Large deviations for quantum spin systems

Large deviations for quantum spin systems

K. Netočný*
F. Redig†

April 6, 2004

Abstract: We consider high temperature KMS states for quantum spin systems on a lattice. We prove a large deviation principle for the distribution of empirical averages $\overline{X}_\Lambda := \frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i$, where the X_i's are copies of a self-adjoint element X (level one large deviations). From the analyticity of the generating function, we obtain the central limit theorem. We generalize to a level two large deviation principle for the distribution of $\frac{1}{|\Lambda|}\sum_{i\in\Lambda}\delta_{X_i}$.

Keywords: large deviation principle, central limit theorem, boundary terms, cluster expansion, Goldon-Thompson inequality.

1 Introduction

Large deviations for classical lattice spin systems constitutes by now a rather complete theory, see e.g. [6], [13]. In particular, for Gibbsian random fields, it is well-known that the relative entropy density governs the large deviations of the empirical measure, see e.g. [5], [13], [4]. The relative entropy density is the Legendre transform of a generating function which is a difference of pressures. For instance, if one studies the large deviations of the magnetization in a Gibbs measure with Hamiltonian H, one has to consider the generating function $F(t) = P(H + th) - P(H)$, where h is a magnetic field Hamiltonian. The Legendre transform of F gives the entropy function I of the large deviations of the magnetization, i.e.,

$$\mathbb{P}_H(\frac{1}{|\Lambda|}\sum_{i\in\Lambda}\sigma_i \simeq a) = e^{-|\Lambda|I(a)}e^{o(|\Lambda|)}$$

where $\sigma_i, i \in \mathbb{Z}^d$ is the value of the lattice spin at site i.

*EURANDOM, Postbus 513, 5600 MB Eindhoven, The Netherlands,
†Faculteit Wiskunde en Informatica and EURANDOM, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, The Netherlands, f.h.j.redig@tue.nl

For quantum lattice spin systems a similar large deviation question can be asked. The σ_i have to be replaced by self-adjoint operators X_i, and the probability measure $\mathbb{P}_H$ has to be replaced by a (KMS)-state. We are then interested in the "probability"

$$\omega\Big(1_A\Big(\frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i\Big)\Big) \tag{1.1}$$

where ω is a KMS state, X_i are copies of an observable X at site i, A is a Borel subset of $[-\|X\|, \|X\|]$, and $\Lambda \subseteq \mathbb{Z}^d$ is a (large) volume.

Surprisingly, such probabilities have not been considered in the literature on quantum spin systems. Laws of large numbers, and central limit theorems have been considered, e.g. in [16], [17]. It is well known that for large volumes Λ, the empirical average $\frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i$ is well-approximated (in the state ω) by $\omega(X_0)\mathrm{Id}$ provided the state ω is mixing. Therefore, it is a natural question to ask whether the probabilities in (1.1) are of the form $\exp(-|\Lambda| \inf_{a\in A} I(a))$, for some entropy function I. In the context of non-interacting bosons, this question has been studied in [10], later generalized in [11] to weakly interacting bosons and fermions, where one considers large deviations of the particle density.

In this paper we prove the large deviation principle for empirical averages of the form $\frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i$ for high temperature KMS states ω (i.e., so-called level-1 large deviations), and give a generalization to level-2 large deviations, i.e., large deviations for distribution of the "measures" $\mathcal{L}_\Lambda = \frac{1}{|\Lambda|}\sum_{i\in\Lambda} \delta_{X_i}$ under the state ω. The existence of the generating function of the large deviations of $\frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i$ is not as obvious as in the classical lattice spin context (unless X commutes with the Hamiltonian of the KMS state). In fact this generating function is not a difference of two pressures, simply obtained by perturbing the Hamiltonian of the original KMS state by a magnetic field Hamiltonian. We show that the entropy function obtained by proceeding as in the classical case is (strictly) larger than the true entropy function. The reason we limit ourselves to high-temperature states is the use of a polymer expansion. This polymer expansion can be set up because we study the large deviations of averages of *one-point* observables.

Our paper is organized as follows. In section 2 we set up basic notation, specify our problem, and state the main result of the paper. In section 3 we consider the easy case of product states, in section 4 we compare with the classical case. In section 5 we show that the classical proof of existence of pressure does not work if we want to show existence of the generating function for the large deviations of empirical averages. In section 6 we set up the cluster expansion, the basic technical tool to obtain both existence, "boundary condition independence" and analyticity of the generating function. In Section 7 we prove the main theorem and point out a generalization to empirical averages of local (not necessarily single site) observables. Finally, in section 8 we prove level 2 large deviations.

2 The Problem

Let M be a finite dimensional algebra of complex matrices. For $\Lambda \subseteq \mathbb{Z}^d$ we define the local algebra

$$\mathcal{U}_\Lambda := \otimes_{i\in\Lambda} M_i \tag{2.1}$$

where each M_i is a copy of M. The algebra of local observables is defined as the inductive limit of the $\mathcal{U}_\Lambda$'s, and is denoted by $\mathcal{U}$. Let $X \in M$ be a fixed self-adjoint element, and consider for $\Lambda \subseteq \mathbb{Z}^d$ the empirical average $\frac{X_\Lambda}{|\Lambda|}$, where

$$X_\Lambda = \sum_{i\in\Lambda} X_i \tag{2.2}$$

and X_i's are copies of X in M_i. Suppose we are given a faithful state ω on $\mathcal{U}$. Given $A \in \mathcal{U}$ such that $A = A^*$, we can consider a probability measure on the spectrum of A, defined by

$$\int_{\sigma(A)} \mathbb{P}_A(dx) f(x) = \omega(f(A)), \tag{2.3}$$

for $f : \sigma(A) \to \mathbb{R}$ continuous. In particular, for $F \subseteq \sigma(A)$ a Borel measurable subset of the spectrum, we have

$$\mathbb{P}_A(F) = \omega(1_F(A)). \tag{2.4}$$

We call $\mathbb{P}_A$ the distribution of A. Given a self-adjoint element $X \in M$, we are interested in the probability measures associated to the empirical averages, i.e.,

$$\mathbb{P}^X_\Lambda := \mathbb{P}_{\frac{1}{|\Lambda|}X_\Lambda}. \tag{2.5}$$

$\mathbb{P}^X_\Lambda$ are probability measures on $[-\|X\|, \|X\|]$, i.e., they have compact support and hence always contain convergent subsequences. If the state ω is sufficiently mixing, then $\mathbb{P}^X_\Lambda$ converges weakly to the Dirac measure $\delta_{\omega(X_0)}$, concentrating on $\omega(X_0)$.

Therefore it is natural to ask whether the sequence $\{\mathbb{P}^X_\Lambda : \Lambda \subseteq \mathbb{Z}^d\}$ satisfies a large deviation principle. This means there exists a lower-semicontinuous convex function $I : [-\|X\|, \|X\|] \to \mathbb{R}$ such that

$$\begin{aligned} \limsup_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathbb{P}^X_\Lambda(F) &\leq -\inf_{x\in F} I(x) \text{ for } F \subseteq \mathbb{R} \text{ closed}, \\ \liminf_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathbb{P}^X_\Lambda(G) &\geq -\inf_{x\in G} I(x) \text{ for } G \subseteq \mathbb{R} \text{ open}. \end{aligned} \tag{2.6}$$

In some sense (2.6) is purely a property of a particular sequence of probability measures with compact support. Therefore, a sufficient condition is the existence of a differentiable generating function

$$\begin{aligned} F(t) &:= \lim_{\Lambda\uparrow\mathbb{Z}^d} \log \frac{1}{|\Lambda|} \int \mathbb{P}^X_\Lambda(dx) e^{t|\Lambda|x} \\ &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega\left(e^{t\sum_{i\in\Lambda} X_i}\right). \end{aligned} \tag{2.7}$$

More precisely, following [6], we have the following standard result:

Proposition 2.8. *If for all $t \in \mathbb{R}$, $F(t)$ exists and is differentiable in t, then the large deviation principle (2.6) holds and the entropy function is*

$$I(x) = \sup_{t\in\mathbb{R}}(xt - F(t)). \tag{2.9}$$

Differentiability in t can be replaced by strict convexity of I. Even if F is not differentiable in t, the large deviation upper bound holds, but the lower bound may fail (see [6] for a counterexample).

We now define what we mean by the central limit theorem in our context.

Definition 2.10. *We say that a collection of operators W_Λ, $\Lambda \subseteq \mathbb{Z}^d$ satisfies the central limit theorem if there exists $\sigma^2 > 0$, such that for all $t \in \mathbb{R}$*

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \omega\left(e^{itW_\Lambda}\right) = e^{-t^2\sigma^2/2} \tag{2.11}$$

Bryc's theorem [1] gives a connection between the large deviation principle and the central limit theorem. In our context this means that if F exists in a neighborhood of the origin in the complex plane, then the central limit theorem (2.11) holds, with $W_\Lambda = \frac{1}{\sqrt{|\Lambda|}}\sum_{i\in\Lambda}(X_i - \omega(X_i))$ but possibly $\sigma^2 = 0$, in which case the statement is empty. If the sum

$$\chi_X^2 = \sum_{i\in\mathbb{Z}^d} \omega\left((X_i - \omega(X)(X_0 - \omega(X))\right)$$

converges absolutely, then $\sigma^2 = \chi_X^2$.

2.1 High temperature KMS states

The states we consider in this paper are KMS states for a translation invariant finite range potential at high temperature. This is a collection of self-adjoint $\Phi(A) \in \mathcal{U}_A$, indexed by finite subsets $A \subseteq \mathbb{Z}^d$ with the following two properties.

1. Translation invariance: $\Phi(A + i) = \tau_i\Phi(A)$
2. Finite range: there exists $R > 0$ such that if $\mathrm{diam}(A) > R$, then $\Phi(A) = 0$.

Later on we will see that we can slightly relax the finite range condition, see (7.1) below. The KMS-state associated to the potential Φ at inverse temperature β is defined as the limit of the finite volume states on $\mathcal{U}_\Lambda$

$$\omega_\Lambda^\beta(X) = \frac{\mathrm{Tr}(Xe^{-\beta H_\Lambda^\Phi})}{\mathrm{Tr}(e^{-\beta H_\Lambda^\Phi})} \tag{2.12}$$

where the Hamiltonians H_Λ^Φ are defined by

$$H_\Lambda^\Phi = \sum_{A\subseteq\Lambda} \Phi(A) \tag{2.13}$$

Remark 2.14. *The KMS-states we consider are defined by the limit of (2.12) as $\Lambda \uparrow \mathbb{Z}^d$. In that way we avoid the question of uniqueness of KMS-states.*

In our context there exists β_0' small enough such that there exists a unique KMS-state, which is possible by the finite range property (or by its generalization (7.1), see e.g. proposition 6.2.45 in [2], but this β_0' depends on the dimension of the single site algebra, and is possibly smaller than the β_0 of our main result stated below.

We can now state our main result.

Theorem 2.15. *1. There exists β_0 independent of X such that for all $\beta < \beta_0$ the generating function*

$$F(z) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega \left(e^{z \sum_{i \in \Lambda} X_i} \right) \tag{2.16}$$

exists and is analytic in a strip $\{z = x + iy \in \mathcal{C} : |y| < \delta\}$.

2. The large deviation principle (2.6) holds.

3. The central limit theorem (2.11) holds for the operators

$$W_\Lambda = \frac{1}{\sqrt{|\Lambda|}} \sum_{i \in \Lambda} (X_i - \omega(X_0)). \tag{2.17}$$

3 Non-interacting case: product states

The simplest situation is the case

$$\omega = \otimes_i \omega_i, \tag{3.1}$$

where ω_i are copies of a faithful state on M, i.e., there exists A such that for $X \in M$:

$$\omega_0(X) = \frac{\mathrm{Tr}(Xe^{-A})}{\mathrm{Tr}(e^{-A})} \tag{3.2}$$

The generating function (2.7) is

$$F(z) = \log \left(\frac{\mathrm{Tr}[e^{zX} e^{-A}]}{\mathrm{Tr}[e^{-A}]} \right), \tag{3.3}$$

which is clearly defined and analytic on the strip $\{z = x + iy \in \mathcal{C} : |y| < \delta\}$ for δ small enough, and

$$\frac{dF}{dz} = \frac{\mathrm{Tr}[Xe^{zX} e^{-A}]}{\mathrm{Tr}[e^{zX} e^{-A}]} \tag{3.4}$$

In that case the distribution of $X_\Lambda/|\Lambda|$ is the same as that of the $\frac{1}{|\Lambda|} \sum_{i \in \Lambda} \tilde{X}_i$ where $\tilde{X}_i$ are i.i.d. with distribution $\mathbb{P}_X$. Hence the large deviation principle (2.6) is clearly satisfied with entropy function

$$I(x) = \sup_{t \in \mathbb{R}} (tx - F(t)). \tag{3.5}$$

4 Comparison with the classical case

In the classical Gibbs formalism, there is a natural way to obtain large deviation probabilities by perturbing the Hamiltonian with a magnetic field potential ("Cramer tilting"). Let us informally follow this procedure in our context. For simplicity we put $\beta = 1$ in this section. If we want to know the probability of the event $\frac{X_\Lambda}{|\Lambda|} \simeq a$, then we perturb the hamiltonian H_Λ^Φ with an external field $h_a X_\Lambda$ to make the value a "typical", i.e., such that

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{\mathrm{Tr}[X_0 e^{-H_\Lambda^\Phi - h_a X_\Lambda}]}{\mathrm{Tr}[e^{-H_\Lambda^\Phi - h_a X_\Lambda}]} = a \tag{4.1}$$

The rate function can then (again informally) be obtained as follows:

$$\begin{aligned} & \omega_\Lambda\left(1_{(a-\epsilon,a+\epsilon)}(\frac{X_\Lambda}{|\Lambda|})\right) \\ = \; & \frac{\mathrm{Tr}\left(e^{-H_\Lambda^\Phi - h_a X_\Lambda}[e^{H_\Lambda^\Phi + h_a X_\Lambda} e^{-H_\Lambda^\Phi}] 1_{(a-\epsilon,a+\epsilon)}(\frac{X_\Lambda}{\Lambda})\right)}{\mathrm{Tr}\left(e^{-H_\Lambda^\Phi - h_a X_\Lambda}\right)} \frac{\mathrm{Tr}[e^{-H_\Lambda^\Phi - h_a X_\Lambda}]}{\mathrm{Tr}[e^{-H_\Lambda^\Phi}]}. \end{aligned} \tag{4.2}$$

Define the pressure

$$P(\Phi) := \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathrm{Tr}[e^{-H_\Lambda^\Phi}], \tag{4.3}$$

and

$$P(\Phi, h_a) := \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathrm{Tr}[e^{-H_\Lambda^\Phi - h_a X_\Lambda}]. \tag{4.4}$$

We can rewrite (4.2) as

$$\omega_\Lambda\left(1_{(a-\epsilon,a+\epsilon)}(\frac{X_\Lambda}{|\Lambda|})\right) = \frac{\mathrm{Tr}[e^{-H_\Lambda^\Phi - h_a X_\Lambda} 1_{(a-\epsilon,a+\epsilon)}(\frac{X_\Lambda}{|\Lambda|})]}{\mathrm{Tr}[e^{-H_\Lambda^\Phi - h_a X_\Lambda}]} e^{|\Lambda|(ah_a + P(\Phi,h_a) - P(\Phi))} e^{o(|\Lambda|)} \tag{4.5}$$

Since h_a is chosen according to (4.1), the first factor in (4.5) is close to one, and we obtain

$$\log \omega_\Lambda\left(1_{(a-\epsilon,a+\epsilon)}(\frac{X_\Lambda}{|\Lambda|})\right) = |\Lambda|(ah_a + P(\Phi, h_a) - P(\Phi)) + o(|\Lambda|) \tag{4.6}$$

This suggests as a rate function

$$\tilde{I}(a) = -[h_a a + P(\Phi, h_a) - P(\Phi)], \tag{4.7}$$

which is the Legendre transform of

$$\tilde{F}(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \frac{\mathrm{Tr}[e^{t\sum_{i\in\Lambda} X_i - H_\Lambda^\Phi}]}{\mathrm{Tr}[e^{-H_\Lambda^\Phi}]}. \tag{4.8}$$

This argument leading to $\tilde{I}, \tilde{F}$ is of course informal, but in the classical case it is easy and standard to make it rigorous in order to obtain the lower bound.

By the notation $\tilde{I}$ we suggest that $\tilde{I}$ is not the entropy function I we are looking for. Indeed, if the large deviation principle (2.6) holds for $X_\Lambda/|\Lambda|$, then the only candidate for I is the Legendre transform of

$$F(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \frac{\mathrm{Tr}_\Lambda[e^{t\sum_{i\in\Lambda} X_i} e^{-H_\Lambda^\Phi}]}{\mathrm{Tr}_\Lambda[e^{-H_\Lambda^\Phi}]} \tag{4.9}$$

By the Golden-Thompson inequality we have

$$\tilde{F}(t) \leq F(t), \tag{4.10}$$

and hence

$$\tilde{I}(x) \geq I(x) \tag{4.11}$$

Therefore, the large deviation principle of Theorem 2.15 implies the following.

Proposition 4.12. *For $G \subseteq \mathbb{R}$ open,*

$$\liminf_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathbb{P}_\Lambda^X(G) \geq - \inf_{x\in G} \tilde{I}(x), \tag{4.13}$$

where $\tilde{I}(x)$ is defined in (4.7).

If X and H_Λ^Φ commute, then (4.10) becomes an equality and $\tilde{I}(x)$ is actually the true entropy function, but in the case $[X, H_\Lambda^\Phi] \neq 0$, the inequality (4.11) can be strict.

Notice that even in the simplest case of product states of the previous section, $I \neq \tilde{I}$ as soon as A and X do not commute. A possible explanation here is that the "perturbed states" obtained by adding a magnetic field potential to the Hamiltonian are not the right states to make the large deviation event typical.

5 Boundary terms

In the previous section we considered as a candidate generating function

$$F_f(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \frac{\mathrm{Tr}_\Lambda[e^{t\sum_{i\in\Lambda} X_i} e^{-\beta H_\Lambda^\Phi}]}{\mathrm{Tr}_\Lambda[e^{-\beta H_\Lambda^\Phi}]}, \tag{5.1}$$

where we now add the subindex f to denote free boundary conditions. The reader might have noticed that we should have written, following (2.7):

$$\begin{aligned} F(t) &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega[e^{t\sum_{i\in\Lambda} X_i}] \\ &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \lim_{\Lambda'\uparrow\mathbb{Z}^d} \log \frac{\mathrm{Tr}_{\Lambda'} \left(\exp\left(t\sum_{i\in\Lambda} X_i\right) \exp\left(-\beta H_{\Lambda'}^\Phi\right)\right)}{\mathrm{Tr}_{\Lambda'} \exp\left(-\beta H_{\Lambda'}^\Phi\right)}, \end{aligned} \tag{5.2}$$

The difference between $F_f(t)$ and $F(t)$ is caused by a boundary term and hence it is expected to vanish in the thermodynamic limit, i.e., we expect that

$$F_f(t) = F(t). \tag{5.3}$$

To be more precise, for $\Lambda' \supset \Lambda$:

$$H^\Phi_{\Lambda'} = H^\Phi_\Lambda + W^\Phi_{\Lambda,\Lambda'} + H^\Phi_{\Lambda'\setminus\Lambda}$$

where

$$W^\Phi_{\Lambda,\Lambda'} = \sum_{A\subseteq\Lambda', A\cap\Lambda\neq\varnothing, A\cap\Lambda^c\neq\varnothing} \Phi(A). \tag{5.4}$$

Remark that $\|W^\Phi_{\Lambda,\Lambda'}\| = O(|\partial\Lambda|)$, and since H^Φ_Λ and $H^\Phi_{\Lambda'\setminus\Lambda}$ commute,

$$\frac{\mathrm{Tr}_{\Lambda'}\left(\exp\left(t\sum_{i\in\Lambda} X_i\right)\exp\left(-\beta H^\Phi_\Lambda - \beta H^\Phi_{\Lambda'\setminus\Lambda}\right)\right)}{\mathrm{Tr}_{\Lambda'}\exp\left(-\beta H^\Phi_\Lambda - \beta H^\Phi_{\Lambda'\setminus\Lambda}\right)} = \frac{\mathrm{Tr}_\Lambda[e^{t\sum_{i\in\Lambda} X_i}e^{-\beta H^\Phi_\Lambda}]}{\mathrm{Tr}_\Lambda[e^{-\beta H^\Phi_\Lambda}]} \tag{5.5}$$

Hence, if we omit the boundary term $W^\Phi_{\Lambda,\Lambda'}$ in (5.2), then we recover $F_f(t)$. The main problem is to omit $W^\Phi_{\Lambda,\Lambda'}$ in the numerator of (5.2), and to prove that the "price" for this omission is of order $(e^{o(|\Lambda|)})$.

This reminds us on the proof of the existence of the pressure, see e.g. [7], [15]. However, in the quantum case this result relies on the inequality

$$|\log\mathrm{Tr}(e^{A+B}) - \log\mathrm{Tr}(e^A)| \leq \|B\|. \tag{5.6}$$

In order to prove (5.3) in a similar way, we would like to have an estimate like

$$|\log\omega(e^{A+B}) - \log\omega(e^A)| \leq \alpha\|B\|, \tag{5.7}$$

for a state ω, where α does not depend on A, B. But such an inequality does not hold!

More precisely, if $\omega = \mathrm{Tr}(e^H\cdot)/\mathrm{Tr}(e^H)$ (we omit for a moment the indices Λ referring to the volume), then

$$|\log\omega(e^{A+B}) - \log\omega(e^A)| = \int_0^1 \frac{d}{dt}\log\omega(e^{A+tB})dt$$

and

$$\frac{d}{dt}\log\omega(e^{A+tB}) = \frac{\mathrm{Tr}\left(\int_0^1 e^{A+tB}e^{-s(A+tB)}e^He^{s(A+tB)}B\ ds\right)}{\mathrm{Tr}\left(\int_0^1 e^{A+tB}e^{-s(A+tB)}e^He^{s(A+tB)}\ ds\right)} =: \Psi(B)$$

In general Ψ is not a state (unless $A+tB$ and H commute), and the norm of Ψ (as a functional of B) will depend on A, B and H, as the following proposition shows.

Proposition 5.8. *For any $X \in \mathcal{U}_\Lambda$ with* $\mathrm{Ker}(X) = \{0\}$*, define*

$$\Psi_X(B) = \frac{Tr(XB)}{Tr(X)} \tag{5.9}$$

Ψ_X defines a continuous functional of C with norm

$$\|\Psi_X\| = \frac{Tr|X|}{|Tr(X)|} \geq 1, \tag{5.10}$$

with $|X| = \sqrt{X^*X}$. In particular, for $X \geq 0$, $\|\Psi_X\| = 1$.

Proof: Put $X = J|X|$, where J is a partial isometry and $|X| = \sqrt{XX^*} \geq 0$. Since $\mathrm{Ker}(X) = \{0\}$, J is a unitary operator, see [8], Theorem 6.1.2. Since $|X| \geq 0$, $\omega_{|X|}(C) := \mathrm{Tr}(\cdot|X|)/\mathrm{Tr}(|X|)$ defines a state. We have:

$$\begin{aligned} \left|\frac{\mathrm{Tr}(CX)}{\mathrm{Tr}(X)}\right| = \left|\frac{\omega_{|X|}(CJ)}{\omega_{|X|}(J)}\right| &\leq \frac{\|C\|\|J\|}{|\omega_{|X|}(J)|} = \|C\| \left|\frac{\mathrm{Tr}[|X|]}{\mathrm{Tr}[J|X|]}\right| \\ &= \left|\frac{\mathrm{Tr}[|X|]}{\mathrm{Tr}(X)}\right| \|C\|, \end{aligned} \tag{5.11}$$

and we obtain

$$\|\Psi_X\| \leq \frac{\mathrm{Tr}(|X|)}{|\mathrm{Tr}(X)|}. \tag{5.12}$$

If we choose $C = J^*$, then

$$\frac{\mathrm{Tr}(CX)}{|\mathrm{Tr}X|} = \frac{\mathrm{Tr}(|X|)}{|\mathrm{Tr}(X)|}, \tag{5.13}$$

so

$$\|\Psi_X\| = \frac{\mathrm{Tr}(|X|)}{|\mathrm{Tr}(X)|}. \tag{5.14}$$

□

This proposition shows that we cannot hope to obtain a useful version of (5.7) in order to show (5.3). Indeed, if X is not positive (the X we are thinking about here is $\int_0^1 e^{A+tB} e^{-s(A+tB)} e^H e^{s(A+tB)}$), then $\|\Psi_X\|$ can be arbitrary large.

Instead we will use a cluster expansion to show the negligibility of the boundary terms.

6 Cluster Expansion

In this section we develop a strategy to prove both existence and analyticity of (5.1) and the equality (5.3), which is based on a quantum cluster expansion. For an introduction to this technique and a comparison of different approaches, see [14]. Here we develop a variant of this expansion, by rewriting the partition function of a quantum model as a partition function of a certain (classical) polymer model. Then, the results on the convergence of the expansion follow whenever the Kotecky-Preiss criterion is satisfied [9].

6.1 Set-up

Rewrite (5.1)

$$\begin{aligned} F_f(t) &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \frac{\mathrm{Tr}_\Lambda \left(e^{t\sum_{i\in\Lambda} X_i} e^{-\beta H_\Lambda^\Phi} \right)}{\mathrm{Tr}_\Lambda e^{t\sum_{i\in\Lambda} X_i}} + \log \mathrm{Tr}(e^{tX}) - P(\beta\Phi) \\ &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega_\Lambda^t(e^{-\beta H_\Lambda^\Phi}) + \log \mathrm{Tr}(e^{tX}) - P(\beta\Phi), \end{aligned} \tag{6.1}$$

where ω_Λ^t is a *product* state on $\mathcal{U}_\Lambda$ defined by

$$\omega_\Lambda^t(Y) = \frac{\mathrm{Tr}_\Lambda(e^{t\sum_{i\in\Lambda} X_i} Y)}{\mathrm{Tr}_\Lambda(e^{t\sum_{i\in\Lambda} X_i})}. \tag{6.2}$$

The product property of the state ω_X^t is crucial and due to the fact that we consider only the averages of a one-point observable. It implies for $A \in \mathcal{U}_{\Lambda'}$, $B \in \mathcal{U}_{\Lambda''}$, $\Lambda' \cap \Lambda'' = \varnothing$:

$$\omega_\Lambda^t(AB) = \omega_\Lambda^t(A)\, \omega_\Lambda^t(B). \tag{6.3}$$

This factorization is crucial to set up the cluster expansion that will allow us to show the existence of the limit

$$\Xi_f(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log Z_\Lambda^{t,\beta} = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega_\Lambda^t(e^{-\beta H_\Lambda^\Phi}) \tag{6.4}$$

Similarly, for $\Lambda' \supset \Lambda$ define $\omega_{\Lambda',\Lambda}^t$ by

$$\omega_{\Lambda',\Lambda}^t(Y) = \frac{\mathrm{Tr}_{\Lambda'}(e^{-t\sum_{i\in\Lambda} X_i} Y)}{\mathrm{Tr}_{\Lambda'}(e^{-t\sum_{i\in\Lambda} X_i} \otimes \mathrm{Id}_{\Lambda'\setminus\Lambda})} \tag{6.5}$$

which is also a product state, and this time we have

$$\begin{aligned} F(t) &= \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \lim_{\Lambda'\uparrow\mathbb{Z}^d} \log \frac{\mathrm{Tr}_{\Lambda'}\big(e^{t\sum_{i\in\Lambda} X_i}\, e^{-\beta H_{\Lambda'}^\Phi}\big)}{\mathrm{Tr}_{\Lambda'}\big(e^{t\sum_{i\in\Lambda} X_i} \otimes \mathrm{Id}_{\Lambda'\setminus\Lambda}\big)} \frac{\mathrm{Tr}_\Lambda\big(e^{t\sum_{i\in\Lambda} X_i}\big)}{\mathrm{Tr}_\Lambda\big(e^{-\beta H_\Lambda^\Phi}\big)} \frac{\mathrm{Tr}_{\Lambda'}\big(e^{-\beta H_\Lambda^\Phi} \otimes \mathrm{Id}_{\Lambda'\setminus\Lambda}\big)}{\mathrm{Tr}_{\Lambda'}\big(e^{-\beta H_{\Lambda'}^\Phi}\big)} \\ &= \log \mathrm{Tr}(e^{tX}) - P(\beta\Phi) + \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \lim_{\Lambda'\uparrow\mathbb{Z}^d} \log \frac{\omega_{\Lambda',\Lambda}^t(e^{-\beta H_{\Lambda'}^\Phi})\, \omega_\Lambda^0(e^{-\beta H_\Lambda^\Phi})}{\omega_{\Lambda'}^0(e^{-\beta H_{\Lambda'}^\Phi})} \end{aligned} \tag{6.6}$$

where $\omega_{\Lambda'}^0 = \omega_{\Lambda'}^{t=0}$ is the trace state on $\mathcal{U}_{\Lambda'}$. The existence of $F(t)$ is equivalent with the existence of

$$\Xi(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \lim_{\Lambda'\uparrow\mathbb{Z}^d} \log \tilde{Z}_{\Lambda',\Lambda}^{t,\beta} \tag{6.7}$$

where

$$\tilde{Z}_{\Lambda',\Lambda}^{t,\beta} = \frac{\omega_{\Lambda',\Lambda}^t(e^{-\beta H_{\Lambda'}^\Phi})\, \omega_\Lambda^0(e^{-\beta H_\Lambda^\Phi})}{\omega_{\Lambda'}^0(e^{-\beta H_{\Lambda'}^\Phi})} \tag{6.8}$$

Moreover, the equality $F_f(t) = F(t)$ will follow from $\Xi_f(t) = \Xi(t)$.

Our strategy is then described as follows.

1. Set up the cluster expansion in order to define $\Xi(t)$, $\Xi_f(t)$. This can be done by properly defining a polymer model and by using the Kotecky-Preiss criterion.

2. Equality of Ξ and Ξ_f follows from the fact that in the expansion, only clusters touching the boundary of Λ will make the difference between $\log Z_\Lambda^{t,\beta}$ and $\log \tilde{Z}_{\Lambda',\Lambda}^{t,\beta}$

3. Analyticity is proved by showing that the polymer weights are analytic in t and satisfy the Kotecky-Preiss criterion in a strip in the complex plane.

6.2 Polymer model

In order to compute $\Xi_f(t)$, we use the idea of the Mayer expansion and rewrite the finite volume expectation $\omega_\Lambda^t(e^{-\beta H_\Lambda^\Phi})$ as the partition function of a polymer gas. Due to the product structure of the state, the polymer weights become independent up to the exclusion, and we can use familiar results on the convergence of a series for the logarithm of such partition functions.

We start by writing the series

$$Z_\Lambda^{t,\beta} = \omega_\Lambda^t\Big[\sum_{n=0}^{\infty} \frac{(-\beta H_\Lambda)^n}{n!}\Big] = \omega_\Lambda^t\Big[\sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \sum_{A_1,\dots,A_n\subseteq\Lambda} \Phi(A_1)\dots\Phi(A_n)\Big] \tag{6.9}$$

that can be cast into the form of a polymer expansion as follows. We use the notation $\Gamma = (A_1,\dots,A_n)$ for any finite sequence of finite sets of sites and the shorthand $\Phi_\Gamma = \Phi(A_1)\dots\Phi(A_n)$. Let G_Γ be the graph over the set of vertices $\{1,\dots,n\}$ such that $1 \le i < j \le n$ are connected by edge whenever $A_i \cap A_j \neq \varnothing$. A sequence $\Gamma' = (A'_1,\dots,A'_k)$ is called a maximally connected subsequence of Γ whenever there is a maximally connected component of the graph G_Γ with the vertex set $i_1 < \dots < i_k$ such that $A'_1 = A_{i_1},\dots,A'_k = A_{i_k}$. The following lemma is then an immediate application of these definitions.

Lemma 6.10. *Let $\{\gamma_\alpha\}_{\alpha\in I}$ be the collection of all maximally connected subsequences of Γ. Then $\Phi_\Gamma = \prod_{\alpha\in I}\Phi_{\gamma_\alpha}$ and the product does not depend on the order.*

Connected sequences of sets are called polymers and we use the symbol $\mathcal{A}_\Lambda$ for the set of all polymers in Λ. Any sequence (respectively set) of polymers $(\gamma_1,\dots,\gamma_n)$, $\gamma_\alpha = (A^\alpha_{k_1},\dots,A^\alpha_{k_\alpha})$, $\alpha = 1,\dots,n$ is called an admissible sequence (respectively set) if $A_i^\alpha \cap A_j^{\alpha'} = \varnothing$ for any i,j and $\alpha \neq \alpha'$. Given any sequence of sets Γ, the collection $\{\gamma_\alpha\}_{\alpha\in I}$ of all maximally connected components of Γ is clearly an admissible set of polymers, but the correspondence is obviously not one-to-one. A simple observation is that there are exactly

$$\frac{(\sum_\alpha k(\alpha))!}{\prod_\alpha k(\alpha)!}$$

sequences Γ such that $\{\gamma_\alpha\}_{\alpha\in I}$ is the collection of all maximally connected subsequences of Γ. Defining $|\Gamma| = k$ for any sequence $\Gamma = (A_1,\dots,A_k)$, we rewrite $Z_\Lambda^{t,\beta}$ as

the partition function of a polymer model:

$$\begin{aligned} Z_\Lambda^{t,\beta} &= \sum_{n=0}^{\infty} \frac{(-\beta)^n}{n!} \sum_{\Gamma:\,|\Gamma|=n} \prod_{\alpha} \omega_\Lambda^t(\Phi_{\gamma_\alpha}) \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\substack{l_1,\dots,l_k\geq 1 \\ \sum_i l_i = n}} \sum_{\substack{\gamma_1,\dots,\gamma_k\in\mathcal{A}_\Lambda \\ |\gamma_1|=l_1,\dots,|\gamma_k|=l_k}} g(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{k} \Big[\frac{(-\beta)^{l_i}}{l_i!}\omega_\Lambda^t(\Phi_{\gamma_i})\Big] \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\gamma_1,\dots,\gamma_n\in\mathcal{A}_\Lambda} g(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{n} \rho^{t,\beta}(\gamma_i) \end{aligned} \tag{6.11}$$

where we have introduced the weights

$$\rho^{t,\beta}(A_1,\dots,A_k) = \frac{(-\beta)^k}{k!} g_C(A_1,\dots,A_k)\, \omega_\Lambda^t(\Phi(A_1)\dots\Phi(A_k)) \tag{6.12}$$

and the indicator functions

$$g(\gamma_1,\dots,\gamma_n) = \begin{cases} 1 & \text{if } (\gamma_1,\dots,\gamma_n) \text{ is admissible} \\ 0 & \text{otherwise} \end{cases} \tag{6.13}$$

$$g_C(A_1,\dots,A_k) = \begin{cases} 1 & \text{if } (A_1,\dots,A_k) \text{ is connected} \\ 0 & \text{otherwise} \end{cases} \tag{6.14}$$

Note that the polymers have been defined as sequences of sets rather than collections of sets. Obviously, the weight $\rho^{t,\beta}(\gamma = A_1,\dots,A_k)$ generically depends on the order of the sets $A_1,\dots,A_k$, whereas it does not depend on Λ as far as $\Lambda \supset \cup_{i=1}^k A_i$, due to the product structure of the state ω_Λ^t. The cluster expansion now reads [9, 12], formally,

$$\begin{aligned} \log Z_\Lambda^{t,\beta} &= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\gamma_1,\dots,\gamma_n\in\mathcal{A}_\Lambda} a_T(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{n} \rho^{t,\beta}(\gamma_i) \\ &= \sum_{C\subseteq\Lambda} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{\gamma_1,\dots,\gamma_n\in\mathcal{A}_C \\ \mathrm{Supp}(\gamma_1,\dots,\gamma_n)=C}} a_T(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{n} \rho^{t,\beta}(\gamma_i) \\ &= \sum_{C\subseteq\Lambda} w^{t,\beta}(C) \end{aligned} \tag{6.15}$$

where we have defined the "cluster" weights $w^{t,\beta}(C)$ by the partial resummation over all sequences $(\gamma_1,\dots,\gamma_n)$, $\gamma_\alpha = (A_1^\alpha,\dots,A_{k(\alpha)}^\alpha)$, $n \geq 1$, such that $\mathrm{Supp}(\gamma_1,\dots,\gamma_n) := \cup_{\alpha=1}^n \cup_{i=1}^{k(\alpha)} A_i^\alpha = C$.

Next we set up a similar expansion for $\tilde{Z}_{\Lambda',\Lambda}^{t,\beta}$. Note first that by taking $t = 0$ in (6.15) we immediately get cluster expansions for both expectations $\omega_\Lambda^0(e^{-\beta H_\Lambda^\Phi})$ and $\omega_{\Lambda'}^0(e^{-\beta H_{\Lambda'}^\Phi})$. For $\omega_{\Lambda',\Lambda}^t(e^{-\beta H_{\Lambda'}^\Phi})$ we repeat the same steps to get

$$\omega_{\Lambda',\Lambda}^t(e^{-\beta H_{\Lambda'}^\Phi}) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\gamma_1,\dots,\gamma_n\in\mathcal{A}_{\Lambda'}} g(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{n} \tilde{\rho}_\Lambda^{t,\beta}(\gamma_i) \tag{6.16}$$

with the notation

$$\tilde{\rho}^{t,\beta}_{\Lambda}(A_1,\dots,A_k) = \frac{(-\beta)^k}{k!} g_C(A_1,\dots,A_k)\, \omega^t_{\Lambda',\Lambda}(\Phi(A_1)\dots\Phi(A_k)) \tag{6.17}$$

In particular, $\tilde{\rho}^{t,\beta}_{\Lambda}(A_1,\dots,A_k)$ does not depend on Λ' if $\cup_i A_i \subseteq \Lambda'$, and

$$\tilde{\rho}^{t,\beta}_{\Lambda}(A_1,\dots,A_k) = \begin{cases} \rho^{t,\beta}(A_1,\dots,A_k) & \text{if } \cup_i A_i \subseteq \Lambda \\ \rho^{0,\beta}(A_1,\dots,A_k) & \text{if } \cup_i A_i \subseteq \Lambda' \setminus \Lambda \end{cases} \tag{6.18}$$

Defining

$$\tilde{w}^{t,\beta}_{\Lambda}(C) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{\gamma_1,\dots,\gamma_n \\ \mathrm{Supp}(\gamma_1,\dots,\gamma_n)=C}} a_T(\gamma_1,\dots,\gamma_n) \prod_{i=1}^{n} \tilde{\rho}^{t,\beta}_{\Lambda}(\gamma_i) \tag{6.19}$$

we also have $\tilde{w}^{t,\beta}_{\Lambda}(C) = w^{t,\beta}(C)$ for any $C \subseteq \Lambda$ and $\tilde{w}^{t,\beta}_{\Lambda}(C) = w^{0,\beta}(C)$ for any $C \subseteq \Lambda' \setminus \Lambda$. Hence, we get from (6.8) that

$$\begin{aligned} \log Z^{t,\beta}_{\Lambda',\Lambda} &= \sum_{C\subseteq\Lambda'} \tilde{w}^{t,\beta}_{\Lambda}(C) + \sum_{C\subseteq\Lambda} w^{0,\beta}(C) - \sum_{C\subseteq\Lambda'} w^{0,\beta}(C) \\ &= \sum_{C\subseteq\Lambda} w^{t,\beta}(C) + \sum_{\substack{C\subseteq\Lambda' \\ C\not\subseteq\Lambda,\, C\not\subseteq\Lambda'\setminus\Lambda}} \tilde{w}^{t,\beta}_{\Lambda}(C) \end{aligned} \tag{6.20}$$

where the first term coincides with the series for $\log Z^{t,\beta}_{\Lambda}$ and the second one is a boundary term summing up the clusters intersecting both sets Λ and $\Lambda' \setminus \Lambda$.

The existence of the cluster weights $w^{t,\beta}(C)$ and $w^{t,\beta}_{\Lambda}(C)$ and upper bounds can be proven under the assumption that the polymer weights are sufficiently damped. We show the Kotecky-Preiss criterion of the convergence [9, 12] to be satisfied in a high-temperature regime:

Lemma 6.21. *Let $a, \beta_0 > 0$ such that*

$$\sum_{B\ni 0} e^{2a|B|} \left(e^{\beta_0\|\Phi_B\|} - 1\right) \le a \tag{6.22}$$

Then one has the upper bounds

$$\sup_{(t,\beta)\in\mathbb{R}\times[0,\beta_0]} \sum_{n=1}^{\infty} \sum_{\substack{A_1,\dots,A_n \\ 0\in\cup_i A_i}} e^{a|\cup_i A_i|} |\rho^{t,\beta}(A_1,\dots,A_n)| \le a \tag{6.23}$$

and

$$\sup_{\Lambda} \sup_{(t,\beta)\in\mathbb{R}\times[0,\beta_0]} \sum_{n=1}^{\infty} \sum_{\substack{A_1,\dots,A_n \\ 0\in\cup_i A_i}} e^{a|\cup_i A_i|} |\tilde{\rho}^{t,\beta}_{\Lambda}(A_1,\dots,A_n)| \le a \tag{6.24}$$

Combining with Theorem 2 in [12], we have the following result.

Proposition 6.25. *Under Condition* (6.22)*, the cluster weights* $w^{t,\beta}(C)$ *are translation-invariant and*

$$\sup_{(t,\beta)\in\mathbb{R}\times[0,\beta_0]} \sum_{C\ni 0} |w^{t,\beta}(C)| \le a \tag{6.26}$$

Similarly,

$$\sup_{\Lambda} \sup_{(t,\beta)\in\mathbb{R}\times[0,\beta_0]} \sum_{C\ni 0} |\tilde{w}^{t,\beta}_{\Lambda}(C)| \le a \tag{6.27}$$

Proof of Lemma 6.21. Using the notation $\underline{\gamma}$ for the image of the sequence $\gamma = (A_1, \dots, A_k)$, we have the upper bound

$$\begin{aligned}
\sum_{\gamma:\,\underline{\gamma}=\{B_1,\dots,B_l\}} |\rho^{t,\beta}(\gamma)| &\le g_C(B_1,\dots,B_l) \sum_{n=1}^{\infty} \frac{\beta^n}{n!} \sum_{\substack{k_1,\dots,k_l\ge 1\\ k_1+\dots+k_l=n}} \binom{n}{k_1,\dots,k_l} \prod_{i=1}^{l} \|\Phi(B_i)\|^{k_i} \\
&= g_C(B_1,\dots,B_l) \prod_{i=1}^{l} \sum_{k=1}^{\infty} \frac{\beta^k}{k!} \|\Phi(B_i)\|^k \\
&= g_C(B_1,\dots,B_l) \prod_{i=1}^{l} (e^{\beta\|\Phi(B_i)\|} - 1)
\end{aligned} \tag{6.28}$$

for any collection of finite sets $\{B_1, \dots, B_l\}$, $l = 1, 2, \dots$ The identical upper bound holds if we replace $\rho^{t,\beta}$ with $\tilde{\rho}^{t,\beta}_{\Lambda}$. Hence, we only need to concentrate on the former weight, for which we have the inequality

$$\sum_{n=1}^{\infty} \sum_{\substack{A_1,\dots,A_n\\ 0\in\cup_i A_i}} e^{a|\cup_i A_i|} |\rho^{t,\beta}(A_1,\dots,A_n)| \le \sum_{l=1}^{\infty} \sum_{\substack{\{B_1,\dots,B_l\}\\ 0\in\cup_i B_i}} g_C(B_1,\dots,B_l) \prod_{i=1}^{l} \varphi(B_i) \tag{6.29}$$

with the shortcut

$$\varphi(B) = e^{a|B|} \left(e^{\beta\|\Phi(B)\|} - 1\right) \tag{6.30}$$

In order to estimate the right-hand side of (6.29), we consider the function

$$Y(\Delta \,|\, B) = \sum_{l\ge 1} \sum_{B\in\{B_1,\dots,B_l\}\subseteq\Delta} g_C(B_1,\dots,B_l) \prod_{i=1}^{l} \varphi(B_i) \tag{6.31}$$

$$= \varphi(B) \sum_{l\ge 0} \sum_{\{B_1,\dots,B_l\}\subseteq\Delta\setminus\{B\}} g_C(B, B_1,\dots,B_l) \prod_{i=1}^{l} \varphi(B_i) \tag{6.32}$$

defined on all pairs of a finite set B and a finite collection Δ of finite sets, $\Delta \ni B$. For this function we derive a recurrent inequality, realizing that the collection $\{B_1, \dots, B_l\}$ obtained from a connected collection $\{B, B_1, \dots, B_l\}$ splits in general into a number

of connected components, each of them intersecting B. By upper-bounding the sums over these components separately we subsequently write:

$$\begin{aligned} Y(\Delta \,|\, B) &\leq \varphi(B) \sum_{m\geq 0} \frac{1}{m!} \sum_{l_1,\dots,l_m\geq 1} \prod_{\alpha=1}^{m} \Big(\sum_{\substack{\{B_1,\dots,B_{l_\alpha}\}\subseteq \Delta\setminus\{B\} \\ B\cap \cup_i B_i \neq \varnothing}} g_C(B_1,\dots,B_{l_\alpha}) \prod_{i=1}^{l_\alpha} \varphi(B_i) \Big) \\ &\leq \varphi(B) \sum_{m\geq 0} \frac{1}{m!} \Big(\sum_{D\colon D\cap B\neq\varnothing} \varphi(D) \sum_{l\geq 0} \sum_{\{B_1,\dots,B_l\}\subseteq\Delta\setminus\{B,D\}} g_C(D,B_1,\dots,B_l) \Big)^m \\ &\leq \varphi(B) \exp\Big[|B| \sup_x \sum_{\substack{D\ni x \\ D\in\Delta\setminus\{B\}}} Y(\Delta\setminus\{B\} \,|\, D) \Big] \end{aligned} \tag{6.33}$$

Since

$$\sum_{n=1}^{\infty} \sum_{\substack{A_1,\dots,A_n \\ 0\in\cup_i A_i}} e^{a|\cup_i A_i|} |\rho^{t,\beta}(A_1,\dots,A_n)| \leq \sup_{\Delta} \sum_{\substack{B\ni 0 \\ B\in\Delta}} Y(\Delta \,|\, B) \tag{6.34}$$

the proof of the lemma is finished by proving by induction in the cardinality of Δ that

$$\sup_x \sum_{\substack{B\ni x \\ B\in\Delta}} Y(\Delta \,|\, B) \leq a \tag{6.35}$$

uniformly in Δ. Indeed, by the induction hypothesis and using inequalities (6.22), (6.33) and the translation-invariance of the potential, we get

$$\sup_x \sum_{\substack{B\ni x \\ B\in\Delta}} Y(\Delta \,|\, B) \leq \sup_x \sum_{B\ni x} \varphi(B)\, e^{a|B|} = \sum_{B\ni 0} e^{2a|B|}\, (e^{\beta\|\Phi(B)\|} - 1) \leq a \tag{6.36}$$

The case $\Delta = \{\varnothing\}$ is trivial. □

6.3 Analyticity of the polymer weights

Under a slightly stronger condition than (6.22), we prove the existence of an analytic continuation of the polymer weights $\rho^{z,\beta}(\gamma)$ to a set $\{z = x + iy \in \mathcal{C};\, |y| < \delta\} \times \{\beta;\, |\beta| < \beta_0\}$, uniformly for all polymers. Since the linear functional ω_Λ^z given by formula (6.2) is not a state anymore for $z \notin \mathbb{R}$ due to the lack of positivity, we write $\rho^{z,\beta}(\gamma)$ in the form

$$\begin{aligned} \rho^{z=x+iy,\beta}(A_1,\dots,A_k) &= \frac{(-\beta)^k}{k!} g_C(A_1,\dots,A_k)\, \omega_\Lambda^z(\Phi(A_1)\dots\Phi(A_k)) \\ &= \frac{(-\beta)^k}{k!} g_C(A_1,\dots,A_k) \\ &\quad \times \Big(\frac{\mathrm{Tr} e^{xX}}{\mathrm{Tr} e^{(x+iy)X}} \Big)^{|\Lambda|} \omega_\Lambda^x(\Phi(A_1)\dots\Phi(A_k)\, e^{i\delta\sum_{i\in\Lambda} X_i}), \end{aligned} \tag{6.37}$$

an identity valid for any $\Lambda \supset \cup_{i=1}^k A_i$. Choosing $\Lambda = \cup_{i=1}^k A_i$ and using the bound

$$\frac{|\mathrm{Tr}\, e^{(x+iy)X}|}{\mathrm{Tr}\, e^{xX}} = |\omega^x_{\{0\}}(e^{iyX})| = |\omega^x_{\{0\}}(\sum_{n=0}^{\infty} \frac{(iy)^n}{n!} X^n)| \tag{6.38}$$

$$\geq 1 - \sum_{n=1}^{\infty} \frac{|y|^n}{n!} \|X\|^n = 2 - e^{|y|\|X\|} \tag{6.39}$$

we obtain the next variant of inequality (6.28):

$$\sum_{\gamma:\, \underline{\gamma}=\{B_1,\ldots,B_l\}} |\rho^{x+iy,\beta}(\gamma)| \leq g_C(B_1,\ldots,B_l) \Big(\frac{1}{2-e^{|y|\|X\|}}\Big)^{|\cup_{i=1}^l B_l|} \prod_{i=1}^{l} (e^{|\beta|\|\Phi_{B_i}\|} - 1)$$

$$\leq g_C(B_1,\ldots,B_l) \prod_{i=1}^{l} \frac{e^{|\beta|\|\Phi_{B_i}\|} - 1}{(2-e^{|y|\|X\|})^{|B_i|}} \tag{6.40}$$

The other steps in the proof of Lemma 6.21 remain unchanged if we replace there the function $\varphi(B)$ with

$$\varphi'(B) = \Big(\frac{e^a}{2-e^{|y|\|X\|}}\Big)^{|B|} (e^{|\beta|\|\Phi(B)\|} - 1) \tag{6.41}$$

and assume condition (6.44) below. As a result, we get the upper bound on the analytic continuation of the polymer weights as

$$\sup_{|y|<\delta,\, |\beta|<\beta_0} \sum_{n=1}^{\infty} \sum_{\substack{A_1,\ldots,A_n \\ 0\in\cup_i A_i}} e^{a|\cup_i A_i|} |\rho^{z=x+iy,\beta}(A_1,\ldots,A_n)| \leq a \tag{6.42}$$

The uniform bounds on the cluster weights in the region $|y| < \delta_0$, $|\beta| < \beta_0$ then follow by [9, 12]. As a consequence, the (partially resummed cluster) weights $w^{z,\beta}(C)$ are analytic by Vitali's theorem and we arrive at the following result:

Proposition 6.43. *Assume there are $a, \delta, \beta_0 > 0$ such that*

$$\sum_{B\ni 0} \Big(\frac{e^{2a}}{2-e^{\delta\|X\|}}\Big)^{|B|} (e^{\beta_0\|\Phi_B\|} - 1) \leq a \tag{6.44}$$

Then the cluster weights $w^{z,\beta}(C)$ are analytic in the region

$$\mathcal{V}_{\delta,\beta_0} = \{(z,\beta) \in \mathcal{C}^2 :\, |Im\, z| < \delta,\, |\beta| < \beta_0\}$$

for all finite sets of sites C. Moreover,

$$\sup_{(z,\beta)\in\mathcal{V}_{\delta,\beta_0}} \sum_{C\ni 0} |w^{z,\beta}(C)| \leq a \tag{6.45}$$

7 Proof of Theorem 2.15

One easily checks that under the assumption

$$\sum_{B\ni 0} e^{\epsilon|B|}\,\|\Phi(B)\| < \infty \tag{7.1}$$

for some $\epsilon > 0$, there exist $a, \delta, \beta_0 > 0$ such that condition (6.44) is satisfied. Moreover, β_0 can be chosen independent of X.

Existence of $F_f(t)$.
It follows from the cluster expansion for $\log Z^{t,\beta}_\Lambda$ due to the translation-invariance of the cluster weights $w^{t,\beta}(C)$. To see this, write

$$\log Z^{t,\beta}_\Lambda - |\Lambda| \sum_{C\ni 0} \frac{w^{t,\beta}(C)}{|C|} = -\sum_{i\in\Lambda} \sum_{\substack{C\ni i\\ C\not\subseteq\Lambda}} \frac{w^{t,\beta}(C)}{|C|} \tag{7.2}$$

By Proposition 6.25, $\sum_{C\ni 0} |w^{t,\beta}(C)| \le a$, and for any $\epsilon > 0$ there exists a finite set of sites $\mathcal{D}$ such that $\sum_{C\ni 0,\, C\not\subseteq\mathcal{D}} |w^{t,\beta}(C)| \le \epsilon$. Introducing

$$\Lambda_0 = \{i \in \Lambda;\, \mathcal{D} + i \subseteq \Lambda\} \tag{7.3}$$

we have

$$\begin{aligned} \Big|\log Z^{t,\beta}_\Lambda - |\Lambda| \sum_{C\ni 0} \frac{w^{t,\beta}(C)}{|C|}\Big| &\le \Big(\sum_{i\in\Lambda_0} + \sum_{i\in\Lambda\setminus\Lambda_0}\Big) \sum_{\substack{C\ni i\\ C\not\subseteq\Lambda}} |w^{t,\beta}(C)| &(7.4)\\ &\le \epsilon|\Lambda_0| + a|\Lambda\setminus\Lambda_0| &(7.5)\end{aligned}$$

and the limit $\lim_{\epsilon\downarrow 0} \lim_{\Lambda\uparrow\mathbb{Z}^d}$, the latter being taken in the van Hove sense so that $\lim |\Lambda_0|/|\Lambda| = 1$, yields

$$\Xi_f(t) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log Z^{t,\beta}_\Lambda = \sum_{C\ni 0} \frac{w^{t,\beta}(C)}{|C|} \tag{7.6}$$

Moreover, we have got the upper bound $\sup_{t\in\mathbb{R}} |\Xi_f(t)| \le a$.

Equality $F_f(t) = f(t)$.
Notice first that the limit $\Lambda' \uparrow \mathbb{Z}^d$ exists,

$$\lim_{\Lambda'\uparrow\mathbb{Z}^d} \log Z^{t,\beta}_{\Lambda',\Lambda} = \log Z^{t,\beta}_\Lambda + \sum_{C\not\subseteq\Lambda,\, C\cap\Lambda\neq\varnothing} \tilde{w}^{t,\beta}_\Lambda(C) \tag{7.7}$$

by the absolute convergence of the second sum. Using the same argument as above, the second term is of order $o(|\Lambda|)$, and we get the equality $\Xi(t) = \Xi_f(t)$.

Analyticity of $F_f(t)$.
We only need to prove the analyticity of the function $\Xi_f(t)$ given by the series (7.6). By Proposition 6.43, all cluster weights have an analytic continuation to the strip

$|\mathrm{Im}\, z| < \delta_0$. Since the series converges there uniformly due to (6.45), $\Xi_f(z)$ is analytic in the strip by Vitali's theorem.

Finally, the existence and the differentiability of $F(t)$ implies both large deviation upper and lower bounds by Gartner-Ellis theorem. Since $F(t)$ has an analytic continuation to a neighborhood of the origin, Bryc's theorem implies the central limit theorem. To see that $\sigma^2 > 0$ for β small enough, consider first the case $\beta = 0$, then

$$\sigma^2 = \left(\frac{d^2}{dt^2} F(t)\right)_{t=0} = \omega_0(X^2) - \omega_0(X)^2$$

where ω_0 is the normalized trace. Hence in that case, $\sigma^2 > 0$ as soon as X has non-trivial spectrum. Therefore, by the convergence of the cluster expansion, the variance $\sigma_\beta^2 = \sigma_0^2 + O(\beta)$ is strictly positive for β small enough. Moreover, it is given by the absolutely converging sum

$$\sigma_\beta^2 = \sum_{i \in \mathbb{Z}^d} \omega\left((X_i - \omega(X_i))(X_0 - \omega(X_0))\right)$$

7.1 A generalization

Our result on the convergence of the cluster expansions for $\Xi_f(t)$ and $\Xi(t)$ can be slightly generalized. We sketch this generalization here without too much details. Let $\{\Phi_k\}_{k=1,\dots,n}$ be a family of potentials and ω be a product state. Then the generating function

$$Z_\Lambda^z = \omega(e^{z_1 H_\Lambda^{\Phi_1}} \dots e^{z_n H_\Lambda^{\Phi_n}}) \tag{7.8}$$

where $z = (z_1, \dots, z_n) \in \mathcal{C}^n$, admits a cluster expansion

$$\log Z_\Lambda^z = \sum_{C \subseteq \Lambda} w^z(C) \tag{7.9}$$

with the cluster weights $w^z(C)$ depending only on $\Phi_i(A), A \subseteq C, i = 1, \dots, n$, and one has the following result.

Proposition 7.10. *Assume that*

$$\sup_x \sum_{B \ni x} e^{2a|B|} (e^{\sum_{i=1}^n \delta_i \|\Phi_i(B)\|} - 1) \leq a \tag{7.11}$$

for some $a, \delta_1, \dots, \delta_n > 0$. Then all cluster weights $w^z(C)$ are analytic in the polydisc $\mathcal{D} = \{z = (z_1, \dots, z_n) : |z_i| \leq \delta_i,\ i = 1, \dots, n\}$ and

$$\sup_{z \in \mathcal{D}} \sup_x \sum_{C \ni x} |w^z(C)| \leq a \tag{7.12}$$

Remark 7.13. *1. Notice that* (7.11) *is the same condition one would write for the convergence of the cluster expansion for* $\log \omega(\exp \sum_i z_i H_\Lambda^{\Phi_i})$.

2. *As a corollary, one obtains the existence and analyticity in a neighborhood of the origin for various kinds of (cumulant) generating functions. In particular, by taking* $\Phi_1 = \Phi$ *translationally invariant,* $z_1 = \beta$, $z_2 = z$, *and* $\Phi_2(B) = \sum_{i\in\mathbb{Z}^d} \tau_i(X)\, 1_{B=D+i}$ *for a fixed set of sites* D *and an operator* $X \in \mathcal{U}_D$, *one gets the existence and analyticity for the (free b.c.) generating function*

$$F_f^X(z) = \lim_{\Lambda\uparrow\mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega_\Lambda^{\Phi,\beta}(e^{z\sum_{i:\, D+i\subseteq\Lambda} \tau_i(X)})$$

where

$$\omega_\Lambda^{\Phi,\beta}(Y) = \frac{Tr(e^{-\beta H_\Lambda^\Phi} Y)}{Tr(e^{-\beta H_\Lambda^\Phi})}$$

3. *A necessary and sufficient condition on the potentials* $\Phi_1, \ldots, \Phi_n$ *to satisfy* (7.11) *with some* $a, \delta_1, \ldots, \delta_n > 0$ *is that there exists* $\epsilon > 0$ *such that*

$$\sup_x \sum_{B\ni x} e^{\epsilon|B|} \|\Phi_i(B)\| < \infty \qquad i = 1, \ldots, n$$

Proposition 7.10 does not give the (full) large deviation principle since the modulus of z_i has to be small. However, it does give the central limit theorem for

$$\frac{H_\Lambda^\Phi - \omega(H_\Lambda^\Phi)}{\sqrt{|\Lambda|}}$$

because for that we only need analyticity in a neighborhood of 0.

8 Level two large deviations

We will now define a random measure which can thought of as the distribution under the state ω of the "measure" $\frac{1}{|\Lambda|}\sum_{i\in\Lambda} \delta_{X_i}$. For $f \in \mathcal{C}([-\|X\|, \|X\|], \mathbb{R})$, and μ a probability measure on $[-\|X\|, \|X\|]$, we write $\langle \mu, f\rangle = \int f d\mu$.

More precisely, for $f_1, \ldots, f_k$ a finite collection of continuous functions on $[-\|X\|, \|X\|]$, and $A_1, \ldots A_k$ Borel sets, define

$$\mathbb{P}(\langle \mathcal{L}_\Lambda, f_1\rangle \in A_1, \ldots, \langle \mathcal{L}_\Lambda, f_k\rangle \in A_k) = \omega\Big(1_{A_1}\Big(\frac{1}{|\Lambda|}\sum_{i\in\Lambda} f_1(X_i)\Big) \ldots 1_{A_k}\Big(\frac{1}{|\Lambda|}\sum_{i\in\Lambda} f_k(X_i)\Big)\Big) \tag{8.1}$$

This formula defines the distribution of a random measure $\mathcal{L}_\Lambda$. Indeed, the sets

$$\{\langle \mathcal{L}_\Lambda, f_1\rangle \in A_1, \ldots \langle \mathcal{L}_\Lambda, f_k\rangle \in A_k : f_i \in \mathcal{C}([-\|X\|, \|X\|], \mathbb{R}), A_i \in \mathcal{B}\} \tag{8.2}$$

are Borel sets in the weak topology on $\mathcal{M}_1([-\|X\|, \|X\|])$, the set of probability measures on $[-\|X\|, \|X\|]$, and they are generating for the Borel-σ-field on $\mathcal{M}_1([-\|X\|, \|X\|])$

The candidate level-2 generating function is then a functional on $\mathcal{C}([-\|X\|, \|X\|], \mathbb{R})$, given by

$$\Psi(f) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \log \mathbb{E}\left(e^{\langle \mathcal{L}_\Lambda, f\rangle}\right) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \log \omega\left(e^{\sum_{i\in\Lambda} f(X_i)}\right) \tag{8.3}$$

And the corresponding candidate large deviation entropy function is its Legendre transform:

$$\mathcal{I}(\mu) = \sup_{f \in \mathcal{C}([-\|X\|, \|X\|], \mathbb{R})} (\langle \mu, f\rangle - \Psi(f) \tag{8.4}$$

We then have the following theorem

Theorem 8.5. *Suppose ω is a high temperature KMS state as in Theorem 2.15.*

1. *The limit defining the generating function $\Psi(f)$ in (8.3) exists and defines a convex $\Psi : \mathcal{C}([-\|X\|, \|X\|], \mathbb{R}) \to \mathbb{R}$.*

2. *The random measures $\mathcal{L}_\Lambda$ satisfy the large deviation principle with rate function $\mathcal{I}$ given by (8.4).*

3. *The relation between $\mathcal{I}$ and I is given by the contraction principle:*

$$I(x) = \inf\{\mathcal{I}(\mu) : \int_{[-\|X\|, \|X\|]} \xi \mu(d\xi) = x\} \tag{8.6}$$

Proof. The existence of the limit defining Ψ follows from Theorem 2.15 and the fact that β_0 does not depend on X, so we can replace X_i by $f(X_i)$.

The large deviation principle follows from Gâteaux differentiability of Ψ. More precisely, for any $f, g \in \mathcal{C}([-\|X\|, \|X\|], \mathbb{R})$, the limit

$$\partial_g \Psi(f) = \lim_{t\to 0} \frac{\Psi(f + tg) - \Psi(f)}{t} \tag{8.7}$$

exists. This can be seen as follows. By the same argument showing the analyticity of $F(z)$ (of 2.16) in a strip $\{z = x + iy : |y| < \delta\}$ one sees that $z \mapsto \Psi(f + zg)$ exists and is analytic in a strip $\{z = x + iy : |y| < \delta\}$, where now δ depends on f and g. This is clearly sufficient to have the existence of the limit (8.7). Then we can apply Corollary 4.5.27 of [3] to conclude the large deviation principle.

Finally, the contraction principle follows from the fact that the distribution of $X_\Lambda/|\Lambda|$ is the distribution of $\int \xi \mathcal{L}_\Lambda(d\xi)$, hence we are in the situation of the standard contraction principle, [3], Theorem 4.2.1. □

Acknowledgements

We are very grateful to A. C. D. van Enter and C. Maes for fruitful discussions and many useful comments.

References

[1] W. Bryc, A remark on the connection between the large deviation principle and the central limit theorem, *Stat. and Prob. Lett.* **18**:253–256 (1993).

[2] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics 2* (Springer-Verlag, Berlin, 1996).

[3] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (Springer-Verlag, New York, 1998).

[4] A. C. D. van Enter, R. Fernández, and A. D. Sokal, Regularity properties and pathologies of position-space renormalization group transformations: Scope and limitations of Gibbsian theory, *J. Stat. Phys.* **72**:879–1167 (1993).

[5] H.-O. Georgii. *Gibbs Measures and Phase Transitions* (Walter de Gruyter & Co., Berlin, 1988).

[6] R. S. Ellis, *Large Deviations and Statistical Mechanics* (Springer-Verlag, New York, 1985).

[7] R. B. Israel, *Convexity in the Theory of Lattice Gases* (Princeton University Press, 1979).

[8] R. V. Kadison, J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras* (Academic Press, New York, London 1983).

[9] R. Kotecky and D. Preiss, Cluster expansion for abstract polymer models, *Commun. Math. Phys.* **103**:491–498 (1986).

[10] J. L. Lebowitz, M. Lenci, and H. Spohn, Large deviations for ideal quantum systems, (math-phys archive 9906014).

[11] G. Gallavotti, J. L. Lebowitz, and V. Mastropietro, Large deviations in rarefied quantum gases, *J. Stat. Phys.* **108**:831–861 (2002).

[12] S. Miracle-Solé, On the convergence of cluster expansions, *Physica A* **279**:244–249 (2000).

[13] S. Olla, Large deviations for Gibbs random fields, *Prob. Th. Rel. Fields* **77**:343–357 (1988).

[14] Y. M. Park, The cluster expansion for classical and quantum lattice systems, *J. Stat. Phys.* **27**:553–576 (1982).

[15] B. Simon, *The Statistical Mechanics of Lattice Gases* (Princeton University Press, 1993).

[16] D. Goderis and P. Vets, Central limit theorem for mixing quantum systems and the CCR-algebra of fluctuations, *Comm. Math. Phys.* **122**:249–265 (1989).

[17] D. Goderis, A. Verbeure, and P. Vets, Noncommutative central limits, *Prob. Th. Rel. Fields* **82**:527–544 (1989).