ON THE VARIATIONAL PRINCIPLE FOR GENERALIZED GIBBS MEASURES

AERNOUT VAN ENTER AND EVGENY VERBITSKIY

Abstract. We present a novel approach to establishing the variational principle for Gibbs and generalized (weak and almost) Gibbs states. Limitations of a thermodynamic formalism for generalized Gibbs states will be discussed. A new class of intuitively Gibbs measures is introduced, and a typical example is studied. Finally, we present a new example of a non-Gibbsian measure arising from an industrial application.

1. Introduction

Gibbs measures, defined as solutions of the Dobrushin-Lanford-Ruelle equations, can equivalently be defined as solutions of a variational principle (at least when they are translation invariant).

Such a variational principle states that when we take as a base measure a Gibbs measure for some potential, or more generally, for some specification, other Gibbs measures for the same potential (specification) are characterized by having a zero relative entropy density with respect to this base measure.

If the base measure is not a Gibbs measure, such a statement need not be true anymore. The construction of Xu [17] provides an example of a "universal" ergodic base measure, such that all translation invariant measures have zero entropy density with respect to it. Even within the reasonably well-behaved class of "almost Gibbs" measures there are examples such that certain Dirac measures have zero entropy density with respect to it, [4]; for a similar weakly Gibbsian example see [14]. One can, however, to some extent circumvent this problem by requiring that both measures share sufficiently many configurations in their support.

For almost Gibbs measures, the measure-one set of good (continuity) configurations have the property that they can shield off any influence from infinity. On the other hand for the strictly larger class of weakly Gibbsian measures, it may suffice that most, but not necessarily all, influences from infinity are blocked by the "good" configurations.

The situation with respect to the variational principle between the class of almost Gibbs measures is much better than with respect to the class of weak Gibbs measures [10].

On the one hand, one expects that a variational principle might hold beyond the class of almost Gibbs measures. For example, infinite-range unbounded-spin systems lack the almost Gibbs property (due to the fact that for a configuration of sufficiently increasing spins the interaction between the origin and infinity is never negligible, whatever happens in between), but a variational principle for such models has been found; on the other hand, the analysis of Külske for the random field Ising model implies that one really needs some extra conditions, or the variational principle can be violated.

The paper is organised as follows. After recalling some basic notions and definitions, we discuss Goldstein's construction of a specification for an arbitrary translation invariant measure. In Section 3, we consider two measures ν and μ such that $h(\nu|\mu) = 0$ and we formulate a sufficient condition for ν to be consistent with a given specification γ for μ. We also consider a general situation of $h(\nu|\mu) = 0$ and recover a result of Föllmer. The new sufficient condition is clarified in the case of almost Gibbs measures in Section 4, and for a particular weak Gibbs measure in Section 5. We also introduce a new class of intuitively weak Gibbs measures. In Section 6, we present an example of a non-Gibbsian measure arising in industrial setting: magnetic and optical data storage.

Acknowledgment. We are grateful to Frank den Hollander, Christof Külske, Jeff Steif, and Frank Redig for useful discussions.

2. Specifications and Gibbs measures

2.1. **Notation.** We work with spin systems on the lattice $\mathbb{Z}^d$, i.e., configurations are elements of the product space $\mathcal{A}^{\mathbb{Z}^d}$, where $\mathcal{A}$ is a finite set (alphabet). The configuration space $\Omega = \mathcal{A}^{\mathbb{Z}^d}$ is endowed with the product topology, making it into a compact metric space. Configurations will denoted by lower-case Greek letters. The set of finite subsets of $\mathbb{Z}^d$ is denoted by $\mathcal{S}$.

For $\Lambda \in \mathcal{S}$ we put $\Omega_\Lambda = \mathcal{A}^\Lambda$. For $\sigma \in \Omega$, and $\Lambda \in \mathcal{S}$, $\sigma_\Lambda \in \Omega_\Lambda$ denotes the restriction of σ to Λ. For σ, η in Ω, $\Lambda \in \mathcal{S}$, $\sigma_\Lambda\eta_{\Lambda^c}$ denotes the configuration coinciding with σ on Λ, and η on Λ^c. For $\Lambda \subseteq \mathbb{Z}^d$, $\mathcal{F}_\Lambda$ denote the σ-algbra generated by $\{\sigma_x |\, x \in \Lambda\}$.

For two translation invariant probability measures μ and ν, define

$$H_\Lambda(\nu|\mu) = H(\nu_\Lambda|\mu_\Lambda) = \sum_{\sigma_\Lambda} \nu(\sigma_\Lambda) \log \frac{\nu(\sigma_\Lambda)}{\mu(\sigma_\Lambda)},$$

if ν_Λ is absolutely continuous with respect to μ_Λ, and $H_\Lambda(\nu|\mu) = +\infty$, otherwise.

The relative entropy density $h(\nu|\mu)$ is defined (provided the limit exists) as

$$h(\nu|\mu) = \lim_{n\to\infty} \frac{1}{|\Lambda_n|} H_{\Lambda_n}(\nu|\mu),$$

where $\{\Lambda_n\}$ is a sequence of finite subsets $\mathbb{Z}^d$, with $\Lambda_n \nearrow \mathbb{Z}^d$ as $n \to \infty$ in van Hove sense. For example, one can take $\Lambda_n = [-n, n]^d$.

A potential $U = \{U(\Lambda, \cdot)\}_{\Lambda\in\mathcal{S}}$ is a family of functions indexed by finite subsets of $\mathbb{Z}^d$ with the property that $U(A, \omega)$ depends only on ω_Λ. A Hamiltonian H^U_Λ is defined by

$$H^U_\Lambda(\sigma) = \sum_{\Lambda'\cap\Lambda\neq\varnothing} U(\Lambda', \sigma).$$

The Hamiltonian H^U_Λ is said to be convergent in σ if the sum on the right hand side is convergent. The Gibbs specification $\gamma^U = \{\gamma^U_\Lambda\}_{\Lambda\in\mathcal{S}}$ is defined by

$$\gamma^U_\Lambda(\omega_\Lambda|\sigma_{\Lambda^c}) = \frac{\exp\Big(-H^U_\Lambda(\omega_\Lambda\sigma_{\Lambda^c})\Big)}{\sum\limits_{\tilde\omega_\Lambda\in\Omega_\Lambda} \exp\Big(-H^U_\Lambda(\tilde\omega_\Lambda\sigma_{\Lambda^c})\Big)}$$

provided H^U_Λ is convergent in every point $\tilde\omega_\Lambda\sigma_{\Lambda^c}$, $\tilde\omega_\Lambda \in \Omega_\Lambda$. Formally, if H^U_Λ is not convergent in every point $\omega \in \Omega$, γ^U is not a specification in the sense of standard Definition 2.1. Nevertheless, in many cases (e.g., weakly Gibbsian measures, see Definition 5.1 below) γ^U can still be viewed as a version of conditional probabilities for μ:

$$\mu(\omega_\Lambda|\sigma_{\Lambda^c}) = \gamma^U_\Lambda(\omega_\Lambda|\sigma_{\Lambda^c}) \quad (\mu\text{-a.s.}).$$

2.2. **Specifications.**

Definition 2.1. *A family of probability kernels $\gamma = \{\gamma_\Lambda\}_{\Lambda\in\mathcal{S}}$ is called a specification if*

a) $\gamma_\Lambda(F|\cdot)$ *is $\mathcal{F}_{\Lambda^c}$-measurable for all $\Lambda \in \mathcal{S}$ and $F \in \mathcal{F}$;*
b) $\gamma_\Lambda(F|\omega) = \mathbb{I}_F(\omega)$ *for all $\Lambda \in \mathcal{S}$ and $F \in \mathcal{F}_{\Lambda^c}$;*
c) $\gamma_{\Lambda'}\gamma_\Lambda = \gamma_{\Lambda'}$ *whenever $\Lambda \subseteq \Lambda'$, and where*

$$\big(\gamma_{\Lambda'}\gamma_\Lambda\big)(F|\omega) = \int \gamma_\Lambda(F|\eta)\gamma_{\Lambda'}(d\eta|\omega).$$

Definition 2.2. *A probability measure μ is called consistent with a specification γ (denoted by $\mu \in \mathcal{G}(\gamma)$) if for every bounded measurable* *function f one has*

$$\int f\,d\mu = \int \gamma_\Lambda(f)\,d\mu. \tag{2.1}$$

If μ is consistent with the specification γ, then γ can be viewed as a version of the conditional probabilities of μ, since (2.1) implies that for any finite Λ

$$\gamma_\Lambda(A|\omega) = \mathbb{E}_\mu(\mathbb{I}_A|\mathcal{F}_{\Lambda^c})(\omega) \quad (\mu\text{-a.s.}),$$

where $\mathcal{F}_{\Lambda^c}$ is the σ-algebra generated by spins outside Λ.

Definition 2.1 requires that for every ω, $\gamma_\Lambda(\cdot|\omega)$ is a probability measure on $\mathcal{F}$, and that the consistency condition (c) is satisfied for **all** $\omega \in \Omega$. In fact, when dealing with the weakly Gibbs measures, these requirements are too strong. Definition 2.1 can be generalized [15, p. 16], and this form is probably more suitable for the weakly Gibbsian formalism.

2.3. **Construction of specifications.** In [7], Goldstein showed that every measure has a specification. In other words, for every measure μ there exists a specification γ in the sense of definition 2.1, such that $\mu \in \mathcal{G}(\gamma)$. Let us briefly recall Goldstein's construction.

Suppose μ is a probability measure on $\Omega = A^{\mathbb{Z}^d}$, and let $\{\Lambda_n\}$, $\Lambda_n \in \mathcal{S}$, be an increasing sequence such that $\cup_n \Lambda_n = \mathbb{Z}^d$. For a finite set $\Lambda \in \mathbb{Z}^d$, and arbitrary $\eta_\Lambda \in A^\Lambda$, $\omega \in A^{\mathbb{Z}^d\setminus\Lambda}$, define

$$\mu(\eta_\Lambda|\omega_{\Lambda^c}) := \mu([\eta_\Lambda]|\mathcal{F}_{\Lambda^c})(\omega),$$

where $[\eta_\Lambda] = \{\zeta \in \Omega : \zeta|_\Lambda = \eta_\Lambda\}$. By the martingale convergence theorem

$$\mu([\eta_\Lambda]|\mathcal{F}_{\Lambda^c})(\omega) = \lim_{n\to\infty} \mu(\eta_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \quad \text{for } \mu - a.e.\ \omega. \tag{2.2}$$

The sequence on the right hand side of (2.2) is defined by elementary conditional probabilities:

$$\mu(\eta_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) = \frac{\mu(\eta_\Lambda\omega_{\Lambda_n\setminus\Lambda})}{\mu(\omega_{\Lambda_n\setminus\Lambda})}$$

Define

$$G_\Lambda = \{\omega : \text{the limit on RHS of (2.2) exists for all } \eta_\Lambda \in \Omega_\Lambda\}, \tag{2.3}$$

and denote this limit by $p_\Lambda(\eta|\omega)$. For $\Lambda \subseteq \Lambda_n$, let

$$Q_\Lambda^{\Lambda_n} = \{\omega \in G_{\Lambda_n} : \sum_{\eta_\Lambda\in\Omega^\Lambda} p_{\Lambda_n}(\eta_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\omega) > 0\}.$$

Finally, let

$$\mathcal{H}_\Lambda = \bigcup_n \bigcap_{j=n}^{\infty} Q_\Lambda^{\Lambda_j},$$

and define γ_Λ by

$$\gamma_\Lambda(\eta|\omega) = \begin{cases} p_\Lambda(\eta|\omega), & \text{if } \omega \in \mathcal{H}_\Lambda \\ (|\Lambda||\mathcal{A}|)^{-1}, & \text{if } \omega \in \mathcal{H}_\Lambda^c \end{cases}. \tag{2.4}$$

Theorem 2.3. *The family $\gamma = \{\gamma_\Lambda\}_{\Lambda\in\mathcal{S}}$ given by (2.4) is a specification, and $\mu \in \mathcal{G}(\gamma)$.*

Suppose γ is a specification, and μ is cosnsistent with γ. Therefore

$$\mu(\cdot|\omega_{\Lambda^c}) = \gamma_\Lambda(\cdot|\omega_{\Lambda^c}) \quad \mu - a.s.$$

Taking (2.2) into account we conclude that

$$\gamma_\Lambda(\eta_\Lambda|\omega_{\Lambda^c}) = \lim_{n\to\infty} \mu(\eta_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \tag{2.5}$$

for all η_Λ and μ-almost all ω. An important problem for establishing the variational principle for generalized Gibbs measures, is determining the set of configurations where the convergence in (2.5) takes place.

3. Properly supported measures

Theorem 3.1. *Let μ be a measure consistent with the specification γ. Suppose that ν is another measure such that $h(\nu|\mu) = 0$. If for ν-almost all ω*

$$\mu(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \to \gamma(\xi_\Lambda|\omega_{\Lambda^c}), \tag{3.1}$$

then ν is consistent with the specification γ, i.e., $\nu \in \mathcal{G}(\gamma)$.

Remark 3.1. Note that by the dominated convergence theorem, convergence in (3.1) is also in $L_1(\nu)$.

Remark 3.2. If μ is an almost Gibbs measure for the specification γ, and ν is a measure concentrating on the points of continuity of γ, i.e., $\nu(\Omega_\gamma) = 1$, then (3.1) holds, see the proof below.

Proof of Theorem 3.1. Suppose that $h(\nu|\mu) = 0$. Then [6, Theorem 15.37] for any $\varepsilon > 0$ and any finite set Λ, and every cube C such that $\Lambda \subseteq C$, there exists Δ, $C \subseteq \Delta$ such that

$$\mu(|f_\Delta - f_{\Delta\setminus\Lambda}|) < \varepsilon, \tag{3.2}$$

where for any finite set V, f_V is the density of $\nu|_V$ with respect to $\mu|_V$:

$$f_V(\omega_V) = \frac{\nu(\omega_V)}{\mu(\omega_V)}.$$

Rewrite (3.2) as follows:

$$\begin{aligned}
\mu(|f_\Delta - f_{\Delta\setminus\Lambda}|) &= \sum_{\eta_\Lambda,\omega_{\Delta\setminus\Lambda}} \mu(\eta_\Lambda\omega_{\Delta\setminus\Lambda})\Big|\frac{\nu(\eta_\Lambda\omega_{\Delta\setminus\Lambda})}{\mu(\eta_\Lambda\omega_{\Delta\setminus\Lambda})} - \frac{\nu(\omega_{\Delta\setminus\Lambda})}{\mu(\omega_{\Delta\setminus\Lambda})}\Big| \\
&= \sum_{\omega_{\Delta\setminus\Lambda}} \nu(\omega_{\Delta\setminus\Lambda})\Big\{\sum_{\eta_\Lambda}\Big|\frac{\nu(\eta_\Lambda\omega_{\Delta\setminus\Lambda})}{\nu(\omega_{\Delta\setminus\Lambda})} - \frac{\mu(\eta_\Lambda\omega_{\Delta\setminus\Lambda})}{\mu(\omega_{\Delta\setminus\Lambda})}\Big|\Big\} \\
&= \sum_{\omega_{\Delta\setminus\Lambda}} \nu(\omega_{\Delta\setminus\Lambda})||\nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda})||_{TV} \\
&= \mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda})||_{TV},
\end{aligned}$$

where $||\cdot||_{TV}$ is the total variation norm.

A measure ν is consistent with the specification γ if

$$\nu_\Lambda(\cdot|\omega_{\Lambda^c}) = \gamma_\Lambda(\cdot|\omega_{\Lambda^c}), \quad \nu - a.e.,$$

or, equivalently,

$$\mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Lambda^c}) - \gamma_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV} = 0.$$

Obviously one has

$$\begin{aligned}
\mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Lambda^c}) &- \gamma_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV} \\
&\le \mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Lambda^c}) - \nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda})||_{TV} \\
&\quad + \mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda})||_{TV} \\
&\quad + \mathbb{E}_\nu||\mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \gamma_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV}.
\end{aligned}$$

By the martingale convergence theorem, the first term

$$\mathbb{E}_\nu||\nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \nu_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV} \to 0, \quad \text{as } \Delta \nearrow \mathbb{Z}^d.$$

The second term tends to 0 due to (3.2), and the third term tends to zero because of our assumptions. □

3.1. **Weakly Gibbs measures which violate the Variational Principle.** Disordered systems studied extensively by Külske [8,9] provide a counterexample to the variational principle for weakly Gibbs measures. In [10], Külske, Le Ny and Redig showed that there exist two weakly Gibbs measures μ^+, μ^- such that

$$h(\mu^+|\mu^-) = h(\mu^-|\mu^+) = 0, \tag{3.3}$$

but μ^+ is not consistent with a (weakly) Gibbsian specification γ^- for μ^-, and vice versa. The novelty and beauty of their Random Field Ising Model example lies in the fact that *both* measures are non-trivial and the relation in (3.3) is symmetric. As already mentioned above, previous almost Gibbs ([4]) and weakly Gibbs [14] example violating the variational principle satisfied $h(\delta|\mu) = 0$ with δ being a Dirac measure.

What happens in the situation when $h(\nu|\mu) = 0$? Suppose ν is consistent with the specification $\tilde{\gamma}$. Then

$$\begin{aligned}\mathbb{E}_\nu ||\mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \tilde{\gamma}_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV} \le & \mathbb{E}_\nu ||\mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda})||_{TV} \\ & + \mathbb{E}_\nu ||\nu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \tilde{\gamma}_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV}.\end{aligned}$$

Again, the first term on the right hand side tends to zero because $h(\nu|\mu)$ =0, and the second term tends to zero, because of the martingale convergence theorem. Therefore

$$||\mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) - \tilde{\gamma}_\Lambda(\cdot|\omega_{\Lambda^c})||_{TV} \to 0,$$

in $L_1(\nu)$, and since the total variation between any two measures is always bounded by 2, we have that

$$\mu_\Lambda(\cdot|\omega_{\Delta\setminus\Lambda}) \to \tilde{\gamma}_\Lambda(\cdot|\omega_{\Lambda^c}), \quad \nu - a.s.$$

Since $\tilde{\gamma}$ is a specification for ν, and since the "infinite" conditional probability for μ are defined as the limits of finite conditional probabilities (provided the limits exist) we conclude that

$$\mu_\Lambda(\cdot|\omega_{\Lambda^c}) = \nu_\Lambda(\cdot|\omega_{\Lambda^c}), \quad \nu - a.s. \tag{3.4}$$

In fact (3.4) was obtained in a different way earlier by Föllmer in [5, Theorem 3.8]. It means that $h(\nu|\mu) = 0$ implies that the conditional probabilities of μ coincide with the conditional probabilities of ν for ν-almost all ω. However, as the counterexample of [10] shows, this result is not suitable for establishing the variational principle for the weakly Gibbsian measures, because the conditional probabilities can converge to a "wrong" specification. Condition (3.1) is instrumental in ensuring that this does not happen.

4. Almost Gibbs Measures

In this section we show that the condition (3.1) holds for all measures ν which are properly supported on the set of continuity points of almost Gibbs specifications.

Definition 4.1. *A specification γ is continuous in ω, if for all $\Lambda \in \mathcal{S}$*

$$\sup_{\sigma,\eta} \left| \gamma_\Lambda(\sigma_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) - \gamma_\Lambda(\sigma_\Lambda|\omega_{\Lambda^c}) \right| \to 0, \ \textit{as } n \to \infty.$$

Denote by Ω_γ the set of all continuity points of γ.

Definition 4.2. *A measure μ is called almost Gibbs, if μ is consistent with a specification γ and $\mu(\Omega_\gamma) = 1$.*

Remark 4.1. Note that we define almost Gibbs measures by requiring only that the specification is continuous almost everywhere. We do not require (as it is usually done, see e.g. [12]) that the corresponding specification is *uniformly non-null*, in other words satsfies a finite energy condition: for any $\Lambda \in \mathcal{S}$, there exist $a_\Lambda, b_\Lambda \in (0,1)$ such that

$$a_\Lambda \le \inf_{\xi_\Lambda,\omega_{\Lambda^c}} \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}) \le \sup_{\xi_\Lambda,\omega_{\Lambda^c}} \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}) \le b_\Lambda.$$

Theorem 4.3. *If μ is an almost Gibbs measure for specification γ, and*

$$\nu(\Omega_\gamma) = 1,$$

then

$$\mu(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \to \gamma(\xi_\Lambda|\omega_{\Lambda^c}) \tag{4.1}$$

for all ξ_Λ and ν-almost all ω.

Proof. Since $\mu \in \mathcal{G}(\gamma)$, μ satisfies the DLR equations for γ, and hence

$$\mu(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}) = \int \gamma_{\Lambda_n}(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta).$$

Similarly

(4.2)

$$\mu(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) = \frac{\mu(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda})}{\sum_{\tilde\xi_\Lambda}\mu(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda})} = \frac{\int \gamma_{\Lambda_n}(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)}{\sum_{\tilde\xi_\Lambda}\int \gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)}.$$

Since γ is a specification, for all ξ, ω, η one has

$$\frac{\gamma_{\Lambda_n}(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})}{\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})} = \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}). \tag{4.3}$$

Let

$$r_n(\omega) = \sup_{\xi_\Lambda,\eta_{\Lambda_n\setminus\Lambda}} |\gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) - \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})|.$$

Therefore, using (4.3), we obtain the following estimate

$$\gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}) - r_n(\omega) \le \mu(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \le \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}) + r_n(\omega).$$

Since $r_n(\omega) \to 0$ for $\omega \in \Omega_\gamma$, we also obtain that for those ω

$$\mu(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}) \to \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}).$$

□

Remark 4.2. In the case μ is a standard Gibbs measure and γ is the corresponding specification, one has $\Omega_\gamma = \Omega$ and hence by repeating the proof of Theorem 4.3 one obtains (3.1) for **all** measures ν.

5. Regular Points of Weakly Gibbs Measures

As we have stressed above the crucial task consists in determining the regular (in the sense of (3.1)) points for μ. In this section we address this problem in the case of weak Gibbs measures. A weakly Gibbsian measure μ is a measure for which one can find a potential convergent on a set of μ-measure 1, but not everywhere convergent.

Definition 5.1. *Let μ be a probability measure on $(\Omega, \mathcal{F})$, and $U = \{U(\Lambda, \cdot)\}$ is an interaction. Then the measure μ is said to be* **weakly Gibbs** *for an interaction U if μ is consistent with γ^U ($\mu \in \mathcal{G}(\gamma^U)$) and*

$$\mu(\Omega_U) = \mu(\{\omega : H^U_\Lambda(\omega) \text{ is convergent } \forall \Lambda \in \mathcal{S}\}) = 1.$$

Is it natural to expect that the set of points Ω_U where the potential is convergent, coincides with the set of points regular in the sense of (3.1))? The definition 5.1 is rather weak. It is not even clear whether in the case of weak Gibbs measures the following convergence holds: for any finite Λ, any ξ_Λ, and μ-almost all ω, η

$$\gamma^U_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) \longrightarrow \gamma^U_\Lambda(\xi_\Lambda|\omega_{\Lambda^c}) \text{ as } \Lambda_n \uparrow \mathbb{Z}^d. \tag{5.1}$$

Note that (5.1) is a natural generalization of a characteristic property of almost Gibbs measures (see definition 4.1).

We suspect that (5.1) does not hold for all weakly Gibbs measures. However, the counterexample should be rather pathalogical. Most of the weakly Gibbs measures known in the literature should be (are) intuitively weak Gibbs as well.

We introduce a class of measures which satisfy (5.1).

Definition 5.2. *A measure μ is called* **intuitively weakly Gibbs** *for an interaction U if μ is weakly Gibbs for U, and there exists a a set $\Omega_U^{reg} \subseteq \Omega_U$ with $\mu(\Omega_U^{reg}) = 1$ and such that*

$$\gamma^U_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) \longrightarrow \gamma^U_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})$$

as $\Lambda_n \uparrow \mathbb{Z}^d$, for all ω, $\eta \in \Omega_U^{reg}$.

This definition of "intuitively" weakly Gibbs measures is new. However, it is very natural, and in fact, this is how the weakly Gibbs measures have been viewed before by one of us, c.f. [2]: "*... The fact that the constraints which act as points of discontinuity often involve configurations which are very untypical for the measure under consideration, suggested a notion of* almost *Gibbsian or* weakly *Gibbsian measures. These are measures whose conditional probabilities are either continuous only on a set of full measure or can be written in terms of an interaction which is summable only on a set of full measure.* Intuitively*, the difference is that in one case the "good" configurations can* *shield off* all *influences from infinitely far away, and in the other case only* almost all *influences.*

The difference between the Gibbs, almost Gibbs, and intuitively weak Gibbs measures is that

$$\gamma_\Lambda^U(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) \longrightarrow \gamma_\Lambda^U(\xi_\Lambda|\omega_{\Lambda^c}) \text{ as } \Lambda_n \uparrow \mathbb{Z}^d,$$

holds

- for **all** ω and **all** η (Gibbs measures);
- for **μ-almost all** ω and **all** η (almost Gibbs measures);
- for **μ-almost all** ω and **μ-almost all** η (intuitively weak Gibbs measures);

A natural question is whether there exists an intuitively weak Gibbs which is not almost Gibbs. An answer is given by the following result.

Theorem 5.3. *Denote by G, AG, WG and IWG the classes of Gibbs, almost Gibbs, weakly Gibbs, and intuitively weakly Gibbs states, respectively. Then*

$$G \subsetneq AG \subsetneq IWG \subseteq WG.$$

Proof of Theorem 5.3. The inclusion $G \subsetneq AG \subsetneq WG$ was first established in [12]. The inclusion $IWG \subseteq WG$ is obvious. In fact, the result of [12] implies $AG \subseteq IWG$ as well.

Let us now show that $AG \neq IWG$. In [12], an example has been provided of a weakly Gibbs measure, which is not almost Gibbs. This example is constructed as follows. Let $\Omega = \{0,1\}^{\mathbb{Z}_+}$ and μ is absolutely continuous with respect to the Bernoulli measure $\nu = B(1/2, 1/2)$ with the density f

$$f(\omega) = \exp(-H^U(\omega)),$$

where H^U is a Hamiltonian for the interaction U, which is absolutely convergent ν-almost everywhere. The interaction is defined as follows. Fix $\rho < 1$ and define

$$U([0,2n],\omega) = \omega_0\omega_{2n}\rho^{n-N_{2n}(\omega)}\mathbb{I}_{\{N_{2n}(\omega)\leq n\}}, \tag{5.2}$$

where

$$N_{2n}(\omega) = \max\{j \geq 1 : \ \omega_{2n}\omega_{2n-1}\dots\omega_{2n-j+1} = 1\}$$

if $\omega_{2n} = 1$ and $N_{2n} = 0$ if $\omega_{2n} = 0$. Moreover, $U(A,\omega) = 0$ if $A \neq [0,2n]$.

It is easy to see that $H^U(\omega) = \sum_{n\geq 0} U([0,2n],\omega)$ is convergent for ν-a.a. ω. However, $H^U(\omega)$ is sufficiently divergent, so that the conditional probabilities

$$\mu(\omega_0 = 1|\omega_1 \ldots \omega_n \ldots) = \frac{\exp(-H^U(1\omega_1\omega_2\ldots))}{1+\exp(-H^U(1\omega_1\omega_2\ldots))},$$
$$\mu(\omega_0 = 0|\omega_1 \ldots \omega_n \ldots) = \frac{1}{1+\exp(-H^U(1\omega_1\omega_2\ldots))}$$

are not continuous μ-almost everywhere. Therefore, μ is not almost Gibbs.

Nevertheless, the exists a set $\Omega' \subseteq \{0,1\}^{\mathbb{Z}_+}$ such that $\mu(\Omega') = 1$ and for every $\omega, \xi \in \Omega'$

$$H^U(1\omega_{0^c}) = H^U(1\omega_{[1,2n]}\omega_{[0,2n]^c}), \quad H^U(1\omega_{[1,2n]}\xi_{[0,2n]^c}) < \infty,$$

and

$$H^U(1\omega_{[1,2n]}\xi_{[0,2n]^c}) \to H^U(1\omega_{[1,2n]}\omega_{[0,2n]^c}), \quad n \to \infty.$$

Define

$$B_k = \{\eta \in \{0,1\}^{\mathbb{Z}_+} : \eta_{2k}\eta_{2k-1}\ldots\eta_{[3k/2]} = 1\},$$
$$B = \bigcap_{K\in\mathbb{N}} \bigcup_{k\geq K} B_k.$$

Clearly,

$$\nu(B_k) \leq 2^{-k/2}, \quad \text{and} \sum_k \nu(B_k) < \infty.$$

Hence, by the Borel-Cantelli lemma $\nu(B) = 0$ and since $\mu \ll \nu$, $\mu(B) = 0$.

Moreover, for every $\omega \in B^c$, $H^U(\omega) < \infty$. Let $\Omega' = B^c$ and consider arbitrary $\omega, \xi \in \Omega'$. Then

$$\begin{aligned}
&|H^U(1\omega_{[1,2n]}\xi_{[0,2n]^c}) - H^U(1\omega_{[1,\infty)})| \\
&= \Big|\sum_{p\geq 0} U([0,2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}) - U([0,2p], 1\omega_{[1,\infty)})\Big| \\
&\leq \sum_{p\geq n+1} \Big|U([0,2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}) - U([0,2p], 1\omega_{[1,\infty)})\Big| \qquad (5.3)\\
&\leq \sum_{p\geq n+1} U([0,2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}) + \sum_{p\geq n+1} U([0,2p], 1\omega_{[1,\infty)})
\end{aligned}$$

The sum $\sum_{p\geq n+1} U([0,2p], 1\omega_{[1,\infty)})$ converges to zero as $n \to \infty$ since it is a remainder of a convergent series for $H^U(1\omega_{[1,\infty)})$. To complete the proof we have to show that

$$S := \sum_{p\geq n+1} U([0,2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}) \qquad (5.4)$$

converges to 0 as well.

For every $\eta \in \Omega' = B^c$ there exists $K = K(\eta)$ such that

$$\eta_{2k}\eta_{2k-1}\dots\eta_{[3k/2]} = 0$$

for all $k \geq K$. Let $K_1 = K(\omega), K_2 = K(\xi)$ and $K = \max(K_1, K_2)$. Suppose $n > K$. Write the sum for S in (5.4) as $S_1 + S_2$, where

$$S_1 = \sum_{p=n+1}^{[4n/3]+1}, \quad S_2 = \sum_{p=[4n/3]+2}^{\infty}.$$

Let us estimate the second sum first. Since $p > n > \max(K_1, K_2) \geq K(\xi)$, we have that

$$\xi_{2p}\xi_{2p-1}\dots\xi_{[3p/2]} = 0.$$

Moreover, since $p \geq [4n/3] + 2 \geq 4n/3 + 1$, one has $[3p/2] > 2n$ and therefore $U([0, 2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c})$ does not depend on $\omega_{[1,2n]}$. Hence

$$S_2 = \sum_{p=[4n/3]+2}^{\infty} U([0, 2p], 1\xi_{[1,\infty]}) \to 0, \text{ as } n \to \infty.$$

Let us now consider the first sum

$$S_1 = \sum_{p=n+1}^{[4n/3]+1} U([0, 2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}).$$

Terms in S_1, in principle, depend on $\omega_{[1,2n]}$. For this one has to have that

$$\xi_{2p}\xi_{2p-1}\dots\xi_{2n+1} = 1,$$

and few of the last bits in $\omega_{[1,2n]}$ are also equal to 1. Suppose $\omega_t = \dots = \omega_{2n} = 1$. Note, however, that since $n > \max(K_1, K_2) \geq K(\omega)$, necessarily $t > [3n/2]$. Therefore,

$$U([0, 2p], 1\omega_{[1,2n]}\xi_{[0,2n]^c}) \leq \rho^{p-(2p-t+1)} \leq \rho^{3n/2-p-2} \leq \rho^{3n/2-4n/3-3} = \rho^{n/6-3},$$

and since $\rho < 1$

$$S_1 \leq n\rho^{n/6-3} \to 0.$$

□

Another example of an intuitively weakly Gibbs measure, which is not almost Gibbs, is the finite absolutely continuous invariant measure of the Manneville–Pomeau map [14]. The reason is that for every ω

$$\gamma_\Lambda(\omega_0 = 1|\omega_{[1,n]}\mathbf{0}_{[n+1,\infty)}) = 0,$$

where $\mathbf{0}$ is a configuration made entirely from zeros. Thus the configurations finishing with an infinite number of zeros, are the bad configurations, causing the discontinuities in conditional probabilities. (One can also show that there are no other such configurations.) Since there is at most a countable number of bad configurations, this set has a μ-measure equal to 0.

Yet another example of a measure which should be IWG is the restriction to a layer of an Ising Gibbs measure.

The reason for this is the following. The example considered above in Theorem 5.3, the absolutely continuous invariant measure for the Manneville-Pomeau map, and the restriction of an Ising model to a layer, have a very similar property in common. Namely, for every "good" configuration ω, there is a finite number $c = c(\omega)$ such that $|U(A,\omega)|$ starts to decay exponentially fast in $\text{diam}(A)$ as soon as $\text{diam}(A) > c(\omega)$. In the example above, $c(\omega) = K(\omega)$. In fact, the "good" configurations are characterized by the property that $c(\omega) < \infty$. This random variable $c(\omega)$ was called a *correlation length*. The main difficulty is in estimating the correlation length $c(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda^n})$ for the "glued" configuration $\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}$ in terms of correlations lengths $c(\xi_\Lambda\omega_{\Lambda^c})$ and $c(\xi_\Lambda\eta_{\Lambda^c})$. Estimates obtained in [13] should provide enough information to deal with this problem in case of the restriction of the Ising model to a layer.

Let us proceed further with the study of regular points of an intuitively weak Gibbs measure μ. We follow the proof of Theorem 4.3. Firstly, one has

$$\frac{\gamma_{\Lambda_n}(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})}{\sum_{\tilde{\xi}_\Lambda}\gamma_{\Lambda_n}(\tilde{\xi}_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})} = \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}). \tag{5.5}$$

Let

$$r_n^\omega(\eta) = \sup_{\xi_\Lambda}\big|\gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda_n\setminus\Lambda}\eta_{\Lambda_n^c}) - \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})\big|.$$

Since γ is a specification, $|r_n| \leq 2$. Moreover, since μ is intuitively weak Gibbs, then for $\omega \in \Omega_U^{reg}$, $r_n^\omega(\eta) \to 0$ for μ-almost all η. Fix $\varepsilon > 0$ and let

$$A_{\varepsilon,n}^\omega = \big\{\eta : |r_n^\omega(\eta)| > \varepsilon\big\}.$$

Then

$$\mu(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}) = \gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})\mu(\omega_{\Lambda_n\setminus\Lambda}))+\int_\Omega r_n^\omega(\eta)\sum_{\tilde{\xi}_\Lambda}\gamma_{\Lambda_n}(\tilde{\xi}_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta),$$

and we continue

$$\left|\frac{\mu(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda})}{\mu(\omega_{\Lambda_n\setminus\Lambda})}-\gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})\right|$$

$$(5.6)\qquad =\left|\frac{(\int_{\Omega\setminus A^\omega_{\varepsilon,n}}+\int_{A^\omega_{\varepsilon,n}})r_n^\omega(\eta)\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)}{\int\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)}\right|$$
$$\le\varepsilon+2\frac{\int_{A^\omega_{\varepsilon,n}}\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)}{\mu(\omega_{\Lambda_n\setminus\Lambda})},$$

where we used that $|r_n^\omega|$ is always bounded by 2.

Let us estimate the remaining integral

$$\int_{A^\omega_{\varepsilon,n}}\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)=\int_\Omega\mathbb{I}_{A^\omega_{\varepsilon,n}}(\eta)\big(\gamma_{\Lambda_n}\mathbb{I}_{\omega_{\Lambda_n\setminus\Lambda}}\big)(\eta)\mu(d\eta),$$

where $\mathbb{I}_{\omega_{\Lambda_n\setminus\Lambda}}$ is the indicator of the cylinder set $\{\zeta:\zeta_{\Lambda_n\setminus\Lambda}=\omega_{\Lambda_n\setminus\Lambda}\}$. The set $A^\omega_{\varepsilon,n}$ is $\mathcal{F}_{\Lambda_n^c}$-measurable, therefore

$$\mathbb{I}_{A_{\varepsilon,n}}\big(\gamma_{\Lambda_n}\mathbb{I}_{\omega_{\Lambda_n\setminus\Lambda}}\big)=\gamma_{\Lambda_n}\big(\mathbb{I}_{A_{\varepsilon,n}}\mathbb{I}_{\omega_{\Lambda_n\setminus\Lambda}}\big)$$

and since μ satisfies the DLR equations with γ, we obtain that

$$\int_{A_{\varepsilon,n}}\sum_{\tilde\xi_\Lambda}\gamma_{\Lambda_n}(\tilde\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda}|\eta_{\Lambda_n^c})\mu(d\eta)=\int\mathbb{I}_{A^\omega_{\varepsilon,n}}(\eta)\mathbb{I}_{\omega_{\Lambda_n\setminus\Lambda}}(\eta)\mu(d\eta)$$
$$=\mu(A^\omega_{\varepsilon,n}\cap\omega_{\Lambda_n\setminus\Lambda}).$$

Therefore, we obtain the following estimate

$$\left|\frac{\mu(\xi_\Lambda\omega_{\Lambda_n\setminus\Lambda})}{\mu(\omega_{\Lambda_n\setminus\Lambda})}-\gamma_\Lambda(\xi_\Lambda|\omega_{\Lambda^c})\right|\le\varepsilon+2\frac{\mu(\omega_{\Lambda_n\setminus\Lambda}\cap A^\omega_{\varepsilon,n})}{\mu(\omega_{\Lambda_n\setminus\Lambda})}.$$

Now, if we can show that for all $\omega\in\Omega_U^{reg}$

$$\frac{\mu(\omega_{\Lambda_n\setminus\Lambda}\cap A^\omega_{\varepsilon,n})}{\mu(\omega_{\Lambda_n\setminus\Lambda})}\to 0,\ \text{ as } \Lambda_n\uparrow\mathbb{Z}^d,$$

we will be able to conclude that all points Ω_U^{reg} are regular in the sense of (3.1).

Let us now turn to the example of an intuitively weak Gibbs measure considered in Theorem 5.3. As usual for the Gibbs formalism, we check the required property only for $\Lambda=\{0\}$. We also let $\Lambda_n=[0,n]$, hence $\Lambda_n\setminus\Lambda=[1,n]$.

Let $\omega\in\Omega_U^{reg}$. Hence the potential is convergent in $\mathbf{1_0}\omega_{[1,\infty)}$, and therefore $\exp(-H(\mathbf{1_0}\omega_{[1,\infty)}))>0$. Choose arbitrary $\varepsilon>0$ such that

$$\varepsilon<\frac{1}{4}\exp(-H(\mathbf{1_0}\omega_{[1,\infty)})).$$

The measure μ is absolutely continuous with respect to the Bernoulli measure $\nu = B(1/2, 1/2)$. First of all, let us show that

$$\frac{\nu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\nu(\omega_{[1,n]})} \to 0, \text{ as } n \to \infty, \tag{5.7}$$

implies

$$\frac{\mu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\mu(\omega_{[1,n]})} \to 0, \text{ as } n \to \infty, \tag{5.8}$$

Since $H(\zeta)$ is non-negative (possibly infinite) for any ζ, one has

$$\begin{aligned}\mu(\omega_{\Lambda_n\setminus\Lambda} \cap A^{\omega}_{\varepsilon,n}) &= \int\limits_{\omega_{\Lambda_n\setminus\Lambda}\cap A^{\omega}_{\varepsilon,n}} \exp(-H(\zeta))\nu(d\zeta) \\ &\le \int\limits_{\omega_{\Lambda_n\setminus\Lambda}\cap A^{\omega}_{\varepsilon,n}} \nu(d\zeta) = \nu(\omega_{\Lambda_n\setminus\Lambda} \cap A^{\omega}_{\varepsilon,n}).\end{aligned}$$

Consider the set $W = \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c$. For every $\eta \in W$ we have

$$\sup_{\xi_0} \Big| \gamma_0(\xi_0 | \omega_{[1,n]}\eta_{[n+1,\infty)}) - \gamma_0(\xi_0|\omega_{[1,\infty)}) \Big| \le \varepsilon.$$

In particular

$$\left| \frac{\exp\big(-H(\mathbf{1}_0\omega_{[1,n]}\eta_{[n+1,\infty)})\big)}{\exp\big(-H(\mathbf{1}_0\omega_{[1,n]}\eta_{[n+1,\infty)})\big)+1} - \frac{\exp\big(-H(\mathbf{1}_0\omega_{[1,\infty)})\big)}{\exp\big(-H(\mathbf{1}_0\omega_{[1,\infty)})\big)+1} \right| \le \varepsilon,$$

where we have used the fact that $H(\mathbf{0}_0\zeta_{[1,\infty)}) = 0$ for all ζ. Note also that since ω, $\eta \in \Omega^{reg}_U$, both $H(\mathbf{1}_0\omega_{[1,\infty)})$, $H(\mathbf{1}_0\omega_{[1,n]}\eta_{[n+1,\infty)})$ are non-negative and finite. Therefore

$$\Big| \exp\big(-H(\mathbf{1}_0\omega_{[1,n]}\eta_{[n+1,\infty)})\big) - \exp\big(-H(\mathbf{1}_0\omega_{[1,\infty)})\big) \Big| \le 4\varepsilon.$$

Hence,

$$\begin{aligned}
\mu(\omega_{[1,n]}) &\geq \mu(\omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c) \\
&= \mu(\mathbf{1}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c) + \mu(\mathbf{0}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c) \\
&= \int\limits_{\mathbf{1}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c} \exp(-H(\zeta))\, \nu(d\zeta) \\
&\qquad + \int\limits_{\mathbf{0}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c} \exp(-H(\zeta))\, \nu(d\zeta) \\
&\geq \Big(\exp(-H(\mathbf{1}_0\omega_{[1,\infty)}) - 4\varepsilon\Big)\nu(\mathbf{1}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c) \\
&\qquad + \nu(\mathbf{0}_0 \cap \omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c) \\
&\geq C\nu(\omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c),
\end{aligned}$$

where $C = \exp(-H(\mathbf{1}_0\omega_{[1,\infty)}) - 4\varepsilon > 0$ (note that $C < 1$). Therefore,

$$\begin{aligned}
\frac{\mu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\mu(\omega_{[1,n]})} &\leq C^{-1}\frac{\nu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\nu(\omega_{[1,n]} \cap (A^{\omega}_{\varepsilon,n})^c)} \\
&= C^{-1}\frac{\nu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\nu(\omega_{[1,n]})}\frac{1}{1 - \frac{\nu(\omega_{[1,n]} \cap A^{\omega}_{\varepsilon,n})}{\nu(\omega_{[1,n]})}},
\end{aligned}$$

and hence (5.7) indeed implies (5.8).

Let us now proceed with the proof of (5.7). Since ν is a symmetric Bernoulli measure, $\nu(\omega_{[1,n]}) = 2^{-n}$.

If $x, y \geq 0$ then

$$\left|\frac{e^{-x}}{1+e^{-x}} - \frac{e^{-y}}{1+e^{-y}}\right| = \left|\frac{1}{1+e^{-x}} - \frac{1}{1+e^{-y}}\right| \leq |x-y|.$$

Therefore, if $\eta \in A^{\omega}_{\varepsilon,n}$, i.e.,

$$\sup_{\xi_0}\Big|\gamma_0(\xi_0|\omega_{[1,n]}\eta_{[n+1,\infty)}) - \gamma_0(\xi_0|\omega_{[1,\infty)})\Big| > \varepsilon,$$

then

$$\Big|H(\mathbf{1}_0\omega_{[1,n]}\eta_{[n+1,\infty)}) - H(\mathbf{1}_0\omega_{[1,\infty)})\Big| > \varepsilon. \tag{5.9}$$

Hence, if we define $B^{\omega}_{\varepsilon,n}$ as a set of points η such that (5.9) holds, we get that $A^{\omega}_{\varepsilon,n} \subseteq B^{\omega}_{\varepsilon,n}$.

To estimate the measure of $B^{\omega}_{\varepsilon,n}$ we have to use the estimates from the proof of Theorem 5.3. Without loss of generality we may assume that n is even, $n = 2n'$. Let us recall the estimate (5.3)

$$
\begin{aligned}
&|H(\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) - H(\mathbf{1}_0\omega_{[1,\infty)})| \\
&\quad \le \sum_{p\ge n+1} U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) + \sum_{p\ge n'+1} U([0,2p],\mathbf{1}_0\omega_{[1,\infty)})
\end{aligned}
$$

The second sum on the right hand side does not depend on η, and converges to 0 as $n' \to \infty$. Therefore, by choosing n' large enough we must have that if $\eta \in B^{\omega}_{\varepsilon,n}$ then

$$\sum_{p\ge n+1} U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) > \frac{\varepsilon}{2}.$$

Let us define a sequence $\delta_p = \rho^{0.1p}$, $p \ge 1$. Since $\rho \in (0,1)$, for sufficiently large n' one has

$$\sum_{p\ge n'+1} \delta_p < \frac{\varepsilon}{2}.$$

Consider the following events,

$$C^{\omega}_p = \Big\{\eta : U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) > \delta_p\Big\}.$$

Obviously,

$$B^{\omega}_{n,\varepsilon} \subseteq \bigcup_{p\ge n'+1} C^{\omega}_p.$$

In general, for arbitrary ζ, $U([0,2p],\zeta) > \delta_p$ if (see (5.2)) $\zeta_0 = \zeta_{2p} = 1$, $N_{2p}(\zeta) \le p$ and and $\rho^{p-N_{2p}(\zeta)} > \rho^{0.1p}$. Therefore,

$$0.9p \le N_{2p}(\zeta) \le p,$$

and hence

$$\nu(\zeta : U([0,2p],\zeta) > \delta_p) \le \sum_{k=[0.9p]}^{p} 2^{-k} \le 2^{-0.9p+2} =: z_p.$$

Let us continue with estimating the probability of C^{ω}_p. If $p > 2n'$, then C^{ω}_p does not depend on ω, and hence using the previous estimate

$$\nu(\omega_{[1,2n']} \cap C^{\omega}_p) \le 2^{-2n'} z_p.$$

For the small values of p, $p \in [n'+1, 2n']$, we have to proceed differently. For such p's the configuration η can "profit" from the last bits (equal to 1) in ω. Since ω is a regular configuration (see Theorem 5.3), for sufficiently large n', $\omega \in G_{n'}$, where

$$G_{n'} = \{\zeta : \zeta_{2k}\ldots\zeta_{[3k/2]} = 0 \quad \forall k \ge n'\}.$$

In particular, it means that at most $n'/2+1$ of the last bits in $\omega_{[1,2n']}$ are equal to 1, and in the worst case, $\omega_{2n'}\dots\omega_{[3n'/2]+1}=1$. From now on we assume that $\omega_{2n'}\dots\omega_{[3n'/2]+1}=1$.

We split the set of "bad" η's as follows:

$$\begin{aligned}C_p^\omega &= \{\eta : U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) > \delta_p\}\\ &= \{\eta : U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) > \delta_p \,\&\, \eta_{2n'+1}\dots\eta_{2p}=0\}\cup\\ &\qquad \{\eta : U([0,2p],\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}) > \delta_p \,\&\, \eta_{2n'+1}\dots\eta_{2p}=1\}\\ &= C_p^{\omega,0}\cup C_p^{\omega,1}.\end{aligned}$$

Again, the set $C_p^{\omega,0}$ does not depend on ω. In fact, $C_p^{\omega,0}$ is not empty only for p's close to $2n'$: on one hand, $p-N_{2p}<0.1p$ and on the other, $N_{2p}\le 2p-2n'$. Together with the fact that $p\in(n',2n']$, this is possible only for $p'\in\left(\frac{2}{1.1}n',2n'\right]$. For any p in this interval, one would need more than $0.9p$ ones, hence making a ν-measure of $C_p^{\omega,0}$ sufficiently small:

$$\nu(C_p^{\omega,0})\le 2^{-0.9p}.$$

Finally, the elements of $C_p^{\omega,1}$ are precisely the configurations which can profit from the fact that the last few bits in $\omega_{[1,2n']}$ are equal to 1. For such η's, in a "glued" configuration $\zeta=\mathbf{1}_0\omega_{[1,2n']}\eta_{[2n'+1,\infty)}$ a continuous interval of 1's is located starting from position $[3n'/2]+1$ and finishing at position $2p$. In order to have a positive contribution from $U([0,2p],\zeta)$ a long run of 1's should not be too long. Namely,

$$N_{2p}(\zeta)=2p-\left[\frac{3n'}{2}\right]\le p,$$

implying that $p\le[3n'/2]$, and hence, $C_p^{\omega,1}$ is empty for $p>[3n'/2]$. For, $p\in[n'+1,[3n'/2]]$ one has

$$U([0,2p],\zeta)=\rho^{p-N_{2p}(\zeta)}=\rho^{[3n'/2]-p}.$$

Hence if $U([0,2p],\zeta)>\rho^{0.1p}$, then $[3n'/2]-p<0.1p$ and hence $p>[3n'/2]/1.1$. Once again that means that $C_p^{\omega,1}$ is empty for $p\in[n'+1,[3n'/2]/1.1-1]$.

Therefore, for $p\in[n'+1,2n']$ we conclude that

$$C_p^{\omega,1}\subseteq\{\eta:\eta_{2n'+1}=\ldots=\eta_{2p}=1\}\quad\text{if}\quad \frac{1}{1.1}\left[\frac{3n'}{2}\right]<p\le\left[\frac{3n'}{2}\right],$$

and $C_p^{\omega,1}=\varnothing$, otherwise. In any case,

$$\nu(C_p^{\omega,1})\le 2^{-2p+2n'}.$$

We obtained that

$$\nu(\omega_{[1,2n']} \cap A^{\omega}_{\varepsilon,n}) \le \nu(\omega_{[1,2n']} \cap B^{\omega}_{\varepsilon,n}) \le \nu\Big(\omega_{[1,2n']} \cap \cup_{p \ge n'+1} C^{\omega}_p\Big)$$
$$\le \sum_{p \ge n'+1} \nu(\omega_{[1,2n']} \cap C^{\omega}_p) = S_1 + S_2 + S_3 + S_4,$$

where S_1, S_2, S_3, S_4 are sums over integer p's in intervals $I_1 = [n' + 1, [3n'/2]/1.1)$, $I_2 = [[3n'/2]/1.1, [3n'/2]]$, $I_3 = [[3n'/2] + 1, 2n']$, and $I_4 = [2n' + 1, \infty)$, respectively. We have the following estimates

$$S_1 = \sum_{p \in I_1} \nu(\omega_{[1,2n']} \cap C^{\omega}_p) = \sum_{p \in I_1} \nu(\omega_{[1,2n']} \cap C^{\omega,0}_p)$$
$$= \sum_{p \in I_1} \nu(\omega_{[1,2n']})\nu(C^{\omega,0}_p) \le 2^{-2n'} \sum_{p \in I_1} 2^{-0.9p}$$
$$\le 2^{-2n'} \frac{2^{-0.9n'}}{1 - 2^{-0.9}} \le 3 \cdot 2^{-2.9n'};$$
$$S_2 = \sum_{p \in I_2} \nu(\omega_{[1,2n']} \cap C^{\omega,0}_p) + \sum_{p \in I_2} \nu(\omega_{[1,2n']} \cap C^{\omega,1}_p)$$
$$\le 2^{-2n'} \sum_{p \in I_2} 2^{-0.9p} + 2^{-2n'} \sum_{p \in I_2} 2^{-2p+2n'}$$
$$\le 2^{-2n'} \cdot 12 \cdot 2^{-\frac{0.9}{1.1} \cdot \frac{3n'}{2}} + 12 \cdot 2^{-\frac{2}{1.1} \cdot \frac{3n'}{2}}$$
$$\le 12 \cdot 2^{-3n'} + 12 \cdot 2^{-2.7n'} \le 12 \cdot 2^{-2.7n'};$$
$$S_3 = \sum_{p \in I_3} \nu(\omega_{[1,2n']} \cap C^{\omega}_p) = \sum_{p \in I_3} \nu(\omega_{[1,2n']} \cap C^{\omega,0}_p)$$
$$= 2^{-2n'} \sum_{p \in I_3} 2^{-0.9p} \le 2^{-2n'} \cdot 12 \cdot 2^{-\frac{0.9 \cdot 3n'}{2}} \le 12 \cdot 2^{-3n'};$$
$$S_4 = \sum_{p \in I_4} \nu(\omega_{[1,2n']} \cap C^{\omega}_p) \le 12 \cdot 2^{-3.8n'}.$$

Finally, we conclude that

$$\frac{\nu(\omega_{[1,2n']} \cap A^{\omega}_{\varepsilon,n})}{\nu(\omega_{[1,2n']})} \le \frac{S_1 + S_2 + S_3 + S_4}{2^{-2n'}} \to 0 \text{ as } n' \to \infty.$$

To summarize our result, we formulate the following theorem.

Theorem 5.4. *Let μ be the (intuitively) weak Gibbs measure, but not almost Gibbs, discussed above in Theorem 5.3, and which has been introduced in* [12]. *Then there exists a set Ω' such that $\mu(\Omega') = 1$ and the following holds:*

- *the potential U is absolutely convergent on Ω';*

- *for all $\omega, \eta \in \Omega'$, any finite Λ and all $\xi_\Lambda \in \Omega_\Lambda$ one has*
$$H_\Lambda(\xi_\Lambda \omega_{\Lambda_n \setminus \Lambda} \eta_{\Lambda_n^c}) \to H_\Lambda(\xi_\Lambda \omega_{\Lambda^c}),$$
$$\gamma_\Lambda(\xi_\Lambda | \omega_{\Lambda_n \setminus \Lambda} \eta_{\Lambda_n^c}) \to \gamma_\Lambda(\xi_\Lambda | \omega_{\Lambda^c}),$$
as $\Lambda_n \to \mathbb{Z}_+$.
- *every $\omega \in \Omega'$ is regular in (Goldstein's) sense: for every ξ_Λ*
$$\mu(\xi_\Lambda | \omega_{\Lambda_n \setminus \Lambda}) \to \gamma_\Lambda^U(\xi_\Lambda | \omega_{\Lambda^c}),$$
as $\Lambda_n \to \mathbb{Z}_+$.

6. Bit-shift channel

A somewhat different kind of non-Gibbsian example comes from an industrial application: data storage on magnetic tape or optical disks (CD, DVD, etc). Before formulating the model precisely, let us explain the mechanism which leads to a non-Gibbsian measure.

The medium for magnetic or optical data storage can be in one of the two states: "high" and "low", or "bright" and "dark". The information is encoded not in the state of the medium itself, but in transitions between these states, and more precisely, in "units of time" between two successive transitions. In the following table the first line indicates the state of the medium (H(igh) or L(ow)) and the second line indicates the corresponding occurrence (1) or absence (0) of transitions:

... L H H H H L L L H H H H L ...
... 1 0 0 0 0 1 0 1 0 0 0 1 ...

An equivalent way to represent the second line is to record the number of zeros between consecutive ones. In the case above, one obtains a sequence $(\ldots, 3, 2, 3, \ldots)$. For technical reasons, in data storage one often uses coding schemes such that the transitions are never too close, but also not too far away from each other. This is achieved by using the so-called *run-length constrained codes*.

When the magnetic medium or optical disk are read, due to various effects like noise, intersymbol interference or clock jittering, the transitions can be erroneously identified, thus producing a time-shift in the detected positions.

Suppose in the example above the following error has occurred: the second transition has been detected one time unit too late. The resulting sequence then is $(\ldots, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, \ldots)$. And the corresponding representation in terms of runs of zeros will be $(\ldots, 4, 1, 3, \ldots)$ instead of $(\ldots, 3, 2, 3, \ldots)$.

The following description of a bit shift channel is due to Shamai and Zehavi, [16].

Let $\mathcal{A} = \{d, \ldots, k\}$, where $d, k \in \mathbb{N}$, $d < k$ and $d \geq 2$. Define $X = \mathcal{A}^{\mathbb{Z}} = \{x = (x_i) : x_i \in \mathcal{A}\}$, $\Omega = \{-1, 0, 1\}^{\mathbb{Z}} = \{\omega = (\omega_i) : \omega_i \in \{-1, 0, 1\}\}$. Consider the following transformation φ defined on $X \times \Omega$ as follows: $y = \varphi(x, \omega)$ with

$$y_i = x_i + \omega_i - \omega_{i-1} \quad \text{for all} \quad i \in \mathbb{Z}.$$

Note that y is a sequence such that $y_i \in \{0, \ldots, k+2\}$ for all i, but not every sequence in $\{0, \ldots, k+2\}^{\mathbb{Z}}$ can be obtained as an image of some $x \in X$, $\omega \in \Omega$. For example, all image sequences $y = \varphi(x, d)$ cannot contain 00. Indeed, suppose $y_i = 0$ for some i. This is possible if and only if $x_i = 2$, $\omega_i = -1$, and $\omega_{i-1} = 1$. But then $y_{i+1} = x_{i+1} + \omega_{i+1} - \omega_i \geq 2 - 1 + 1 = 2$.

Since φ is a continuous (in the product topology) transformation the set $Y = \varphi(X \times \Omega)$ is a so-called sofic shift, see [11].

Suppose μ and π are product Bernoulli measure on X and Ω with

$$\mu(j) = p_j, \ j = d, \ldots, k, \ \pi(-1) = \pi(1) = \epsilon, \ \pi(0) = 1 - 2\epsilon.$$

The measure μ describes the source of information and π describes the jitter (noise).

Let $\nu = (\mu \times \pi) \circ \varphi^{-1}$ be a corresponding factor measure on Y defined by

$$\nu(C) := (\mu \times \pi)\big(\varphi^{-1} C\big) \text{ for any Borel measurable } A \subseteq Y.$$

Despite the fact that some configurations are forbidden in Y, in other words, we have some "hard-core" constraints, there is a rich theory of Gibbs measures for sofic subshifts. One of the equivalent ways to define Gibbs measures is as follows. We say that an invariant measure ρ on Y is Gibbs for a Hölder continuous function $\varphi : Y \to \mathbb{R}$ and constants P and $C > 1$ such that for any $y \in Y$ one has

$$C^{-1} \leq \frac{\rho([y_0, y_1, \ldots, y_n])}{\exp\big(\sum_{k=0}^{n} \varphi(\sigma^k y) - (n+1)P\big)} \leq C, \tag{6.1}$$

where $\sigma : Y \to Y$ is the left shift.The function φ is often called a *potential*, and has a role analogous to that of $f_U(\cdot) = \sum_{0 \in A} U(A, \cdot)/|A|$ for standard lattice systems. The constant P in (6.1) is in fact the pressure of φ.

Now, (6.1), often called the *Bowen-Gibbs* property, implies that for all $y \in Y$

$$\varphi(y) - C_1 \leq \log \rho(y_0 | y_1, \ldots, y_n) \leq \varphi(y) + C_1, \tag{6.2}$$

for some positive constant C_1. Since Y is compact, and φ is continuous, we conclude that for every y and all $n \in \mathbb{N}$ the logarithm of the conditional probability $\rho(y_0|y_1,\ldots,y_n)$ is bounded from below and above.

It turns out that $\nu=(\mu\times\pi)\circ\varphi^{-1}$ is not Gibbs. As usual in the study of non-Gibbsianity we have to indicate a bad configuration. In our case, configuration 02^∞ is a bad configuration for ν. Consider cylinder $[y_0,\ldots,y_n]$ where

$$y_0=0,\ y_1=\ldots=y_n=2$$

Then effectively there is a unique preimage of this cylinder. Indeed $y_0=0$, and as we have seen above, this is possible only for

$$x_0=2,\ \omega_0=-1,\ \omega_{-1}=1.$$

For the next position $i=1$ we have

$$2=y_1=x_1+\omega_1-\omega_0=x_1+\omega_1+1.$$

Again, since $\omega_1+1\geq 0$ and $x_1\geq 2$, this is possible if and only if $\omega_1=-1$ and $x_1=2$. But then $x_2=2$ and $\omega_2=-1$, and so on. Therefore

$$\varphi^{-1}([0,\underbrace{2,2,\ldots,2}_{n\text{ times}}])\subseteq[\underbrace{2,2,\ldots,2}_{n+1\text{ times}}]\times[\underbrace{-1,-1,\ldots,-1}_{n+1\text{ times}}],$$

and hence

$$\nu([0,2,2,\ldots,2])\leq\mu([2,2,\ldots,2])\pi([-1,-1,\ldots,-1])=(p_2\epsilon)^{n+1}.$$

On the other hand, cylinder $[2,2,\ldots,2]$ has many preimages. For example, with appropriate choice of ω's cylinders of the form

$$[x_1,\ldots,x_n]=[2,\ldots,2,3,2,\ldots,2]\subseteq X$$

will project into $[2,\ldots,2]$. Indeed, if j is the position of 3 in $[x_1,\ldots,x_n]$, then the choice $\omega_0=\omega_1=\ldots=\omega_{j-1}=0$, and $\omega_j=\omega_{j+1}=\ldots=\omega_n=-1$ will suffice. Therefore

$$\nu([2,2,\ldots,2])\geq\sum_{j=1}^{n}p_2^{n-1}p_3(1-2\epsilon)^j\epsilon^{n-j+1},$$

and for $\epsilon<1/3$, one has

$$\nu([2,2,\ldots,2])\geq np_2^{n-1}p_3\epsilon^{n+1},$$

and therefore

$$\nu(0|2,2,\ldots,2)=\frac{\nu([0,2,2,\ldots,2])}{\nu([2,2,\ldots,2])}\leq\frac{C}{n},$$

and thus the logarithm of $\nu(0|2,2,\ldots,2)$ is not uniformly bounded from below, and hence there is no Hölder continuous φ such that (6.2) is valid for ν, and hence, ν is not Gibbs.

A slightly more accurate analysis shows that ν is not Gibbs for $\epsilon > 1/3$ as well.

An interesting open problem is the computation of the capacity of the bit-shift channel with a fixed jitter measure π. This problem reduces to the computation of the entropy of the transformed measure ν for an arbitrary input measure μ. In [1] an efficient algorithm was proposed for Bernoulli measures μ. This algorithm produces accurate (to arbitrary precision) numerical lower and upper bounds on the entropy of ν.

7. Discussion

In this paper we addressed the problem of finding sufficient conditions under which $h(\nu|\mu) = 0$ implies that ν is consistent with a given specification γ for μ. In particular, the question is interesting in the case of an almost or a weakly Gibbs measure μ. Intuition developed in [4, 10, 14] shows that ν must be concentrated on a set of "good" configurations for measure μ. In the case μ is almost Gibbs, ν must be concentrated on the continuity points Ω_γ, [10]. A natural generalization to the case of a weakly Gibbs measure μ for potential U would be to assume that ν is concentrated on the convergence points of the Hamiltonian H^U. However, this is not true as the counterexample of [10] shows.

We weakened and generalized the conditions under which we can prove the first part of the Variational Principle. Moreover, we introduced the class of Intuitively Weak Gibbs measures, which is strictly larger than the almost Gibbs class, but contained in the Weak Gibbs class.

The example considered in this paper shows (and we conjecture the same type of behaviour for other interesting examples of weakly Gibbs measures) that some weak Gibbs measures are more regular than was thought before.

References

[1] Stan Baggen and Vladimir Balakirsky, *An efficient algorithm for computing the entropy of output sequence for bitshift channels*, Proceedings of the 24^{th} Symposium on Information Theory in Benelux (2003), 157-164.

[2] Aernout C. D. van Enter, *The renormalization-group peculiarities of Griffiths and Pearce: what have we learned?*, Mathematical results in statistical mechanics (Marseilles, 1998), 1999, pp. 509–526.

[3] Aernout C. D. van Enter, Roberto Fernández, and Alan D. Sokal, *Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory*, J. Statist. Phys. **72** (1993), no. 5-6, 879–1167.

[4] Aernout C. D. van Enter and Senya B. Shlosman, *(Almost) Gibbsian description of the sign fields of SOS fields*, J. Statist. Phys. **92** (1998), no. 3-4, 353–368.
[5] Hans Föllmer, *On entropy and information gain in random fields*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **26** (1973), 207–217.
[6] Hans-Otto Georgii, *Gibbs measures and phase transitions*, de Gruyter Studies in Mathematics, vol. 9, Walter de Gruyter & Co., Berlin, 1988.
[7] Sheldon Goldstein, *A Note on Specifications*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **46** (1978), 45–51.
[8] Christof Külske, *(Non-) Gibbsianness and phase transitions in random lattice spin models*, Markov Process. Related Fields **5** (1999), no. 4, 357–383.
[9] ———, *Weakly Gibbsian representations for joint measures of quenched lattice spin models*, Probab. Theory Related Fields **119** (2001), no. 1, 1–30.
[10] Christof Külske, Arnaud Le Ny, and Frank Redig, *Relative entropy and variational properties of generalized Gibbsian measures*, Ann. Probab. **32** (2004), no. 2, 1691–1726.
[11] Douglas Lind and Brian Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.
[12] Christian Maes, Annelies van Moffaert, and Frank Redig, *Almost Gibbsian versus weakly Gibbsian measures*, Stoch. Proc. Appl. **79** (1999), 1–15.
[13] Christian Maes, Frank Redig, and Annelies Van Moffaert, *The restriction of the Ising model to a layer*, J. Statist. Phys. **96** (1999), no. 1-2, 69–107.
[14] Christian Maes, Frank Redig, Floris Takens, Annelies van Moffaert, and Evgeny Verbitski, *Intermittency and weak Gibbs states*, Nonlinearity **13** (2000), no. 5, 1681–1698.
[15] Chris Preston, *Random fields* (1976), ii+200. Lecture Notes in Mathematics, Vol. 534.
[16] Shlomo Shamai (Shiltz) and Ephraim Zehavi, *Bounds on the Capacity of the Bit-Shift Magnetic Recording Channel*, IEEE Trans. Inform. Theory **37** (1991), no. 3, 863–872.
[17] Shaogang Xu, *An ergodic process of zero divergence-distance from the class of all stationary processes*, J. Theoret. Probab. **11** (1998), no. 1, 181–195.

A.C.D.van Enter, Institute for Theoretical Physics, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
E-mail address: `a.c.d.van.enter@phys.rug.nl`

E.A. Verbitskiy, Philips Research Laboratories, Prof. Holstlaan 4 (WO 2), 5656 AA Eindhoven, The Netherlands
E-mail address: `evgeny.verbitskiy@philips.com`