spaces. Denote by* $(\Omega, \mathcal{F}_0, \mu)$ *the corresponding product space with the product* σ*-algebra where the distribution* μ *has the form* $\mu(d\omega^{(1)} d\omega^{(2)}) = \mu^1(d\omega^{(1)})\mu^{(2)}(d\omega^{(2)}|\omega^{(1)})$ *with a probability measure on the first space and a probability kernel from the first to the second space.*

Suppose that $(\mathcal{T}_i^{(k)})_{i=0,1,2,\dots}$ *are two decreasing sequence of* σ*-algebras on the spaces* $\Omega^{(k)}$ *such that (a) the measure* $\mu^{(1)}$ *is trivial on the tail-*σ*=algebra* $\mathcal{F}_\infty^{(1)} := \bigcap_{i=0,1,2,\dots} \mathcal{T}_i^{(1)}$

and (b) the measure $\mu^{(2)}(\,\cdot\,|\omega^{(1)})$ *is trivial on the tail-*σ*=algebra* $\mathcal{F}^{(2)}_\infty := \bigcap_{i=0,1,2,\dots} \mathcal{T}^{(2)}_i$ *for any* $\mu^{(1)}$*-a.e.* $\omega^{(1)}$*.*

Suppose that Z *is a real random variable on* Ω *such that* $\mu\left(\exp(tZ)\right) < \infty$ *for all real* t*. Assume that* Z *has uniformly bounded martingale differences*

$$\begin{aligned} &\|\mu(Z|\mathcal{T}^{(1)}_0 \otimes \mathcal{T}^{(2)}_i) - \mu(Z|\mathcal{T}^{(1)}_0 \otimes \mathcal{T}^{(2)}_{i+1})\|_\infty \le M_i, \quad \forall i = 0, 1, \dots \\ &\|\mu(Z|\mathcal{T}^{(1)}_j \otimes \mathcal{T}^{(2)}_\infty) - \mu(Z|\mathcal{T}^{(1)}_{j+1} \otimes \mathcal{T}^{(2)}_\infty)\|_\infty \le L_j, \quad \forall j = 0, 1, \dots \end{aligned} \tag{A.3}$$

Then we have the exponential concentration estimate

$$\mu\Big(Z - \mu(Z) \ge a\Big) \le \exp\left(-\frac{a^2}{2\sum_{i=0}^\infty (M_i^2 + L_i^2)}\right) \tag{A.4}$$

Proof of Lemma A.1. We will show that

$$\mu\Big(e^{t(Z-\mu(Z))}\Big) \le e^{\frac{t^2}{2}\sum_{i=0}^\infty M_i^2} \tag{A.5}$$

From this the estimate on the probabilities follows in the standard way from the exponential Markov inequality saying that for all $t \ge 0$ in the form $\mu\left(Z - \mu(Z) \ge a\right) \le e^{-ta}\mu\Big(e^{t(Z-\mu(Z|\mathcal{T}_\infty))}\Big)$ by optimizing the bound (A.5) over t.

Now, to show (A.5) one puts $t \equiv 1$ without loss and estimates the Laplace transform

$$\begin{aligned} &\mu\Big(e^{Z-\mu(Z|\mathcal{T}_1)}e^{\mu(Z|\mathcal{T}_1)-\mu(Z)}\Big) = \mu\Big(\mu[e^{Z-\mu(Z|\mathcal{T}_1)}|\mathcal{T}_1] \times e^{\mu(Z|\mathcal{T}_1)-\mu(Z)}\Big) \\ &\le \Big\|\mu[e^{Z-\mu(Z|\mathcal{T}_1)}|\mathcal{T}_1]\Big\|_\infty \mu\Big(e^{\mu(Z|\mathcal{T}_1)-\mu(Z|\mathcal{T})}\Big) \end{aligned} \tag{A.6}$$

The supremum over the conditional Laplace transform of the first martingale difference is estimated in terms of the uniform bound M_0. Estimating the higher terms of the Taylor series of the exponential by the use of the uniform bound from the hypothesis and noting that the linear term vanishes one gets that

$$\Big\|\mu[e^{Z-\mu(Z|\mathcal{T}_1)}|\mathcal{T}_1]\Big\|_\infty \le e^{M_0} - M_0 \le e^{\frac{M_0^2}{2}} \tag{A.7}$$

where the latter bound is elementary. From that we get by iteration

$$\mu\Big(e^{Z-\mu(Z)}\Big) \le e^{\frac{1}{2}\sum_{i=0}^{N-1} M_i^2}\mu\Big(e^{Y_N}\Big) \tag{A.8}$$

where $Y_N = \mu(Z|\mathcal{T}_N) - \mu(Z)$. To show (A.1) we show that $\lim_{N\uparrow\infty}\mu\Big(e^{Y_N}\Big) = 1$. To see this, note at first that, by the backwards Martingale theorem (see e.g. Bauer Theorem 60.8) we know that, μ-a.s. $\lim_{N\uparrow\infty}\mu(Z|\mathcal{T}_N) = \mu(Z|\mathcal{T}_\infty)$. But since we assumed that μ

is trivial on $\mathcal{T}_\infty$ this means $\lim_{N\uparrow\infty} Y_N = 0$ μ-a.s. So one has $\lim_{N\uparrow\infty} \mu(e^{Y_N} 1_{Y_N \le \lambda}) = 0$ for all fixed λ. But from this follows the convergence of the full integrals because of the uniform estimate

$$\begin{aligned} &\sup_{N=0,1,\dots} \mu\Big(e^{Y_N} 1_{Y_N \ge \lambda}\Big) \le e^{-\lambda} \sup_{N=0,1,\dots} \mu\Big(e^{2Y_N}\Big) \\ &\le e^{-\lambda} \mu\Big(e^{2\mu(Z|\mathcal{T}_N)}\Big)^{\frac{1}{2}} \mu\Big(e^{-2\mu(Z|\mathcal{T}_\infty)}\Big)^{\frac{1}{2}} \le e^{-\lambda} \mu\Big(e^{2Z}\Big)^{\frac{1}{2}} \mu\Big(e^{-2Z}\Big)^{\frac{1}{2}} < \infty \end{aligned} \tag{A.9}$$

where the last inequality is Jensen's inequality.

Proof of Lemma 3. We need to show that $\mu\Big(e^{t(Z-\mu(Z))}\Big) \le e^{\frac{t^2}{2}\sum_{i=0}^{\infty}(M_i^2+L_i^2)}$.

Now, we write the Laplace transform as

$$\mu\Big(e^{t(Z-\mu(Z))}\Big) = \mu\Big(e^{Z-\mu(Z|\mathcal{T}_0^{(1)}\otimes\mathcal{T}_\infty^{(2)})} e^U\Big) \tag{A.10}$$

with $U = \mu(Z|\mathcal{T}_0^{(1)} \otimes \mathcal{T}_\infty^{(2)}) - \mu(Z)$. With the same arguments as the ones leading to (A.8) one gets that

$$\mu\Big(e^{Z-\mu(Z)}\Big) \le e^{\frac{1}{2}\sum_{i=0}^{N-1} M_i^2} \mu\Big(e^{V_N} e^U\Big) \le e^{\frac{1}{2}\sum_{i=0}^{\infty} M_i^2} \Big[\mu\Big(e^U\Big) + \mu\Big((e^{V_N} - 1)e^U\Big)\Big] \tag{A.11}$$

where $V_N = \mu(Z|\mathcal{T}_0^{(1)} \otimes \mathcal{T}_N^{(2)}) - \mu(Z|\mathcal{T}_0^{(1)} \otimes \mathcal{T}_\infty^{(2)})$. We can apply the martingale decomposition for $\mu\Big(e^U\Big)$ from which follows that $\mu\Big(e^U\Big) \le e^{\frac{1}{2}\sum_{i=0}^{\infty} L_i^2}$, using tail-triviality. So, we need to show that the second term in the last parenthesis converges to zero with $N \uparrow \infty$. But this follows from the backwards martingale convergence theorem, tail triviality and existence of all exponential moments in an analogous fashion as in the proof of Lemma A.1.

Acknowledgments: I am grateful to M.Baake for his motivation of the study of diffraction patterns and to A.Bovier for valuable discussions about concentration inequalities. This work was supported by the DFG Schwerpunkt 'Wechselwirkende stochastische Systeme hoher Komplexität'.

REFERENCES:

[BaaHoe00] M.Baake, M.Höffe, *"Diffraction of Random Tilings: Some Rigorous Results"*, J. Stat. Phys. **99**, Nos. 1/2, 219-261 (2000)

[BaaMoo98] M.Baake and R.V.Moody, *"Diffractive point sets with entropy"*, J. Phys. A **31**, 9023-9039 (1998)

[BaaMoo01] M.Baake and R.V.Moody, *"Weighted Dirac combs with pure point diffraction"*, preprint (2001)

[Ce01] Filippo Cesi, *"Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields"*, Probab. Theor. Rel. Fields **120**, 569-584 (2001)

[Do68] R.L.Dobrušin, *"Description of a random field by means of conditional probabilities and conditions for its regularity"*, Teor. Verojatnost. i Primenen **13** 201-229 (1968)

[DS84] R.L.Dobrushin and S.B.Shlosman, in *Statistical physics and dynamical systems (Köszeg, 1984)*, 371-403, Birkhäuser, Boston, Boston, MA, 1985

[EnMi92] A.C.D. van Enter, J.Miekisz, *"How should one define a (weak) crystal?"*, J. Stat. Phys. **66**, 1147-1153 (1992)

[Geo88] H.O.Georgii, *"Gibbs measures and Phase transitions"*, de Gruyter, Berlin (1988)

[He00] *"Properties of Models for aperiodic solids"*, D.J.L.Herrmann, Ph.D. thesis, Nijmegen (2000)

[Hof95a] A.Hof, *"Diffraction by aperiodic structures at high temperatures"*, J. Phys. A **28**, 57-62 (1995)

[Hof95b] A.Hof, *"On diffraction by aperiodic structures"*, Commun. Math. Phys. **169**, 25-43 (1995)

[K99b] C.Külske, *(Non-) Gibbsianness and Phase Transitions in Random Lattice Spin Models*, Markov Proc. Rel. Fields **5** 357-383 (1999)

[K01a] C.Külske, *Weakly Gibbsian Representations for joint measures of quenched lattice spin models*, Probab. Theor. Rel. Fields **119**, 1-30 (2001).

[K01b] C.Külske, *"Universal bound on the selfaveraging of random diffraction measures"*, WIAS-preprint 676, available as preprint math-ph/0109005 at http://lanl.arXiv.org/

[Le01] M.Ledoux, *"The concentration of measure phenomenon"*, Mathematical Surveys and Monographs 89, American Mathematical Society (2001)

[LT91] M.Ledoux and M.Talagrand, *Probability in Banach spaces*, Springer, Berlin (1991)

[Ma98] K.Marton, *"Measure concentration for a class of random processes"*, Probab. Theor. Rel. Fields **110**, 427-439 (1998)

[M00] M.Schlottmann, Martin, *"Generalized model sets and dynamical systems"*, in: Directions in mathematical quasicrystals, 143-159, CRM Monogr.Ser., 13, AMS, Providence, (2000)

[Sa00] P.-M.Samson, *"Concentration of measure inequalities for Markov chains and Φ-mixing processes"*, Ann.Probab. **28** no. 1, 416-461 (2000)

[SZ92] D.W.Stroock, B.Zegarlinski, *"The Equivalence of the Logarithmic Sobolev Inequality and the Dobrushin-Shlosman Mixing Condition"*, Commun. Math. Phys. **144**, 303-323 (1995)

[Ta96] M.Talagrand, *"A New Look at Independence"*, Ann.Probab. **24**, 1-34 (1996)