Journal of Statistical Physics, Vol. 118, Nos. 3/4, February 2005 (© 2005)
DOI: 10.1007/s10955-004-8821-5

Chains with Complete Connections: General Theory, Uniqueness, Loss of Memory and Mixing Properties

Roberto Fernández[1] **and Grégory Maillard**[1]

Received April 19, 2004; accepted September 20, 2004

We introduce a statistical mechanical formalism for the study of discrete-time stochastic processes with which we prove: (i) General properties of extremal chains, including triviality on the tail σ-algebra, short-range correlations, realization via infinite-volume limits and ergodicity. (ii) Two new sufficient conditions for the uniqueness of the consistent chain. The first one is a transcription of a criterion due to Georgii for one-dimensional Gibbs measures, and the second one corresponds to the Dobrushin's criterion in statistical mechanics. (iii) Results on loss of memory and mixing properties for chains in the Dobrushin regime. These results are complementary to those existing in the literature, and generalize the Markovian results based on the Dobrushin ergodic coefficient.

KEY WORDS: Discrete-time stochastic processes; chains with complete connections; uniqueness criteria, mixing rates, Markov chains.

1. INTRODUCTION

Chains with complete connections is the name coined by ref. 19 for discrete-time stochastic processes whose dependence on the past is not necessarily Markovian. The theory of these processes has many points in common with the theory of Gibbs measures in statistical mechanics – particularly, the existence of phase transitions. Nevertheless there is a clear difference, at the formal level, between both theories. Indeed, processes are described in terms of *single-site* transition probabilities, while Gibbs measures are characterized by their conditional probabilities for *arbitrary* finite regions (specifications). In this paper we propose a natural way to

[1]Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen, Site Colbert, F-76821 Mont Saint Aignan, France; e-mail: {roberto.fernandezgregory.maillard}@univ-rouen.fr

reduce this asymmetry, by introducing a statistical–mechanical framework for the study of processes. This framework establishes a more direct relation between both theories, which allows us to reproduce, for chains with complete connections, a number of benchmark Gibbsian results.

We present three types of results. First, we obtain general properties of extremal chains for any type of alphabet, namely triviality on the tail σ-algebra, short-range correlations, realization via infinite-volume limits and ergodicity. Second, we produce some new sufficient conditions for the uniqueness of the consistent chain. On the one hand, we obtain a transcription of a criterion given by ref. 8 for one-dimensional Gibbs fields, which generalizes Ruelle's (ref. 20) work. This criterion comes close to be optimal for the latter, for instance it pinpoints the absence of phase transition for two-body spin models with a $1/r^{2+\varepsilon}$-interaction, for all $\varepsilon > 0$. The criterion imposes no restriction on the type of alphabet. On the other hand we prove a "one-sided" Dobrushin's criterion, which corresponds to a well known uniqueness criterion in statistical mechanics (see, for instance, ref. 21, Chapter V). This criterion is valid for systems with a compact metric alphabet. We exhibit simple examples where Dobrushin's criterion applies but that fall outside the scope of most other known uniqueness criteria (see refs. 1, 10, 12, 13, 22 and 24).

Our third type of results refer to loss of memory and mixing properties of chains in the Dobrushin regime. Our results, obtained along the lines of a similar Gibbsian theory (again we refer the reader to Chapter V of ref. 21), are complementary, both in their precision and in their range of applicability, to similar results available in the literature (See refs. 2, 11 and references therein). The results depend on a *sensitivity matrix* that generalizes the Dobrushin ergodic coefficient of Markov chains.

Our approach is based on a notion analogous to the specifications in statistical mechanics, which we call *left interval-specifications* (LIS). These are kernels for regions in the form of intervals which depend on the preceding history of the process. In contrast, Gibbsian specifications involve arbitrary finite regions and depend of the configuration on the whole exterior of the region. This amounts, in one dimension, to a dependence on both past and future. The difference is, of course, a consequence of the "one-sidedness" associated to a stochastic (time) evolution, as compared with the lack of favored direction in the spatial description provided by a Gibbs measure.

The description in terms of LIS is totally equivalent to the traditional description in terms of transition probabilities (=LIS singletons). We show this in our first theorem. But, as this paper illustrates, our approach has the advantage of allowing us to "import", in a natural manner, notions, techniques and arguments from statistical mechanics. It may also be useful in the opposite direction, namely to explore the consequences of known properties of chains for the theory of Gibbs measures. As a step in this direction, in a companion paper[6] we study conditions under which chains and Gibbs measures can be identified. On a more conceptual level, we believe that our statistical mechanical approach is more appropriate to study the general situation where several different chains are consistent with the same transition probabilities (see refs. 3, 4 and 16). Statistical mechanics is the framework developed, precisely, to study this phenomenon which corresponds to the appearance of (first-order) phase transitions.

2. PRELIMINARIES

We consider a measurable space $(E, \mathcal{E})$ and a subset $\Omega \subset E^{\mathbb{Z}}$. The exponent $\mathbb{Z}$ stands, in fact, for any countable set with a total order. The group structure of $\mathbb{Z}$ will play no role, except in Theorem 3.9 where $\mathbb{Z}$ acts by isomorphisms. The elements of $\mathbb{Z}$ are called *sites*, and those of Ω (*admissible*) *configurations*. The space E is sometimes called *alphabet*. We endow Ω with the projection $\mathcal{F}$ of the product σ-algebra associated to $E^{\mathbb{Z}}$. When we invoke topological notions (e.g., compactness) the σ-algebra $\mathcal{E}$ is assumed to be Borelian. We adopt the following notation

- Let $\Lambda \subset \mathbb{Z}$. For a configuration $\sigma \in E^{\mathbb{Z}}$ we denote $\sigma_\Lambda = (\sigma_i)_{i\in\Lambda} \in E^\Lambda$. The set of admissible configurations in Λ is $\Omega_\Lambda \triangleq \{\sigma_\Lambda \in E^\Lambda : \exists\, \omega \in \Omega \text{ with } \omega_\Lambda = \sigma_\Lambda\}$, while $\mathcal{F}_\Lambda$ is the sub-σ-algebra of $\mathcal{F}$ generated by the cylinders with base in Ω_Λ. If $\Delta \subset \mathbb{Z}$ with $\Lambda \cap \Delta = \emptyset$, $\omega_\Lambda \sigma_\Delta$ denotes the configuration on $\Lambda \cup \Delta$ coinciding with ω_i for $i \in \Lambda$ and with σ_i for $i \in \Delta$.

- We denote $\mathcal{S}_b$ the set of finite intervals of $\mathbb{Z}$. When $\Lambda = [k, n] \in \mathcal{S}_b$ we shall also use the "sequence" notation: $\omega_k^n \triangleq \omega_{[k,n]} = \omega_k, \ldots, \omega_n$; $\Omega_k^n \triangleq \Omega_{[k,n]}$; etc. If $\Lambda = [k, +\infty[$, the notation will be analogous but with $+\infty$ as upper limit.

- If $n \in \mathbb{Z}$, $\mathcal{F}_{\leqslant n} \triangleq \mathcal{F}_{]-\infty,n]}$. For every $\Lambda \in \mathcal{S}_b$ we denote $l_\Lambda \triangleq \min \Lambda$; $m_\Lambda \triangleq \max \Lambda$; $\Lambda_- =]-\infty, l_\Lambda - 1]$.

- For kernels associated to a LIS (defined below), $\lim_{\Lambda\uparrow V} f_\Lambda$ is the limit of the net $\{f_\Lambda, \{\Lambda\}_{\Lambda\in\mathcal{S}_b}, \Lambda \subset V, \subset\}$, for V an infinite interval of $\mathbb{Z}$. If μ a measure on $(\Omega, \mathcal{F})$ and h a $\mathcal{F}$-measurable function, we will write $\mu(h)$ instead of $E_\mu(h)$.

Definition 2.1 (LIS). A **left interval-specification** f on $(\Omega, \mathcal{F})$ is a family of probability kernels $\{f_\Lambda\}_{\Lambda\in\mathcal{S}_b}$, $f_\Lambda : \mathcal{F}_{\leqslant m_\Lambda} \times \Omega \longrightarrow [0, 1]$ such that for all Λ in $\mathcal{S}_b$,

(a) For each $A \in \mathcal{F}_{\leqslant m_\Lambda}$, $f_\Lambda(A \mid \cdot)$ is $\mathcal{F}_{\Lambda_-}$-measurable.

(b) For each $B \in \mathcal{F}_{\Lambda_-}$ and $\omega \in \Omega$, $f_\Lambda(B \mid \omega) = 1\!\!1_B(\omega)$.

(c) For each $\Delta \in \mathcal{S}_b : \Delta \supset \Lambda$,

$$f_\Delta f_\Lambda = f_\Delta \quad \text{on } \mathcal{F}_{\leqslant m_\Lambda}, \tag{2.1}$$

that is, $(f_\Delta f_\Lambda)(h \mid \omega) = f_\Delta(h \mid \omega)$ for each $\mathcal{F}_{\leqslant m_\Lambda}$-measurable function h and configuration $\omega \in \Omega$.

These conditions are analogous to those defining a *specification* in the theory of Gibbs measures (see ref. 9, for instance). Two important differences should be highlighted, however, both being a consequence of the "directional" character of the notion of process. First, the LIS kernels act only on functions measurable towards the left, while Gibbsian specifications have no similar constraint. As a consequence, LIS kernels involve only conditioning with respect to the past [property (b)], while Gibbsian kernels condition with respect to the whole exterior of Λ. Second, LIS kernels are defined only for intervals while Gibbsian kernels are defined for all finite sets of sites.

Property (c) is usually labeled *consistency*. There and in the sequel we adopt the standard notation for a composition of probability kernels or of a probability kernel with a measure. Explicitly, (2.1) means that

$$\iint h(\xi)\, f_\Lambda(d\xi \mid \sigma)\, f_\Delta(d\sigma \mid \omega) = \int h(\sigma)\, f_\Delta(d\sigma \mid \omega)$$

for each $\mathcal{F}_{\leqslant m_\Lambda}$-measurable function h and configuration $\omega \in \Omega$.

Definition 2.2 (left interval-consistency). A probability measure μ on $(\Omega, \mathcal{F})$ is said to be consistent with a LIS f if for each $\Lambda \in \mathcal{S}_b$

$$\mu f_\Lambda = \mu \quad \text{on } \mathcal{F}_{\leqslant m_\Lambda}. \tag{2.2}$$

Such a measure μ is called a chain with complete connections, or simply a chain, consistent with the LIS f. The family of these measures will be denoted $\mathcal{G}(f)$.

Remark 2.3. A *Markov LIS of range k* is a LIS such that each function $f_\Lambda(A \mid \cdot)$ is measurable with respect to $\mathcal{F}_{[l_\Lambda - k, l_\Lambda - 1]}$, for each $A \in \mathcal{F}_\Lambda$. A chain consistent with such a LIS is a *Markov chain of range k*.

Remark 2.4. *Chains with complete connections* is the original nomenclature introduced by Onicescu and Mihoc (ref. 19). These objects have been later reintroduced under a panoply of names, some associated to particular additional properties, others to notions later proven to be equivalent. Among them we mention: *chains of infinite order*,[10] *g-measures*,[15] list processes,[17] *uniform martingales* or *random Markov processes*.[14]

3. RESULTS ON GENERAL FRAMEWORK

We start by making the connection with the traditional definition of chains based on singleton kernels.

Theorem 3.1 (Singleton consistency for chains). Let $(f_i)_{i\in\mathbb{Z}}$ be a family of probability kernels $f_i : \mathcal{F}_{\leqslant i} \times \Omega \to [0,1]$ such that for each $i \in \mathbb{Z}$

(a) For each $A \in \mathcal{F}_{\leqslant i}$, $f_i\,(A \mid \cdot)$ is $\mathcal{F}_{\leqslant i-1}$-measurable.

(b) For each $B \in \mathcal{F}_{\leqslant i-1}$ and $\omega \in \Omega$, $f_i\,(B \mid \omega) = \mathbb{1}_B(\omega)$.

Then the LIS $f = \{f_\Lambda\}_{\Lambda\in\mathcal{S}_b}$ defined by

$$f_\Lambda = f_{l_\Lambda}\, f_{l_\Lambda+1} \cdots f_{m_\Lambda} \tag{3.1}$$

is the unique LIS such that $f_{\{i\}} = f_i$ for all $i \in \mathbb{Z}$. Furthermore,

$$\mathcal{G}(f) = \Big\{\mu : \mu f_i = \mu, \text{ for all } i \text{ in } \mathbb{Z}\Big\}. \tag{3.2}$$

In particular, the theorem shows that any LIS f enjoys the factorization property

$$f_\Lambda = f_{\{l_\Lambda\}}\, f_{\{l_\Lambda+1\}} \cdots f_{\{m_\Lambda\}} \tag{3.3}$$

on $\mathcal{F}_{\leqslant m_\Lambda}$ for each $\Lambda \in \mathcal{S}_b$. By recurrence this yields

$$f_{[l,m]} = f_{[l,n]}\, f_{[n+1,m]} \tag{3.4}$$

for any $l, n, m \in \mathbb{Z}$ with $l \leqslant n < m$.

The following three theorems establish relations among extremality, triviality, mixing properties and infinite-volume limits similar to those valid for Gibbs measures or, more generally, for measures consistent with specifications. Their proofs, presented in Section 6, are patterned on the Gibbsian proofs, taking care of the one-sided measurability of the LIS kernels.

Theorem 3.2 (Extremality and triviality). Let $f=(f_\Lambda)_{\Lambda\in\mathcal{S}_b}$ be a left interval-specification on $(\Omega,\mathcal{F})$. Denote by $\mathcal{F}_{-\infty}\triangleq\bigcap_{k\in\mathbb{Z}}\mathcal{F}_{\leqslant k}$ the tail σ-algebra. Then

(a) $\mathcal{G}(f)$ is a convex set.

(b) A measure μ is extreme in $\mathcal{G}(f)$ if and only if μ is trivial on $\mathcal{F}_{-\infty}$.

(c) Let $\mu\in\mathcal{G}(f)$ and $\nu\in\mathcal{P}(\Omega,\mathcal{F})$ such that $\nu\ll\mu$. Then $\nu\in\mathcal{G}(f)$ if and only if there exists a $\mathcal{F}_{-\infty}$-measurable function $h\geqslant 0$ such that $\nu=h\mu$.

(d) Each $\mu\in\mathcal{G}(f)$ is uniquely determined (within $\mathcal{G}(f)$) by its restriction to the tail σ-algebra $\mathcal{F}_{-\infty}$.

(e) Two distinct extreme elements μ,ν of $\mathcal{G}(f)$ are mutually singular on $\mathcal{F}_{-\infty}$.

Theorem 3.3 (Triviality and short-range correlations). For each probability measure on $(\Omega,\mathcal{F})$, the following statements are equivalent.

(a) μ is trivial on $\mathcal{F}_{-\infty}$.

(b) $\lim_{\Lambda\uparrow\mathbb{Z}}\sup_{B\in\mathcal{F}_{\Lambda_-}}|\mu(A\cap B)-\mu(A)\mu(B)|=0$, for all cylinder sets A in $\mathcal{F}$.

(c) $\lim_{\Lambda\uparrow\mathbb{Z}}\sup_{B\in\mathcal{F}_{\Lambda_-}}|\mu(A\cap B)-\mu(A)\mu(B)|=0$, for all $A\in\mathcal{F}$.

Theorem 3.4 (Infinite volume limits). Let f be a LIS, μ an extreme point of $\mathcal{G}(f)$ and $(\Lambda_n)_{n\geqslant 1}$ a sequence of regions in $\mathcal{S}_b$ such that $\Lambda_n\uparrow\mathbb{Z}$. Then

(a) $f_{\Lambda_n}h\to\mu(h)$ μ-a.s. for each bounded local function h on Ω.

(b) If Ω is a compact metric space, then for μ-almost all $\omega\in\Omega$, $f_{\Lambda_n}h\to\mu(h)$ for all continuous local functions h on Ω.

As customary, we are calling a function *local* if it is $\mathcal{F}_\Lambda$-measurable for some finite $\Lambda\subset\mathbb{Z}$.

The following theorem is the only result in the paper where we consider translation invariance. We briefly recall the relevant notions. We consider the (right) shift $\tau(i)=i+1$. (More generally, the same theory applies to any action of $\mathbb{Z}$ on $\mathbb{Z}$ by isomorphisms. In the case of k-shifts such theory leads to k-periodic objects). The shift induces actions on configurations, measurable sets, measurable functions and measures that we denote with the same symbol: for $\omega\in\Omega$, $\tau(\omega)=\big(\omega_{i-1}\big)_{i\in\mathbb{Z}}$, for $A\in\mathcal{F}$, $\tau A=\{\tau\omega:$

$\omega \in A\}$, for h $\mathcal{F}$-measurable, $(\tau h)(\omega) = h(\tau^{-1}\omega)$, and for a measure μ on $(\Omega, \mathcal{F})$, $(\tau\mu)(h) = \mu(\tau^{-1}h)$. Objects invariant under the action of the shift are called shift-invariant. We denote $\mathcal{I}$ the σ-algebra of all shift-invariant measurable sets, and $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ the set of shift-invariant probability measures on $(\Omega, \mathcal{F})$. A measure in $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ is ergodic if it is trivial on $\mathcal{I}$.

For $k \in \mathbb{Z}$ and $\Lambda \subset \mathbb{Z}$ we denote $\Lambda + k = \{i + k : i \in \Lambda\}$. A LIS f is shift-invariant or stationary if

$$f_{\Lambda+1}(\tau A \mid \tau\omega) = f_\Lambda(A \mid \omega)$$

for each $\Lambda \in \mathcal{S}_b$ and $\omega \in \Omega$. We denote $\mathcal{G}_{\text{inv}}(f)$ the family of shift-invariant chains consistent with a LIS f.

Theorem 3.5 (Ergodic chains). Let f be a shift-invariant LIS.

(a) A chain $\mu \in \mathcal{G}_{\text{inv}}(f)$ is extreme in $\mathcal{G}_{\text{inv}}(f)$ if and only if μ is ergodic.

(b) Let $\mu \in \mathcal{G}_{\text{inv}}(f)$. If $\nu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ is such that $\nu \ll \mu$, then $\nu \in \mathcal{G}_{\text{inv}}(f)$.

(c) $\mathcal{G}_{\text{inv}}(f)$ is a face of $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$. More precisely, if $\mu, \nu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ and $0 < s < 1$ are such that $s\mu + (1-s)\nu \in \mathcal{G}_{\text{inv}}(f)$ then $\mu, \nu \in \mathcal{G}_{\text{inv}}(f)$.

4. UNIQUENESS RESULTS

We shall prove two types of uniqueness results. We start with the counterpart of a criterion proven by Georgii (ref. 8) for measures determined by specifications.

Theorem 4.1 (One-sided boundary-uniformity). Let f be a LIS for which there exists a constant $c > 0$ satisfying the following property: For every $m \in \mathbb{Z}$ and every cylinder set $A \in \mathcal{F}_{-\infty}^{m}$ there exists an integer $n < m$ such that

$$f_{[n,m]}(A \mid \xi) \geqslant c\, f_{[n,m]}(A \mid \eta) \quad \text{for all } \xi, \eta \in \Omega. \tag{4.1}$$

Then there exists at most one chain consistent with f.

The main virtue of this criterion is its generality. Other existing uniqueness criteria (See refs. 1, 10, 12, 13, 22 and 24) require that the space E have particular properties (finite, countable, compact), and that the kernels satisfy appropriate non-nullness hypotheses. The reader is referred to Fernández and Maillard (ref. 7) and Maillard (ref. 18) for a detailed survey of uniqueness criteria. In fact, many of these are based on summability properties of the sequence of variations:

$$\mathrm{var}_j(f_{\{i\}}) \triangleq \sup\left\{ \left| f_{\{i\}}(\xi_i \mid \xi_{-\infty}^{i}) - f_{\{i\}}(\eta_i \mid \eta_{-\infty}^{i}) \right| : \xi, \eta \in \Omega_{-\infty}^{i},\ \xi_j^i = \eta_j^i \right\} \tag{4.2}$$

for $j < i$.

Proposition 4.2. Assume that E is a countable set and $\mathcal{E}$ the discrete σ-algebra. A LIS f satisfies the one-sided boundary-uniformity condition (4.1) if it is uniformly non-null:

$$\inf_{i\in\mathbb{Z}} \inf_{\omega\in\Omega_{\leqslant i}} f_{\{i\}}\big(\omega_i \,\big|\, \omega_{-\infty}^{i-1}\big) > 0\,, \tag{4.3}$$

and satisfies

$$\sup_{n\in\mathbb{Z}} \sum_{i\geqslant n} \mathrm{var}_n \left(f_{\{i\}}\right) < +\infty\,. \tag{4.4}$$

We observe that when f is stationary the last condition amounts to summable variations: $\sum_{j<0} \mathrm{var}_j \left(f_{\{0\}}\right) < +\infty$.

Our second type of uniqueness result corresponds to the Dobrushin's criterion for specifications. The required mathematical setting is richer. We choose a bounded distance d on E and take $\mathcal{E}$ as the associated Borel σ-algebra. We endow $E^{\mathbb{Z}}$ with the product topology (so $\mathcal{F}$ is also Borel) and $\Omega \subset E^{\mathbb{Z}}$ with the restricted topology. The choice of distance is dictated by the type of measures to be analyzed. For finite, or countable, alphabets the canonical choice is the discrete distance $d_{\mathrm{disc}}(a,b) = 1$ if $a \neq b$ and 0 otherwise.

Definition 4.3. A LIS on f on $(\Omega, \mathcal{F})$ is continuous if the functions $\Omega \ni \omega \longrightarrow f_\Lambda(A \mid \omega)$ are continuous for all $\Lambda \in \mathcal{S}_b$ and all $A \in \mathcal{F}_\Lambda$.

In the case of specifications, continuity is associated with Gibbsianness (non-nullness is also needed, see, e.g., the discussion in Section 2.3.3 in ref. 23 Fernández and Sokal, 1993). For E finite, continuity is equivalent to $\lim_{j\to-\infty} \mathrm{var}_j \left(f_{\{i\}}\right) = 0$.

Remark 4.4. If the LIS f is continuous and the space Ω is compact, then there always exists at least one compatible chain. Indeed, the probability measures on a compact space form a (weakly) compact set. Hence, if $(\Lambda_n)_{n\in\mathbb{N}} \subset \mathcal{S}_b$ is any exhausting sequence of regions and $(\omega^{(n)})_{n\in\mathbb{N}} \subset \Omega$ any sequence of pasts, the sequence of measures $f_{\Lambda_n}(\cdot \mid \omega^{(n)})$, $n \in \mathbb{N}$, has some accumulation point. Continuity ensures that such a limit belongs to $\mathcal{G}(f)$. Therefore, for continuous LIS on a compact space of configurations, Theorems 4.1 and 4.6 determine conditions for the existence of *exactly* one compatible measure.

For every measurable function h, the d-oscillation of h with respect to the site $j \leqslant i$, is defined by

$$\delta_j^d(h) \triangleq \sup\left\{ \frac{|h(\xi)-h(\eta)|}{d\left(\xi_j,\eta_j\right)} : \xi,\eta \in \Omega_{-\infty}^{\tau(h)},\ \xi \overset{\neq j}{=} \eta \right\}, \tag{4.5}$$

with the convention $0/0=0$. Here $\tau(h)$ is the minimal $i\in\mathbb{Z}$ such that h is $\mathcal{F}_{\leqslant i}$–measurable and we introduced the notation

$$\xi \overset{\neq j}{=} \eta \quad \Longleftrightarrow \quad \xi_i = \eta_i\,,\ \forall\, i \neq j \tag{4.6}$$

("ξ equal to η off j"). We introduce also the space of functions of *bounded d-oscillations*:

$$\mathcal{B}_d \triangleq \left\{ \mathcal{F}\text{-measurable } h : \sup_{j\in\mathbb{Z}} \delta_j^d(h) < \infty \right\}, \tag{4.7}$$

and its restrictions

$$\mathcal{B}_d(\Lambda) \triangleq \left\{ h\in\mathcal{B}_d : h\ \mathcal{F}_\Lambda\text{-measurable} \right\}$$

for $\Lambda \subset \mathbb{Z}$. The most general version of Dobrushin's strategy allows the use of a "tiling" of $\mathbb{Z}$ by finite intervals. These intervals V must be chosen so that there is an appropriate control of the "sensitivity" of the averages f_V to the configuration in V_-.

Definition 4.5 (*d*-sensitivity estimator). Let $V \in \mathcal{S}_b$ and f_V a probability kernel on $\mathcal{F}_{\leqslant m_V} \times \Omega$. A d-sensitivity estimator for f_V is a non-negative matrix $\alpha^V = \left(\alpha_{ij}^V\right)_{i,j\in\mathbb{Z}}$ such that $\alpha_{ij}^V = 0$ if $i \notin V$ or $j \notin V_-$ and

$$\delta_j^d\left(f_V h\right) \leqslant \sum_{i\in V} \delta_i^d(h)\, \alpha_{ij}^V \tag{4.8}$$

for all $j \in V_-$ and $\mathcal{F}_V$-measurable functions $h \in \mathcal{B}_d$.

Theorem 4.6 (One-sided Dobrushin). Let f be a continuous LIS. If there exist a countable partition $\mathcal{P}$ of $\mathbb{Z}$ into finite intervals such that for each $V \in \mathcal{P}$ there exists a d-sensitivity estimator α^V for f_V with

$$\sum_{j \in V_-} \alpha_{ij}^V < 1 \tag{4.9}$$

for all $i \in \mathbb{Z}$, then there exists at most one chain consistent with f.

This criterion is not directly comparable with existing uniqueness results, which are based on the rates of variations (4.2). As an illustration, Example 4.7 below exhibits processes with arbitrarily slow power-law variation rates, but that nevertheless satisfy Dobrushin's criterion.

In particular, the partition can be trivial, namely $\mathcal{P} = \{\{i\} : i \in \mathbb{Z}\}$. In the stationary case, the estimators for such a partition are of the form $\alpha_{ij}^{\{i\}} = \alpha(i-j)$ for a certain function α on the integers that takes value zero for non-positive integers. Dobrushin's criterion becomes then $\sum_{n \geqslant 1} \alpha(-n) < 1$.

The customary way to construct d-sensitivity estimators for kernels f_V is resorting to the Vaserstein–Kantorovich–Rubinstein (VKR) distance between measures on $\mathcal{F}_V$ for the distance $d_V(\omega_V, \sigma_V) \triangleq \sum_{i \in V} d(\omega_i, \sigma_i)$. If we denote $\mathring{f}_V$ the projection of each kernel f_V over Ω_V:

$$\mathring{f}_V\left(A \,\middle|\, \omega_{-\infty}^{l_V-1}\right) \triangleq f_V\left(\{\sigma_V \in A\} \,\middle|\, \omega_{-\infty}^{l_V-1}\right), \quad \forall A \in \mathcal{F}_V,\ \forall \omega_{-\infty}^{l_V-1} \in \Omega_{-\infty}^{l_V-1},$$

then the VKR distances between these projections are

$$\begin{aligned} &\left\| \mathring{f}_V\left(\cdot \,\middle|\, \xi_{-\infty}^{l_V-1}\right) - \mathring{f}_V\left(\cdot \,\middle|\, \eta_{-\infty}^{l_V-1}\right) \right\|_{d_V} \\ &\quad = \sup\left\{ \left| \mathring{f}_V\left(h \,\middle|\, \xi_{-\infty}^{l_V-1}\right) - \mathring{f}_V\left(h \,\middle|\, \eta_{-\infty}^{l_V-1}\right) \right| : h \in \mathcal{B}_d(V),\ \mathrm{osc}_V(h) \leqslant 1 \right\} \end{aligned} \tag{4.10}$$

where $\mathrm{osc}_V(h) = \sup\{|h(\sigma_V) - h(\omega_V)| / d_V(\sigma_V, \omega_V)\}$. In fact, an optimal coupling argument (see, for instance, ref. 5 Section 11.8), yields the identity

$$\begin{aligned} &\left\| \mathring{f}_V\left(\cdot \,\middle|\, \xi_{-\infty}^{l_V-1}\right) - \mathring{f}_V\left(\cdot \,\middle|\, \eta_{-\infty}^{l_V-1}\right) \right\|_{d_V} \\ &\quad = \inf\left\{ \int d(\sigma_V, \omega_V)\, \rho(d\sigma_V, d\omega_V) : \rho \in \mathcal{P}(\Omega \times \Omega) \right. \\ &\qquad\qquad \left. \text{with marginals } \mathring{f}_V\left(\cdot \,\middle|\, \xi_{-\infty}^{l_V-1}\right) \text{ and } \mathring{f}_V\left(\cdot \,\middle|\, \eta_{-\infty}^{l_V-1}\right) \right\}. \end{aligned} \tag{4.11}$$

The VKR (canonical) d-estimator is defined by the coefficients

$$C_{ij}^{V}(f) \triangleq \sup_{\substack{\xi,\eta\in\Omega_{-\infty}^{l_V-1} \\ \xi \overset{\neq j}{=} \eta}} \frac{\left\| \overset{\circ}{f}_V\,(\,\cdot\,|\,\xi) - \overset{\circ}{f}_V\,(\,\cdot\,|\,\eta)\right\|_{d_V}}{d\left(\xi_j,\eta_j\right)}, \quad i\in V,\, j\in V_- \tag{4.12}$$

and $C_{ij}^{V}(f)=0$ otherwise.

If the partition is trivial and d is the discrete metric, each $\left\|\cdot\right\|_{d_{\{i\}}}$ coincides with the variational norm. If the alphabet E is countable, this means

$$C_{ij}^{\{i\}} = \delta_j(f_i) \triangleq \delta_j^{d_{\rm disc}}(f_i) = \sup\left\{|f_i(\xi)-f_i(\eta)| : \xi,\eta\in\Omega\,,\ \xi \overset{\neq j}{=} \eta\right\}. \tag{4.13}$$

and a sufficient condition for Dobrushin's criterion (4.1) is, therefore,

$$\sum_{j<i} \delta_j(f_i) < 1\,, \quad i\in\mathbb{Z}\,. \tag{4.14}$$

Furthermore, the one-sided Dobrushin criterion enlarges the scope of uniqueness criteria. The following example gives a LIS which does not satisfy any known uniqueness criteria except one-sided Dobrushin's.

Example 4.7. Consider the 2-letter alphabet $E=\{0,1\}$ and a shift-invariant LIS defined by singletons

$$f\left(\omega_0=1\,\middle|\,\omega_{-\infty}^{-1}\right) = \sum_{i\leqslant 0} a_i\,\omega_i\,, \tag{4.15}$$

for a sequence $\{a_i\}_{i\leqslant 0}$ of non-negative numbers. The estimators (4.13) yield a sensitivity matrix

$$\alpha_{ij} = \delta_j\left(f_{\{i\}}\right) = a_{i-j} \tag{4.16}$$

for $i>j$, and zero otherwise. Theorem (4.6) is therefore applicable as long as $\sum_{i\leqslant 0} a_i<1$. On the other hand, for each $0<\varepsilon<1$, the choice

$$a_{-k} = \frac{1-\varepsilon}{M_\varepsilon}\,\frac{1}{k^{1+\varepsilon}} \tag{4.17}$$

with $M_\varepsilon = \sum_{k\geqslant 1} k^{-(1+\varepsilon)}$, satisfies

$$\mathrm{var}_j(f_{\{i\}}) \geqslant \frac{1}{(i-j-1)^\varepsilon}$$

for $i-j\geqslant 2$. Thus, this LIS is not covered by any uniqueness criteria on the continuity rate except one-sided Dobrushin's.

Besides the absence of non-nullness hypotheses, an advantage of Dobrushin's criterion is that it determines a regime where mixing properties can be determined, as we discuss in next section.

To conclude, we remark that in fact the two uniqueness criteria given in Theorems 4.1 and 4.6 give a very strong form of uniqueness.

Definition 4.8 (HUC). A LIS f on $(\Omega, \mathcal{F})$ satisfies a **hereditary uniqueness condition** (HUC) if for all intervals of the form $\Gamma = [k, +\infty[$, $k\in\mathbb{Z}$, and configurations $\omega\in\Omega$, the LIS $f^{(\Gamma,\omega)}$ defined by

$$f_\Lambda^{(\Gamma,\omega)}(\cdot\,|\,\xi) = f_\Lambda(\cdot\,|\,\omega_{\Gamma_-}\,\xi_\Gamma)\,, \qquad \Lambda\in\mathcal{S}_b, \quad \Lambda\subset\Gamma \tag{4.18}$$

admits at most one consistent unique chain.

The two criteria given above involve bounds valid for all past conditions. They remain, therefore, valid if only particular pasts are considered as in (4.18). This observation proves the following corollary.

Corollary 4.9. If a LIS satisfies the hypotheses of either Theorem 4.1 or Theorem 4.6, then it also satisfies a HUC.

We remark that, for similar reasons, the criteria of refs. 10, 13 and 22 also imply the validity of a HUC.

5. RESULTS ON LOSS OF MEMORY AND MIXING PROPERTIES

We place ourselves in the framework needed for the one-sided Dobrushin criterion – E with a topology defined by a bounded metric d, $\mathcal{E}$ its Borel σ-algebra, Ω topologized with the restricted product topology – and take up all the related notions – d-oscillations, functions of bounded oscillations, sensitivity estimators. To improve readability, we write the results only for a trivial partition $\mathcal{P}$. Versions for more general partitions, of potential interest for coarse-graining arguments, can be obtained in a straightforward manner from our proofs by replacing sites by blocks of sites.

Definition 5.1. A **d-sensitivity matrix** for a LIS f is a matrix of the form

$$\alpha_{ij} \triangleq \begin{cases} \alpha_{ij}^{\{i\}} & \text{if } i > j \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

where each $\alpha_{ij}^{\{i\}}$ is a d-sensitivity estimator for f_i, $i \in \mathbb{Z}$.

For all $\Lambda \in \mathcal{S}_b$ we define the Λ-projection

$$(P_\Lambda)_{kj} = \begin{cases} 1 & \text{if } k = j \text{ and } k \in \Lambda \\ 0 & \text{otherwise .} \end{cases}$$

Theorem 5.2 (Loss of memory). Let f be a continuous LIS and (α_{ij}) a d-sensitivity matrix for f. Then,

(i) For every $\Lambda \in \mathcal{S}_b$, $j < l_\Lambda$ and $h \in \mathcal{B}_d(\Lambda)$,

$$\delta_j^d\,(f_\Lambda h) \leqslant \sum_{k \in \Lambda} \delta_k^d(h) \left[\sum_{l=1}^{|\Lambda|} (P_\Lambda \alpha)^l \right]_{kj} \tag{5.2}$$

(ii) Assume that there exist a function $F : \mathbb{Z}^2 \to \mathbb{R}^+$ satisfying the triangular inequality $F(i, j) \leqslant F(i, k) + F(k, j)\ \forall i, j, k \in \mathbb{Z}$ such that

$$\gamma_i \triangleq \sum_{j<i} \alpha_{ij}\, e^{F(i,j)} < 1\,, \tag{5.3}$$

for each $i \in \mathbb{Z}$. Then, for each $\Lambda \in \mathcal{S}_b$, $h \in \mathcal{B}_d(\Lambda)$ and $j < l_\Lambda$.

$$\delta_j^d\,(f_\Lambda h) \leqslant \frac{\gamma_\Lambda}{1 - \gamma_\Lambda} \sum_{k \in \Lambda} \delta_k^d(h)\, e^{-F(k,j)}\,, \tag{5.4}$$

with $\gamma_\Lambda = \max_{i \in \Lambda} \gamma_i$.

Remark 5.3. In the Markovian case $\alpha_{ij} = 0$ if $|i - j| > 1$. Then expression (5.2) implies that for $h \in \mathcal{F}_{\{n\}}$

$$\delta_{-1}^d\Big(f_{[0,n]}(h)\Big) \leqslant \gamma^n\, \delta_n^d(h) \tag{5.5}$$

with $\gamma = \sup_i \sum_j \alpha_{ij}$. For d discrete and estimators (4.12), γ is known as the **Dobrushin ergodic coefficient**. If, in addition, E is countable, $\Omega = E^{\mathbb{Z}}$ and f shift-invariant, then

$$\gamma = 1 - \min_{\sigma_{-1},\omega_{-1}\in E} \sum_{\omega_0\in E} f_{\{0\}}\big(\omega_0 \,\big|\, \sigma_{-1}\big) \wedge f_{\{0\}}\big(\omega_0 \,\big|\, \omega_{-1}\big) . \tag{5.6}$$

Remark 5.4. If the alphabet E is countable and the metric discrete we can use the estimators (4.13). With this choice, (5.4) implies

$$\begin{aligned} &\delta_j(f_i) \leqslant \mathrm{const}\, e^{-F(i,j)} \\ &\qquad \Longrightarrow \quad \delta_{-n}[f_{[0,m]}(A)] \leqslant \mathrm{const}\, e^{-F(m,-n)}\,,\ A\in \mathcal{F}_{\{m\}}\,. \end{aligned} \tag{5.7}$$

Published loss-of-memory results (refs. 2 and 11) resort instead to the variations (4.2). Comparisons can only be made through the obvious inequalities

$$\delta_j[f_{\{i\}}(h)] \leqslant \mathrm{var}_j[f_{\{i\}}(h)] \leqslant \sum_{k\leqslant j} \delta_k[f_{\{i\}}(h)]\,.$$

For LIS with an exponentially decaying dependence on the past, (5.7) implies an exponential loss of memory with an identical rate, in terms either of oscillations or of variations. This should be contrasted with the results in ref. 2 where there is an infinitesimal loss of rate. LIS with a power-law dependence can be treated by taking $F(i, j) = c\, \log(1 + |i - j|)$. In terms of variations, the loss of memory implied by (5.7) is also a power law but with a power decreased by one unit. Bressaud, Fernández and Galves (ref. 2) obtain, instead, the same power.

Furthermore, it is relatively simple to construct examples falling outside the scope of all preexisting loss-of-memory results, but for which Theorem 5.2 applies. It is the case, for instance, of the LIS constructed in Example (4.7) which is not covered by the results of Iosifescu (ref. 11) or of ref. 2.

The following mixing results form the LIS version of a well known chapter in the theory for Gibbs measures (see, for example, Chapter V in ref. 21). Their proofs, presented in Section 8, follow the guidelines of the statistical mechanical proofs. They require a compact Ω. We observe that example (4.15)–(4.17) shows that our results are complementary to those existing in the literature, which are based on variations rather than oscillations (See ref. 2, and references therein).

Theorem 5.5. Assume Ω compact and let f and $\widetilde{f}$ be two LIS on $(\Omega, \mathcal{F})$ with f continuous and with a unique consistent measure. Assume also that for each $i \in \mathbb{Z}$ there exists a measurable function b_i on Ω such that

$$\left\| \overset{\circ}{f}_{\{i\}} (\cdot \mid \omega) - \overset{\circ}{\widetilde{f}}_i (\cdot \mid \omega) \right\|_d \leqslant b_i(\omega) \tag{5.8}$$

for every configuration $\omega \in \Omega_{-\infty}^{i-1}$. Then, for all $\mu \in \mathcal{G}(f)$, $\widetilde{\mu} \in \mathcal{G}(\widetilde{f})$ and $\Lambda \in \mathcal{S}_b$

$$\left|\mu(h) - \widetilde{\mu}(h)\right| \leqslant \sum_{k \in \Lambda \cup \Lambda_-} \widetilde{\mu}(b_k)\, \delta_k^d \left(f_{[k+1, m_\Lambda]} h\right) \tag{5.9}$$

for every $h \in \mathcal{B}_d(\Lambda)$.

Let us denote $D \triangleq \sup_{x,y \in E} d(x,y)$ and for a measure μ on $\mathcal{F}$ and $\mathcal{F}$-measurable functions h_1 and h_2

$$\mathrm{Cor}_\mu (h_1, h_2) \triangleq \left| \mu(h_1 h_2) - \mu(h_1)\mu(h_2) \right|.$$

Theorem 5.6. Assume Ω compact and let f be a LIS on $(\Omega, \mathcal{F})$ that is continuous and with a unique consistent measure. Let μ be the unique probability measure in $\mathcal{G}(f)$. Then for every $\Lambda, \Delta \in \mathcal{S}_b$ such that $m_\Delta < l_\Lambda$,

$$\mathrm{Cor}_\mu (h_1, h_2) \leqslant \frac{D^2}{4} \sum_{k \leqslant m_\Delta} \delta_k^d \left(f_{[k+1, m_\Lambda]} h_1\right) \delta_k^d \left(f_{[k+1, m_\Delta]} h_2\right) \tag{5.10}$$

for all functions $h_1 \in \mathcal{B}_d(\Lambda)$ and $h_2 \in \mathcal{B}_d(]-\infty, m_\Delta])$.

Our next corollary offers a more quantitative consequence of this theorem. For a matrix $\left(A_{kj}\right)_{k,j \in \mathbb{Z}}$ with nonnegative entries, we denote

$$\left[\frac{A}{1-A}\right]_{kj} \triangleq \sum_{n \geqslant 1} \left[A^n\right]_{kj} . \tag{5.11}$$

These are well-defined sums on $[0, +\infty]$.

Corollary 5.7. Consider the hypotheses of the previous theorem and let (α_{ij}) be a d-sensitivity matrix for f.

(i) If $h_1 \in \mathcal{B}_d(\Lambda)$ and $h_2 \in \mathcal{B}_d(]-\infty, m_\Delta])$,

$$\mathrm{Cor}_\mu\,(h_1,h_2) \leqslant \frac{D^2}{4} \sum_{k \leqslant m_\Delta} \sum_{l \in \Lambda} \delta_l^d(h_1) \left[\frac{P_\Lambda \alpha}{1-P_\Lambda \alpha}\right]_{lk} \delta_k^d\left(f_{[k+1,m_\Delta]} h_2\right). \quad (5.12)$$

(ii) If $h_1 \in \mathcal{B}_d(\Lambda)$ and $h_2 \in \mathcal{B}_d(\Delta)$,

$$\mathrm{Cor}_\mu\,(h_1,h_2) \leqslant \frac{D^2}{4} \sum_{l \in \Delta} \sum_{m \in \Lambda} \delta_m^d(h_1)\, \delta_l^d(h_2)\, A_{ml}\,, \quad (5.13)$$

where

$$A_{ml} \triangleq \left[\frac{P_\Lambda \alpha}{1-P_\Lambda \alpha}\right]_{ml} + \sum_{k \leqslant m_\Delta} \left[\frac{P_\Lambda \alpha}{1-P_\Lambda \alpha}\right]_{mk} \left[\frac{P_\Lambda \alpha}{1-P_\Lambda \alpha}\right]_{lk}.$$

The following proposition is useful to estimate the different matrices appearing in this corollary.

Proposition 5.8. If (α_{ij}) is a matrix satisfying (5.3), then for each $\Lambda \in \mathcal{S}_b$

$$\left[\frac{(P_\Lambda \alpha)}{1-(P_\Lambda \alpha)}\right]_{kj} \leqslant \frac{\gamma_\Lambda}{1-\gamma_\Lambda}\, e^{-F(k,j)}. \quad (5.14)$$

6. PROOFS FOR THE GENERAL FRAMEWORK

6.1. Singleton Consistency for Chains

The fact that the objects defined by (3.1) are kernels from $\mathcal{F}_{\leqslant m_\Lambda} \times \Omega$ to the interval [0, 1] follows immediately from the properties of the kernels f_i. Their normalization is proven by induction, using the fact that

$$f_{\{i\}}(1\,|\,\cdot) = f_{\{i\}}\big(\Omega_{\leqslant i}\,\big|\,\cdot\big) = 1$$

and the inductive step

$$f_\Lambda\big(\Omega_{\leqslant m_\Lambda}\,|\,\omega\big) = f_{[l_\Lambda, m_\Lambda-1]}\Big(\big(f_{m_\Lambda}\big(\Omega_{\leqslant m_\Lambda}\big)\,\Big|\,\omega\Big) = f_{[l_\Lambda, m_\Lambda-1]}\big(1\,\big|\,\omega\big) = 1\,,$$

for $\omega \in \Omega_{\leqslant l_\Lambda}$.

Properties (a) and (b) of the definition 2.1 of LIS are an immediate consequence of similar properties of the kernels f_i. To prove consistency, we first remark that for $l \leqslant m \leqslant p$, $\omega \in \Omega$ and any $\mathcal{F}_{\leqslant p}$-measurable function h,

$$\begin{aligned}\big(f_{[l,m]}\, f_{[l,p]}\big)(h \mid \omega) &= f_{[l,m]}\Big(f_{[l,p]}(h)\,\Big|\,\omega\Big)\\ &= f_{[l,p]}(h \mid \omega)\, f_{[l,m]}(1 \mid \omega)\\ &= f_{[l,p]}(h \mid \omega)\,. \end{aligned} \tag{6.1}$$

The second equality is due to property (b) of Definition 2.1 plus the fact that $f_{[l,p]}(h \mid \cdot)$ is $\mathcal{F}_{\leqslant l-1}$-measurable. The last equality is the just proven normalization. Identity (6.1) justifies the last equality in the following string of identities, valid for $l \leqslant m < p$,

$$f_{[l,p]}\, f_{[l,m]} = f_{[l,m]}\, f_{[m+1,p]}\, f_{[l,m]} = f_{[l,m]}\, f_{[l,p]} = f_{[l,p]}\,. \tag{6.2}$$

The other equalities are simply due to definition 3.2. A similar identity is trivially true for $l \leqslant m = p$. Consistency follows for, if $\Delta \supset \Lambda$:

$$f_\Delta\, f_\Lambda = f_{[l_\Delta, l_\Lambda - 1]}\, f_{[l_\Lambda, m_\Delta]}\, f_{[l_\Lambda, m_\Lambda]} = f_{[l_\Delta, l_\Lambda - 1]}\, f_{[l_\Lambda, m_\Delta]} = f_\Delta\,.$$

We used (6.2) in the middle identity and we assumed $l_\Delta < l_\Lambda$, otherwise we revert to (6.2).

The remainder of the proof relies on the following observation valid for any measure μ on $\mathcal{F}$ and any$\Lambda \in \mathcal{S}_b$:

$$\mu\, f_i = \mu\,,\ \forall i \in \Lambda \quad \Longrightarrow \quad \mu\, f_\Lambda = \mu\,. \tag{6.3}$$

This is proven by induction on the cardinality of Λ through the identity

$$\mu\, f_\Lambda = \mu\, f_{l_\Lambda}\, f_{[l_\Lambda+1, m_\Lambda]} = \mu\, f_{[l_\Lambda+1, m_\Lambda]}\,.$$

Property (6.3) directly proves the non-trivial inclusion in (3.2). Furthermore, it yields uniqueness. Indeed, consider a LIS $(g_\Lambda)_{\Lambda \in \mathcal{S}_b}$ consistent with the family $(f_i)_{i \in \mathbb{Z}}$. By (6.3) g_Λ must be consistent with f_Λ for each $\Lambda \in \mathcal{S}_b$. But then, if $\omega \in \Omega$ and h is $\mathcal{F}_{\leqslant m_\Lambda}$-measurable

$$g_\Lambda\big(h \,\big|\, \omega\big) = g_\Lambda\Big(f_\Lambda(h)\,\Big|\,\omega\Big) = f_\Lambda\big(h \,\big|\, \omega\big)\, g_\Lambda\big(1 \,\big|\, \omega\big) = f_\Lambda\big(h \,\big|\, \omega\big)\,.$$

The second identity is a consequence of the $\mathcal{F}_{l_\Lambda - 1}$-measurability of $f_\Lambda(h|\cdot)$ plus property (b) of Definition 2.1. The last equality is the normalization of g_Λ.

6.2. Extreme Chains

We start with general results on probability kernels.

Proposition 6.1. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{F}$, π a probability kernel on $\mathcal{B} \times \Omega$ and $\mu \in \mathcal{P}(\Omega, \mathcal{F})$ such that $\mu\pi = \mu$ on $\mathcal{B}$. Then:

(i) The system

$$\mathcal{I}_\pi^{\mathcal{B}}(\mu) \triangleq \Big\{A \in \mathcal{B} : \pi(A \,|\, \cdot) = \mathbb{1}_A(\,\cdot\,)\; \mu\text{-a.s.}\Big\} \tag{6.4}$$

is a σ-algebra.

(ii) For all $\mathcal{B}$-measurable functions $h : \Omega \to [0, +\infty[$,

$$(h\mu)\,\pi = h\mu \text{ on } \mathcal{B} \quad \text{if and only if} \quad h \text{ is } \mathcal{I}_\pi^{\mathcal{B}}(\mu)\text{-measurable}\,. \tag{6.5}$$

Proof. (i) Clearly $\Omega \in \mathcal{I}_\pi^{\mathcal{B}}(\mu)$. For each $A \in \mathcal{I}_\pi^{\mathcal{B}}(\mu)$,

$$\pi(A^c \,|\, \cdot) = 1 - \pi(A \,|\, \cdot) = 1 - \mathbb{1}_A \;\; (\mu\text{-a.s.}) = \mathbb{1}_{A^c} \;\; (\mu\text{-a.s.})\,.$$

Likewise, for each sequence $(A_n)_{n \in \mathbb{N}}$ of disjoint sets in $\mathcal{I}_\pi^{\mathcal{B}}(\mu)$,

$$\pi\big(\cup A_n \,\big|\, \cdot\big) = \sum_{n \in \mathbb{N}} \pi(A_n \,|\, \cdot) = \sum_{n \in \mathbb{N}} \mathbb{1}_{A_n} \;\; (\mu\text{-a.s.}) = \mathbb{1}_{\cup A_n} \;\; (\mu\text{-a.s.})\,.$$

Finally, if $A, B \in \mathcal{I}_\pi^{\mathcal{B}}(\mu)$, then

$$\pi(A \cap B \,|\, \cdot) \leqslant \pi(A \,|\, \cdot) \wedge \pi(B \,|\, \cdot) = \mathbb{1}_A \wedge \mathbb{1}_B \;\; (\mu\text{-a.s.}) = \mathbb{1}_{A \cap B} \;\; (\mu\text{-a.s.})$$

and, by the consistency of μ with π,

$$\mu\Big(\mathbb{1}_{A \cap B} - \pi(A \cap B \,|\, \cdot)\Big) = \mu(A \cap B) - \mu\pi(A \cap B) = 0\,.$$

Thus

$$\pi(A \cap B) = \mathbb{1}_{A \cap B}\; \mu\text{-a.s.}$$

(ii) Let us assume that $(h\mu)\pi = h\mu$ on $\mathcal{B}$. To prove necessity it suffices to show that $\{h \geqslant c\} \in \mathcal{I}_{\pi}^{\mathcal{B}}(\mu)$, for all $c > 0$. Let us fix some $c > 0$ and denote $g = 1\!\!1_{h \geqslant c}$. We have

$$\mu\Big((1-g)\,h\,\pi(g)\Big) = (h\,\mu)\big(\pi(g)\big) - \mu\big(g\,h\,\pi(g)\big) = (h\,\mu)(g) - \mu\big(g\,h\,\pi(g)\big)$$
$$= \mu\Big(g\,h\,\big(1-\pi(g)\big)\Big).$$

But $gh \geqslant cg$ and $1-\pi(g) \geqslant 0$, hence

$$\mu\Big((1-g)\,h\,\pi(g)\Big) \geqslant c\,\mu\Big(g\big(1-\pi(g)\big)\Big) = c\,\mu\big(\pi(g)\big) - c\,\mu\big(g\,\pi(g)\big)$$
$$= c\,\mu\Big((1-g)\,\pi(g)\Big).$$

We obtain that $\mu\left(1\!\!1_{\{h<c\}}\,(h-c)\,\pi(g)\right) \geqslant 0$, which implies $1\!\!1_{\{h<c\}}\pi(g) = 0$ μ-a.s. Therefore,

$$\pi(g) = g\,\pi(g) + 1\!\!1_{\{h<c\}}\,\pi(g) \leqslant g \;\; \mu\text{-a.s.}$$

Furthermore, $\mu\big(g - \pi(g)\big) = 0$ by the consistency of μ with π. This fact, together with the previous inequality, allows us to conclude that $\pi(g) = g$ μ-a.s., that is $\{h \geqslant c\} \in \mathcal{I}_{\pi}^{\mathcal{B}}(\mu)$.

Conversely, assume that h is $\mathcal{I}_{\pi}^{\mathcal{B}}(\mu)$-measurable. By the standard machinery of measure theory sufficiency follows if we show for all $A \in \mathcal{I}_{\pi}^{\mathcal{B}}(\mu)$ that $(1\!\!1_A\,\mu)\pi = 1\!\!1_A\mu$ on $\mathcal{B}$. If $B \in \mathcal{B}$,

$$(1\!\!1_A\,\mu)\,\pi(B) = (1\!\!1_A\,\mu)\,\pi(A \cap B) + (1\!\!1_A\,\mu)\,\pi(B \setminus A)$$
$$\leqslant \mu\pi(A \cap B) + (1\!\!1_A\,\mu)\,\pi\left(A^c\right).$$

The consistency of μ with π implies that the second term of the last line is zero. Thus we have proved that

$$(1\!\!1_A\mu)\,\pi(B) \leqslant (1\!\!1_A\mu)(B). \tag{6.6}$$

By the same token,

$$(1\!\!1_A\mu)\,\pi\left(B^c\right) \leqslant (1\!\!1_A\mu)\left(B^c\right). \tag{6.7}$$

But the consistency of μ with π implies that the sum of the LHS of (6.6) and (6.7) equals the sum of the corresponding RHS, namely $\mu(A)$. We conclude that $(1\!\!1_A\mu)\,\pi(B) = (1\!\!1_A\mu)(B)$. ∎

Corollary 6.2. Let Π be a non-empty set of probability kernels π defined on $\mathcal{F}_\pi \times \Omega$, where $\mathcal{F}_\pi$ is a sub-σ-algebra of $\mathcal{F}$. Let us denote

$$\mathcal{G}(\Pi) = \Big\{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \mu\,\pi = \mu \text{ on } \mathcal{F}_\pi \text{ for all } \pi \in \Pi\Big\} \tag{6.8}$$

and for each $\mu \in \mathcal{G}(\Pi)$,

$$\mathcal{I}_\Pi(\mu) = \bigcap_{\pi \in \Pi} \mathcal{I}_\pi^{\mathcal{F}_\pi}(\mu) \tag{6.9}$$

be the σ-algebra of all μ-almost surely Π-invariant sets. Then μ is trivial on $\mathcal{I}_\Pi(\mu)$ if μ is extreme in $\mathcal{G}(\Pi)$.

Proof. Suppose μ is not trivial on $\mathcal{I}_\Pi(\mu)$ and take $A \in \mathcal{I}_\Pi(\mu)$ such that $0 < \mu(A) < 1$. The measures

$$\nu = \mu(\,\cdot \mid A) \triangleq h\,\mu \quad \text{with } h = \frac{1\!\!1_A}{\mu(A)}$$

and

$$\nu' = \mu(\,\cdot \mid A^c) \triangleq h'\mu \quad \text{with } h' = \frac{1\!\!1_{A^c}}{\mu(A^c)}$$

satisfy $\nu \neq \nu'$ and $\mu = \mu(A)\,\nu + \mu(A^c)\,\nu'$. The functions h and h' are $\mathcal{I}_\pi^{\mathcal{F}_\pi}(\mu)$-measurable, for all $\pi \in \Pi$. Thus, (ii) of Proposition 6.1 implies that ν, $\nu' \in \mathcal{G}(\Pi)$, a fact that contradicts the extremality of μ. ∎

Lemma 6.3. Let f be a LIS defined on $(\Omega, \mathcal{F})$ and $\mu \in \mathcal{G}(f)$. Let us denote by $\mathcal{F}_{-\infty}^{\mu}$ the μ-completion of $\mathcal{F}_{-\infty}$. Then

$$\bigcap_{n \geqslant 0} \mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leqslant k}}(\mu) = \mathcal{F}_{-\infty}^{\mu} \tag{6.10}$$

for each $k \in \mathbb{Z}$ and

$$\bigcap_{\Lambda \in \mathcal{S}_b} \mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leqslant m_\Lambda}}(\mu) = \mathcal{F}_{-\infty}^{\mu}\,. \tag{6.11}$$

Proof. Identity (6.10) follows from the observation that for each $B\in\bigcap_n \mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leqslant k}}(\mu)$ the set $A\triangleq\bigcap_n\left\{f_{[k-n,k]}(B\,|\,\cdot\,)=1\right\}$ satisfies $A=B$ μ-a.s. and $A\in\mathcal{F}_{-\infty}$. Equality (6.11) is a consequence of (6.10) because

$$\bigcap_{\Lambda\in\mathcal{S}_b}\mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leqslant m_\Lambda}}(\mu)=\bigcap_{k\in\mathbb{Z}}\bigcap_{n\geqslant 0}\mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leqslant k}}(\mu). \quad \blacksquare$$

Proof of Theorem 3.2. (a) It is immediate.

(b) ($\Rightarrow$) The implication follows readily from Corollary 6.2 and the fact that, by (6.11), $\bigcap_{\Lambda\in\mathcal{S}_b}\mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leqslant m_\Lambda}}(\mu)$ is μ-trivial if and only if μ is trivial on $\mathcal{F}_{-\infty}$.

(c) ($\Rightarrow$) Let $\mu,\,\nu\in\mathcal{G}(f)$ such that $\nu\ll\mu$. There exists a $\mathcal{F}$-measurable non-negative function g such that

$$\nu=g\,\mu\,.$$

Let us consider, for each $k\in\mathbb{Z}\,\mu_k\triangleq\mu\big|_{\mathcal{F}_{\leqslant k}}$ and $\nu_k\triangleq\nu\big|_{\mathcal{F}_{\leqslant k}}$. As in particular $\nu_k\ll\mu_k$ on $\mathcal{F}_{\leqslant k}$, there exists $g_k\geqslant 0$, $\mathcal{F}_{\leqslant k}$-measurable, satisfying $\nu_k=g_k\,\mu_k$ on $\mathcal{F}_{\leqslant k}$). All we have to prove is that

$$g_k \text{ is } \mathcal{F}_{-\infty}^{\mu}\text{-measurable} \quad \forall\, k\in\mathbb{Z}\,. \tag{6.12}$$

Indeed, by the reverse martingale theorem $g_k=g$ μ-a.s. Therefore, g inherits the $\mathcal{F}_{-\infty}^{\mu}$-measurability and, thus, it is μ-a.s. equal to a $\mathcal{F}_{-\infty}$-measurable function.

To prove (6.12) we observe that since $\nu\in\mathcal{G}(f)$,

$$g_k\,\mu_k\,f_{[k-n,k]}=g_k\,\mu_k$$

on $\mathcal{F}_{\leqslant k}$ for all $n\in\mathbb{N}$. As g_k is $\mathcal{F}_{\leqslant k}$-measurable, we conclude from Proposition 6.4 that g_k is $\bigcap_n\mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leqslant k}}(\mu)$-measurable. Its, $\mathcal{F}_{-\infty}^{\mu}$-measurability follows, hence, from (6.10).

(b) ($\Leftarrow$) Assume μ is a trivial measure on $\mathcal{F}_{-\infty}$ and suppose that there exist $s{:}0<s<1$ and $\nu,\nu'\in\mathcal{G}(f)$ such that $\mu=s\,\nu+(1-s)\,\nu'$. As $\nu,\nu'\ll\mu$, by (c) ($\Rightarrow$) there exist $\mathcal{F}_{-\infty}$-measurable functions $h,h'\geqslant 0$ such that $\nu=h\mu$ and $\nu'=h'\mu$. But the triviality of μ on $\mathcal{F}_{-\infty}$ implies that $h=h'=1$ μ-a.s. Thus $\mu=\nu=\nu'$.

(c) ($\Leftarrow$) This is an immediate consequence of Proposition 6.1 plus the fact that h is $\mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leqslant m_\Lambda}}(\mu)$-measurable for all $\Lambda\in\mathcal{S}_b$.

(d) Let $\mu, \nu \in \mathcal{G}(f)$ such that $\mu = \nu$ on $\mathcal{F}_{-\infty}$. Consider $\widetilde{\mu} \triangleq \frac{1}{2}\mu + \frac{1}{2}\nu \in \mathcal{G}(f)$. Since $\mu \ll \widetilde{\mu}$ and $\nu \ll \widetilde{\mu}$, assertion (b) implies that $\mu = f\widetilde{\mu}$ and $\nu = g\widetilde{\mu}$ for $\mathcal{F}_{-\infty}$-measurable functions f and g. But $\mu = \nu = \widetilde{\mu}$ on $\mathcal{F}_{-\infty}$, so $f = g$ μ-a.s. and therefore $\mu = \nu$.

(e) It is an immediate consequence of (b) and (d). ∎

6.3. Triviality and Short-Range Correlations

The proofs involve standard arguments. We include them for completeness.

Proof of Theorem 3.3. (a) ⇒ (c) Let $A \in \mathcal{F}$ and $k \in \mathbb{Z}$. Since $\mathcal{F}_{-\infty} = \bigcap_{n \geqslant 1} \mathcal{F}_{\leqslant k-n}$, the reverse martingale theorem yields

$$\mu\big(A \,\big|\, \mathcal{F}_{\leqslant k-n}\big) \xrightarrow[n\to+\infty]{L^1(\mu)} \mu\big(A \,\big|\, \mathcal{F}_{-\infty}\big). \tag{6.13}$$

The assumed triviality of μ on $\mathcal{F}_{-\infty}$ implies that $\mu\,(A \,|\, \mathcal{F}_{-\infty}) = \mu(A)$ μ-a.s. We deduce that for each $\varepsilon > 0$, there exists $\Delta \in \mathcal{S}_b$ such that

$$\mu\Big(\big|\mu\left(A \,|\, \mathcal{F}_{\Delta_-}\right) - \mu(A)\big|\Big) < \varepsilon\,. \tag{6.14}$$

Hence, for all $\Lambda \in \mathcal{S}_b : \Lambda \supset \Delta$,

$$\begin{aligned}
\sup_{B \in \mathcal{F}_{\Lambda_-}} \Big|\mu(A \cap B) - \mu(A)\,\mu(B)\Big| &\leqslant \sup_{B \in \mathcal{F}_{\Delta_-}} \Big|\mu(A \cap B) - \mu(A)\,\mu(B)\Big| \\
&= \Big|\mu\Big(\big[\mu\left(A \,|\, \mathcal{F}_{\Delta_-}\right) - \mu(A)\big]\, \mathbb{1}_B\Big)\Big| \\
&\leqslant \mu\Big(\big|\mu\left(A \,|\, \mathcal{F}_{\Delta_-}\right) - \mu(A)\big|\Big) \\
&< \varepsilon.
\end{aligned}$$

(b) ⇒ (a) Fix $B \in \mathcal{F}_{-\infty}$ and consider $\mathcal{D} \triangleq \{A \in \mathcal{F} : \mu(A \cap B) = \mu(A)\mu(B)\}$. It is straightforward to see that $\mathcal{D}$ is a λ-system. By assumption $\mathcal{D}$ contains all cylinder events, so $\mathcal{D} = \mathcal{F}$ [Dynkin's π-λ theorem]. In particular $B \in \mathcal{D}$, thus $\mu(B) = (\mu(B))^2$ and thereby $\mu(B) = 0$ or 1. ∎

Proof of Theorem 3.4. (a) Let h be a bounded local function on Ω. As μ is consistent with f, $f_{\Lambda_n} h$ coincides with $\mu\left(h \,|\, \mathcal{F}_{(\Lambda_n)_-}\right)$, μ-a.s., for n sufficiently large. Therefore, by the reverse martingale convergence theorem we conclude that

$$f_{\Lambda_n} h \xrightarrow[n\to+\infty]{} \mu\,(h \,|\, \mathcal{F}_{-\infty}) \quad \mu\text{-a.s.}$$

This implies assertion (a) because μ is trivial on $\mathcal{F}_{-\infty}$.

(b) It is a consequence of assertion (a) and the fact that if Ω is compact and metric, the space of local continuous functions on Ω contains a countable subset which is dense with respect to the uniform-norm. ∎

6.4. Ergodicity

We need a well known result from ergodic theory. See, for instance, ref. 9, Theorem 14.5, for a proof.

Theorem 6.4. (a) A probability measure $\mu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ is extreme in $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ if and only if μ is ergodic.

(b) Let $\mu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ and $\nu \in \mathcal{P}(\Omega, \mathcal{F})$ such that $\nu \ll \mu$, then $\nu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ if and only if $\exists\, h \geqslant 0$, $\mathcal{I}$-measurable: $\nu = h\mu$.

Lemma 6.5. Let $\mu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$, then $\mathcal{I} \subset \mathcal{F}_{-\infty}$ μ-*a.s.* More precisely, for each $A \in \mathcal{I}$ there exists $B \in \mathcal{F}_{-\infty}$ such that $\mu(A \Delta B) = 0$.

Proof. Let $A \in \mathcal{I}$ and $(B_n)_{n \geqslant 1}$ be a sequence of cylinder sets such that $\mu(A \Delta B_n) \leqslant 2^{-n}$ for all $n \geqslant 1$. Since $\mu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$, we have that

$$\mu\left(A \Delta \tau^i B_n\right) = \mu\left(\tau^i A \Delta \tau^i B_n\right) = \mu\left(A \Delta B_n\right) \leqslant 2^{-n}$$

for each $i \in \mathbb{N}$ (τ^i is the ith-iterate of τ). Consider $\Lambda_n \uparrow \mathbb{Z}$ such that $B_n \in \mathcal{F}_{\Lambda_n}$. For each $n \geqslant 1$ we choose $i(n) \geqslant 0$ such that $\Lambda_n \cap (\Lambda_n - i(n)) = \emptyset$. Each set $C_n \triangleq \tau^{i(n)} B_n$ belongs to $\mathcal{F}_{(\Lambda_n)_-}$ and satisfies $\mu\left(A \Delta C_n\right) \leqslant 2^{-n}$. Therefore, the set $C \triangleq \bigcap_{m \geqslant 1} \bigcup_{n \geqslant m} C_n$ belongs to $\mathcal{F}_{-\infty}$ and satisfies

$$\mu(A \Delta C) \leqslant \mu\left(\bigcap_{m \geqslant 1} \bigcup_{n \geqslant m} A \Delta C_n\right) \leqslant \lim_{m \to +\infty} \sum_{n \geqslant m} 2^{-n} = 0. \quad ∎$$

Proof of Theorem 3.5. (a) Let us consider the probability kernel T on $\mathcal{F} \times \Omega$ defined by

$$T(A \mid \omega) = 1\!\!1_A(\tau\omega)$$

for every $A \in \mathcal{F}$ and every $\omega \in \Omega$.

To prove necessity we introduce

$$\mathcal{K}(\mu) \triangleq \left(\bigcap_{\Lambda \in \mathcal{S}_b} \mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leqslant m_\Lambda}}(\mu) \right) \bigcap \mathcal{I}_T^{\mathcal{F}}(\mu)\,.$$

By (6.11) and Lemma 6.5, $\mathcal{K}(\mu)$ is the μ-completion of $\mathcal{I}$. Therefore Corollary 6.2 implies that each μ extreme in $\mathcal{G}_{\rm inv}(f)$ is trivial on $\mathcal{I}$.

For the sufficiency, suppose that μ is trivial on $\mathcal{I}$ and consider a decomposition $\mu = s\,\nu + (1-s)\,\nu'$ with $0 < s < 1$ and $\nu,\ \nu' \in \mathcal{G}_{\rm inv}(f)$. Then there exist $\mathcal{F}$-measurable $h,\ h' \geqslant 0$ such that $\nu = h\mu$ and $\nu' = h'\mu$. Since $\mu,\ \nu,\ \nu' \in \mathcal{P}_{\rm inv}(\Omega, \mathcal{F})$, Proposition 6.1 applied to $\mathcal{I}_T^{\mathcal{F}}(\mu)$ implies that $h,\ h'$ are measurable with respect to the μ-completion of $\mathcal{I}$. Hence the triviality of μ on $\mathcal{I}$ assure that $h = h' = 1$ μ-a.s. Thus $\mu = \nu = \nu'$.

(b). Theorem 6.4 (b) implies that there exists $h \geqslant 0$, $\mathcal{I}$-measurable such that $\nu = h\mu$. By Lemma 6.5 h is $\mathcal{F}_{-\infty}$-measurable, so Theorem 3.2 (b) implies that $\nu \in \mathcal{G}(f)$. Therefore $\nu \in \mathcal{G}_{\rm inv}(f)$.

(c) It is an immediate consequence of (b). ∎

7. PROOFS ON UNIQUENESS

7.1. One-sided Boundary-Uniformity

Lemma 7.1. If uniqueness condition (4.1) is satisfied, then $\nu \geqslant c\,\mu,\ \forall\,\mu, \nu \in \mathcal{G}(f)$.

Proof. Let A be a cylinder set and n an integer such that (4.1) holds. If μ and ν are consistent with f,

$$\begin{aligned} \nu(A) &= \iint f_{[-n,m]}(A \mid \xi)\, \mu(d\eta)\, \nu(d\xi) \\ &\geqslant c \iint f_{[-n,m]}(A \mid \eta)\, \mu(d\eta)\, \nu(d\xi) \\ &= c\,\mu(A). \quad ∎ \end{aligned}$$

Proof of Theorem 4.1. We shall prove that every element of $\mathcal{G}(f)$ is extreme. Let $\mu \in \mathcal{G}(f)$ and $B \in \mathcal{F}_{-\infty}$ such that $\mu(B) > 0$. Define

$$\nu \triangleq \mu(\cdot \mid B) = \frac{1\!\!1_B}{\mu(B)}\,\mu\,.$$

By Theorem 3.2 (c), $\nu \in \mathcal{G}(f)$. By the preceding lemma $0 = \nu(B^c) \geqslant c\,\mu(B^c)$, so $\mu(B) = 1$. ∎

Proof of Proposition 4.2. Call $m(f)$ the infimum (4.3) and $V(f)$ the supremum (4.4). Through an elementary logarithmic inequality we have that for each $i, j \in \mathbb{Z}$ with $i > j$ and each $\xi, \eta \in \Omega_{\leqslant i}$ with $\xi_j^i = \eta_j^i$,

$$\frac{f_{\{i\}}\big(\xi_i \,\big|\, \xi_{-\infty}^{i-1}\big)}{f_{\{i\}}\big(\eta_i \,\big|\, \eta_{-\infty}^{i-1}\big)} \geqslant \exp\Big(-\frac{\operatorname{var}_j(f_{\{i\}})}{m(f)}\Big). \tag{7.1}$$

Applying the factorization (3.3) we conclude that for each $n, m \in \mathbb{Z}$ with $n < m$ and each $\xi, \eta \in \Omega_{\leqslant m}$ with $\xi_n^m = \eta_n^m$,

$$\frac{f_{[n,m]}\big(\xi_n^m \,\big|\, \xi_{-\infty}^{n-1}\big)}{f_{[n,m]}\big(\eta_n^m \,\big|\, \eta_{-\infty}^{n-1}\big)} \geqslant e^{-V(f)/m(f)}. \quad \blacksquare \tag{7.2}$$

7.2. Dobrushin Uniqueness

The following bound is the basic tool of the theory.

Lemma 7.2 (Multisite dusting lemma). Let $V \in \mathcal{S}_b$, f_V a probability kernel on $\mathcal{F}_{\leqslant m_V} \times \Omega$ and α^V is a d-sensitivity estimator for f. Then,

$$\delta_j^d\,(f_V h) \quad \begin{cases} = 0 & \text{if } j \in V \\ \leqslant \delta_j^d(h) + \sum\limits_{k \in V} \delta_k^d(h)\, \alpha_{kj}^V & \text{if } j \in V_-, \end{cases} \tag{7.3}$$

for every continuous function h on $V \cup V_-$.

Remark 7.3. The name of the lemma comes from a picturesque interpretation due to Michael Aizenman reported in ref. 21: If the oscillations are interpreted as "dust" and the averages f_V as applications of a (multisite) "duster", the lemma says that no dust remains in V after dusting the sites there [first line of (7.3)], but the dust has been spread over the remaining sites [second line of (7.3)]. The estimators give the fraction blown from site to site. In this picture, Dobrushin condition (4.9) means that some dust stays in the duster, a fact that allows for an eventual total cleaning.

Proof. The first line in (7.3) just expresses the fact that the average $f_V h$ is $\mathcal{F}_{V_-}$-measurable. The second line shows two contributions: The first one due to the direct dependence of h on the configuration at the site j, and the second to the sensibility of the f_V-averages to the configuration on the past instant j. To separate both contributions we introduce a family of auxiliary functions $h_{V,\omega}(\sigma_V) \triangleq h\left(\omega_{V_-}\, \sigma_V\right)$ for each $\omega \in \Omega$ ("freezing" at ω). For $j \in V_-$ and $\xi,\ \eta \in \Omega_{V_-}$ such that $\xi \overset{\neq j}{=} \eta$, we have

$$\begin{aligned} &\left| f_V(h \mid \xi) - f_V(h \mid \eta) \right| \\ &\quad \leqslant \left| \overset{\circ}{f}_V \left(h_{V,\xi} - h_{V,\eta} \mid \xi \right) \right| + \left| \overset{\circ}{f}_V \left(h_{V,\eta} \mid \xi \right) - \overset{\circ}{f}_V \left(h_{V,\eta} \mid \eta \right) \right|. \end{aligned} \tag{7.4}$$

If we divide throughout by $d(\xi_j, \eta_j)$ and use the estimator bound (4.8) we obtain, upon taking the necessary suprema, the second line in (7.3). ∎

We now fix a partition $\mathcal{P}$ of $\mathbb{Z}$ into finite intervals and denote, for each $\Lambda \subset \mathcal{S}_b$,

$$\Lambda^* = \bigcup \left\{ V \in \mathcal{P} : \Lambda \cap V \neq \emptyset \right\}.$$

Let $n(\Lambda)$ denote the number of elements of $\mathcal{P}$ forming Λ^*.

Proposition 7.4. Consider a LIS f and d-sensitivity estimators α^V for f_V for each $V \in \mathcal{P}$.

(i) For every $j \in \Lambda^*_-$ and $h \in \mathcal{B}_d(\Lambda^* \cup \Lambda^*_-)$,

$$\delta_j^d\,(f_{\Lambda^*} h) \leqslant \delta_j^d(h) + \sum_{k \in \Lambda^*} \delta_k^d(h) \left[\sum_{l=1}^{n(\Lambda)} (P_{\Lambda^*}\alpha)^l \right]_{kj}. \tag{7.5}$$

(ii) If Dobrushin condition (4.9) is satisfied, then for every $j \in \Lambda^*_-$ and $h \in \mathcal{B}_d(\Lambda^*)$,

$$\delta_j^d\,(f_{\Lambda^*} h) \leqslant \sum_{k \in \Lambda^*} \delta_k^d(h) \left[\frac{P_{\Lambda^*}\alpha}{1 - P_{\Lambda^*}\alpha} \right]_{kj}. \tag{7.6}$$

Proof. We only need to prove (7.5). Inequality (7.6) is then obtained by bounding the sum in the RHS of (7.5) by the limit $n(\Lambda) \to \infty$, which is finite under Dobrushin condition.

We proceed by induction on $n(\Lambda)$. The case $n(\Lambda) = 1$ is just the multisite dusting lemma. Suppose the inequality valid for all Λ with $n(\Lambda) =$

n. Consider Δ such that $\Delta^* = \bigcup_{i=1}^{n+1} V_i$, where the $V_i \in \mathcal{P}$, $i = 1, \ldots, n+1$ are labeled so that $m_{V_i} = l_{V_{i+1}-1}$. Denote $\Lambda^* = \bigcup_{i=1}^{n} V_i$. Let $j \in \Delta^*_-$ and $h \in \mathcal{B}_d(\Delta^* \cup \Delta^*_-)$. By the factorization property (3.4) of the LIS, $\delta_j^d\,(f_{\Delta^*}h) = \delta_j^d\,\big(f_{\Lambda^*} f_{V_{n+1}} h\big)$. Therefore, by the inductive hypothesis,

$$\delta_j^d\,(f_{\Delta^*}h) \leqslant \delta_j^d(f_{V_{n+1}}h) + \sum_{k \in \Lambda^*} \delta_k^d\,\big(f_{V_{n+1}}h\big) \left[\sum_{l=1}^{n} (P_{\Lambda^*}\alpha)^l\right]_{kj},$$

and the multisite dusting Lemma 7.2 yields

$$\begin{aligned}\delta_j^d\,(f_{\Delta^*}h) \leqslant{}& \delta_j^d(h) + \sum_{m \in V_{n+1}} \delta_m^d(h)\,\big[P_{V_{n+1}}\alpha\big]_{mj} \\ &+ \sum_{k \in \Lambda^*} \left(\delta_k^d(h) + \sum_{m \in V_{n+1}} \delta_m^d(h)\,\big[P_{V_{n+1}}\alpha\big]_{mk}\right) \left[\sum_{l=1}^{n} (P_{\Lambda^*}\alpha)^l\right]_{kj}.\end{aligned}$$

We now observe that, given the restrictions in the sites being summed over, we can replace in the RHS P_{Λ^*} and $P_{V_{n+1}}$ by P_{Δ^*}. Furthermore, for $m \in V_{n+1}$, $l \in \mathbb{N}$,

$$\sum_{i=1}^{n} \sum_{k \in V_i} [P_{\Delta^*}\alpha]_{mk} \left[(P_{\Delta^*}\alpha)^l\right]_{kj} = \left[(P_{\Delta^*}\alpha)^{l+1}\right]_{mj}.$$

The last two displays imply that

$$\delta_j^d\,(f_{\Delta^*}h) \leqslant \delta_j^d(h) + \sum_{k \in \Delta^*} \delta_k^d(h) \left[\sum_{l=1}^{n+1} \big(P_{\Delta_*}\alpha\big)^l\right]_{kj}. \quad \blacksquare$$

Proof of Theorem 4.6. Let us label the elements of the partition so that $\mathcal{P} = \{V_i : i \in \mathbb{Z}\}$ and $m_{V_i} = l_{V_{i+1}-1}$, $i \in \mathbb{Z}$. Let us denote $V_{m-i}^n = \bigcup_{l=m-i}^{n} V_l$ for every integer n, m, i with $m - i \leqslant n$. Let μ, $\nu \in \mathcal{G}(f)$ and consider a local function h of d-bounded variations. Pick $m, n \in \mathbb{Z}$ such that $h \in \mathcal{B}_d(V_m^n)$. The consistency of both μ an ν with $f_{V_{m-i}^n}$, for an integer $i > 0$, imply

$$\big|\nu(h) - \mu(h)\big| \leqslant \iint \Big|f_{V_{m-i}^n}\,(h \,|\, d\xi) - f_{V_{m-i}^n}\,(h \,|\, d\eta)\Big|\;\nu(d\xi)\,\mu(d\eta).$$

Therefore, by the continuity of f and (7.6),

$$\begin{aligned}\left|\nu(h)-\mu(h)\right| &\leqslant \sum_{j\in(V_{m-i})_-} \delta_j^d\left(f_{V_{m-i}^n}h\right)\iint d(\xi_j,\eta_j)\,\nu(d\xi)\,\mu(d\eta)\\ &\leqslant D\sum_{k\in\Lambda}\delta_k^d(h)\sum_{j\in(V_{m-i})_-}\left[\frac{P_\Lambda\alpha}{1-P_\Lambda\alpha}\right]_{kj}.\end{aligned}$$

Under condition (4.9) the series on the RHS is summable, hence the bound converges to zero as $i\to\infty$. ∎

8. PROOFS ON LOSS OF MEMORY AND MIXING

Proof of Theorem 5.2 and Proposition 5.8. Part (i) of Theorem 5.2 is just (7.5). The triangular property of F implies that for every natural n and every $k\in\Lambda^*$,

$$\begin{aligned}\left[(P_{\Lambda^*}\alpha)^n\right]_{kj}e^{F(k,j)} &= \sum_{i_1,\dots,i_{n-1}\in\Lambda^*}\alpha_{ki_1}\,\alpha_{i_1i_2}\dots\alpha_{i_{n-1}j}\,e^{F(k,j)}\\ &\leqslant \sum_{i_1,\dots,i_{n-1}\in\Lambda^*}\alpha_{ki_1}\,e^{F(k,i_1)}\,\alpha_{i_1i_2}\,e^{F(i_1,i_2)}\dots\alpha_{i_{n-1}j}\,e^{F(i_{n-1},j)}.\end{aligned}$$

Therefore, applying the definition (5.3) we obtain

$$\left[(P_{\Lambda^*}\alpha)^n\right]_{kj}\leqslant\gamma_{\Lambda^*}^n\,e^{-F(k,j)}. \tag{8.1}$$

This yields (5.14) upon summation over n. Combining (5.14) with (7.6), we obtain (5.4). ∎

Proof of Theorem 5.5. Fix $\Lambda\in\mathcal{S}_b$ and $h\in\mathcal{B}_d(\Lambda)$. Using the consistency of μ and $\widetilde{\mu}$ respectively with f and $\widetilde{f}$, we have that, for each $n\in\mathbb{N}$,

$$\begin{aligned}\left|\mu(h)-\widetilde{\mu}(h)\right| &\leqslant \left|\mu\left(f_{[m_\Lambda-n,m_\Lambda]}h\right)-\widetilde{\mu}\left(f_{[m_\Lambda-n,m_\Lambda]}h\right)\right|\\ &\quad+\left|\widetilde{\mu}\left(f_{[m_\Lambda-n,m_\Lambda]}h\right)-\widetilde{\mu}\left(\widetilde{f}_{[m_\Lambda-n,m_\Lambda]}h\right)\right|.\end{aligned} \tag{8.2}$$

We estimate separately each term on the right as n tends to infinity. The compactness of Ω implies that $f_{[m_\Lambda-n,m_\Lambda]}(h\,|\,\omega)\to\mu(h)$ for each $\omega\in$

Ω as $n \to \infty$ (see Remark 4.4). Therefore, by dominated convergence (h is continuous, hence bounded)

$$\left|\mu\left(f_{[m_\Lambda - n, m_\Lambda]}h\right) - \widetilde{\mu}\left(f_{[m_\Lambda - n, m_\Lambda]}h\right)\right| \xrightarrow[n\to\infty]{} 0. \tag{8.3}$$

To bound the last term in (8.2) we telescope using the factorization property (3.4) for LIS:

$$\left|\widetilde{\mu}\big(f_{[m_\Lambda - n, m_\Lambda]}h\big) - \widetilde{\mu}\big(\widetilde{f}_{[m_\Lambda - n, m_\Lambda]}h\big)\right| \leqslant \left|\mu\big(f_{\{m_\Lambda\}}h\big) - \widetilde{\mu}\big(\widetilde{f}_{\{m_\Lambda\}}h\big)\right| + \sum_{k=m_\Lambda - n}^{m_\Lambda - 1}\left|\widetilde{\mu}\big(f_{[k, m_\Lambda]}h\big) - \widetilde{\mu}\big(\widetilde{f}_{\{k\}}f_{[k+1, m_\Lambda]}h\big)\right|. \tag{8.4}$$

The definition (4.10)/(4.11) of the VKR distance, implies that

$$\left|(f_k\, g)(\omega) - (\widetilde{f}_k\, g)(\omega)\right| \leqslant \delta_k^d(g)\left\|\overset{\circ}{f}_{\{k\}}(\cdot\,|\,\omega) - \overset{\circ}{\widetilde{f}}_{\{k\}}(\cdot\,|\,\omega)\right\|_d,$$

for all $k \in \mathbb{Z}$, $\omega \in \Omega_{-\infty}^{k-1}$ and $g \in \mathcal{B}_d(]-\infty, k])$. Hypothesis (5.8) implies

$$\left|\widetilde{\mu}\big(f_{\{k\}}\, g - \widetilde{f}_{\{k\}}\, g\big)\right| \leqslant \widetilde{\mu}\,(b_k)\,\delta_k^d(g). \tag{8.5}$$

Combining (8.4) and (8.5) we obtain

$$\left|\widetilde{\mu}\big(f_{[m_\Lambda - n, m_\Lambda]}h\big) - \widetilde{\mu}\big(\widetilde{f}_{[m_\Lambda - n, m_\Lambda]}h\big)\right| = \sum_{k=m_\Lambda - n}^{m_\Lambda - 1}\widetilde{\mu}\,(b_k)\;\delta_k^d\big(f_{[k+1, m_\Lambda]}\big) + \widetilde{\mu}\,(b_i)\;\delta_i^d(h). \tag{8.6}$$

To obtain (5.9) we insert this bound in (8.2), let n tend to infinity and use (8.3). ∎

Proof of Theorem 5.6. Fix $\Lambda, \Delta \in \mathcal{S}_b$ with $m_\Delta < l_\Lambda$, $h_1 \in \mathcal{B}_d(\Lambda)$ and $h_2 \in \mathcal{B}_d(\Delta)$. Without loss of generality, we can suppose that $h_2 \geqslant 0$, $h_2 \neq 0$ and $\mu(h_2) = 1$ since both sides of (5.10) are invariant under adding a constant to h_2 and both multiply in the same way if h_2 is multiplied by a positive constant. We then can write

$$\mathrm{Cor}_\mu(h_1, h_2) = \left|\nu(h_1) - \mu(h_1)\right| \tag{8.7}$$

where ν is the probability measure defined by

$$\nu = h_2\,\mu\,. \tag{8.8}$$

1st stage: We construct a LIS $\widetilde{f}$ for ν on $]-\infty, m_\Lambda]$. For every $k \in\,]-\infty, m_\Lambda]$, let us define

$$\widetilde{f}_k = g_k\, f_{\{k\}} \tag{8.9}$$

with

$$g_k = \begin{cases} 1 & \text{if } k \in [m_\Delta + 1, m_\Lambda] \\ \dfrac{f_{[k+1,m_\Delta]}\big(h_2 \,\big|\, \cdot\big)}{f_{[k,m_\Delta]}\big(h_2 \,\big|\, \cdot\big)}\, 1\!\!1_A & \text{if } k \in]-\infty, m_\Delta]\,, \end{cases} \tag{8.10}$$

where

$$A = \Big\{\omega \in \Omega : f_{[k,m_\Delta]}\big(h_2 \,\big|\, \omega\big) > 0\Big\}\,.$$

It is clear that the kernels $\widetilde{f}_k$ satisfy the hypotheses of Theorem (3.1), hence they uniquely define a LIS $\widetilde{f}$ on $]-\infty, m_\Lambda]$. The same theorem shows that the consistency of ν with each $\widetilde{f}_k$, $k \in]-\infty, m_\Lambda]$ is all that has to be checked in order to prove that ν is consistent with $\widetilde{f}$.

If $k \in [m_\Delta + 1, m_\Lambda]$, this consistency is a consequence of the following sequence of identities, valid for every $h \in \mathcal{F}_{\leqslant k}$:

$$\nu\big(\widetilde{f}_k(h)\big) = \mu\big(h_2\, f_{\{k\}}(h)\big) = \mu\big(f_{\{k\}}(h_2\, h)\big) = \mu(h_2\, h) = \nu(h)\,. \tag{8.11}$$

The third inequality is due to the $\mathcal{F}_{\leqslant k-1}$ measurability of h_2 and the fourth one to consistency.

For $k \in]-\infty, m_\Delta]$ we observe that for $h \in \mathcal{F}_{\leqslant k}$,

$$\begin{aligned} \nu\big(\widetilde{f}_k(h)\big) &= \mu\big(h_2\, f_{\{k\}}(g_k\, h)\big) \\ &= \mu\Big(f_{[k,m_\Delta]}\big[h_2\, f_{\{k\}}(g_k\, h)\big]\Big) \\ &= \mu\Big(f_{\{k\}}(g_k\, h)\, f_{[k,m_\Delta]}(h_2)\Big)\,, \end{aligned}$$

the next-to-last inequality being a consequence of the consistency of μ with f and the last one of the $\mathcal{F}_{\leqslant k-1}$-measurability of $f_{\{k\}}(g_k\, h \,|\, \cdot)$. Upon inserting the definition of g_k [second line in (8.10)] we see that there is a term $f_{[k,m_\Delta]}$ in the denominator that can be pulled to the left because of its $\mathcal{F}_{\leqslant k-1}$-measurability. This produces a cancellation with an analogous term in the numerator. We thus obtain

$$\begin{aligned}\nu\big(\widetilde{f}_k(h)\big) &= \mu\Big(f_{\{k\}}\big[h\, f_{[k,m_\Delta]}(h_2)\big]\Big) = \mu\Big(f_{\{k\}}\big[f_{[k,m_\Delta]}(h_2\, h)\big]\Big)\\ &= \mu(h_2\, h) = \nu(h)\,. \end{aligned} \tag{8.12}$$

The third inequality is due to the $\mathcal{F}_{\leqslant k}$-measurability of h and the fourth one to the consistency of μ with f. Identities (8.11) and (8.12) prove that ν is consistent with $\widetilde{f}$ on $]-\infty, m_\Lambda]$.
2nd stage: For every $k \in \Lambda \cup \Lambda_-$ and $\omega \in \Omega_{-\infty}^{k-1}$, we construct $b_k(\omega)$ such that

$$\Big\| \overset{\circ}{f}_k(\cdot\,|\,\omega) - \overset{\circ}{\widetilde{f}}_k(\cdot\,|\,\omega)\Big\|_d \leqslant b_k(\omega)\,. \tag{8.13}$$

For starters, we can take

$$b_k = 0 \quad \forall\, k \in [m_\Delta + 1, m_\Lambda]\,, \tag{8.14}$$

because $\overset{\circ}{f}_k(\cdot\,|\,\omega) = \overset{\circ}{\widetilde{f}}_k(\cdot\,|\,\omega)$, for $k \in [m_\Delta + 1, m_\Lambda]$ and $\omega \in \Omega_{-\infty}^{k-1}$.

We fix $k \in \Delta \cup \Delta_-$ and $\omega \in \Omega_{-\infty}^{k-1}$ and consider the set $\Omega_k^\omega = \{\omega_k \in \Omega_{\{k\}} : \omega_{-\infty}^k \in \Omega_{-\infty}^k\}$ with the restricted topology and Borel σ-algebra. To abbreviate the notation we introduce the function $u : \Omega_k^\omega \to \mathbb{R}$ defined by

$$u(x) \triangleq g_k\Big(\omega_{-\infty}^{k-1}\, x\Big) = \frac{f_{[k+1,m_\Delta]}\big(h_2\,\big|\,\omega_{-\infty}^{k-1}\, x\big)}{f_{[k,m_\Delta]}\big(h_2\,\big|\,\omega\big)}\, \mathbb{1}_A \tag{8.15}$$

and the measure

$$\alpha \triangleq \overset{\circ}{f}_k(\cdot\,|\,\omega) \tag{8.16}$$

on Ω_k^ω. Notice that

$$\overset{\circ}{\widetilde{f}}_k(\cdot\,|\,\omega) - \overset{\circ}{f}_k(\cdot\,|\,\omega) = u\,\alpha - \alpha\,. \tag{8.17}$$

We denote, for each $\mathcal{F}_{\{k\}}$-measurable function h,

$$m_h \triangleq \frac{\sup h + \inf h}{2}\,.$$

Claim (*i*)

$$\left\| h - m_h D \right\|_\infty \leqslant \frac{D}{2}\,\delta_k^d(h) \tag{8.18}$$

Indeed,

$$\begin{aligned}\sup_{x\neq y}\frac{h(x)-h(y)}{2d(x,y)} &\geqslant \sup_{x\neq y}\frac{h(x)-h(y)}{2D}\\ &= \frac{1}{2D}\,[\sup h - \inf h]\\ &= \frac{1}{D}\Big[\sup h - \Big(\frac{\sup h}{2}+\frac{\inf h}{2}\Big)\Big]\\ &= \frac{1}{D}\,\|h-m_h\|_\infty\,.\end{aligned}$$

Claim (*ii*)

$$\left\| \overset{\circ}{f}_k\,(\cdot\,|\,\omega) - \overset{\circ}{\tilde{f}}_k\,(\cdot\,|\,\omega)\right\|_d \leqslant \frac{D}{2}\,\alpha(|u-1|)\,. \tag{8.19}$$

Indeed, for $h\in\mathcal{B}_d(\{k\})$ with $\delta_k^d(h)\leqslant 1$ we have

$$\Big|u\,\alpha(h)-\alpha(h)\Big| = \Big|\alpha\big[(u-1)(h-m_h)\big]\Big| \leqslant \alpha\,(|u-1|)\,\left\|h-m_h\right\|_\infty\,.$$

From this and (8.18), assertion (8.19) follows.

We now use Schwarz's inequality to bound

$$\alpha(|u-1|) = \alpha\big(|u-\alpha(u)|\big) \leqslant \Big[\alpha\Big((u-\alpha(u))^2\Big)\Big]^{\frac{1}{2}}\,,$$

and since $\alpha(u)$ minimizes $x\longmapsto\alpha\Big((u-x)^2\Big)$, we obtain $(x=m_u)$

$$\alpha(|u-1|) = \Big[\alpha\Big((u-m_u)^2\Big)\Big]^{\frac{1}{2}} \leqslant \|u-m_u\|_\infty\,. \tag{8.20}$$

The combination of (8.18)–(8.20) gives (8.13) with

$$
\begin{aligned}
b_k(\omega) &\triangleq \frac{D^2\,\delta_k^d(u)}{4} \\
&= \frac{D^2}{4}\,\delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right)\,\frac{1\!\!1_A(\omega)}{f_{[k,m_\Delta]}(h_2\,|\,\omega)}\,.
\end{aligned}
\tag{8.21}
$$

3rd stage: We estimate $\nu\,(b_k)$. From (8.21):

$$
\nu(b_k) = \frac{D^2}{4}\,\delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right)\,\mu\left(\frac{h_2\,1\!\!1_A}{f_{[k,m_\Delta]}h_2}\right).
$$

By consistency, $\mu = \mu\, f_{[k,m_\Delta]}$, hence the last factor is bounded by $\mu(A) \leqslant 1$. From this and (8.14) we conclude that

$$
\nu(b_k)\quad \begin{cases} = 0 & \text{if } k \in [m_\Delta + 1, m_\Lambda] \\ \leqslant \dfrac{D^2}{4}\,\delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right) & \text{if } k \in \Delta \cup \Delta_-\,. \end{cases}
\tag{8.22}
$$

In view of (8.7), (8.13) and (8.22) imply (5.10) by Theorem 5.5. ▮

Proof of Corollary 5.7. Part (i) follows from (7.6) and (5.10), and part (ii) from (7.5) and (5.12). ▮

ACKNOWLEDGMENTS

R.F. wishes to thank Michael Aizenman for many enlightening discussions on Dobrushin's criterion. The authors are thankful to J. Steif for pointing out a mistake in a previous version of this paper. They also thank the referees for their careful and encouraging criticism.

REFERENCES

1. H. Berbee. Chains with complete connections: Uniqueness and Markov representation. *Prob. Th. Rel. Fields* **76**:243–253, (1987).
2. X. Bressaud, R. Fernández, and A. Galves. Decay of correlations for non Hölderian dynamics. a coupling approach. *Elect. J. Prob.* **4**, 1999 (http://www.math.washington.edu/~ejpecp/).
3. N. Berger, C. Hoffman, and V. Sidoravicius. Nonuniqueness for specifications in $\ell^{2+\epsilon}$. Preprint arXiv:mat.PR/0312344 (2003).
4. M. Bramson and S. A. Kalikow. Nonuniqueness in g-functions. *Israel J. Math.* **84**:153–160 (1993).

5. R. M. Dudley. *Real analysis and probability* (Cambridge University Press, 2002).
6. R. Fernández and G. Maillard. Chains with complete connections and one-dimensional Gibbs measures. *Elect. J. Prob.*, **9** (2004). (http://www.math.washington.edu/~ejpecp/).
7. R. Fernández and G. Maillard. Chains and specifications. *Markov Proc. Rel. Fields*, **10**:435–456 (2004).
8. H.-O. Georgii. Stochastische Felder und ihre Anwendung auf Interaktionssysteme. *Lecture Notes, Institut für Angewandte Mathematik, Universität Heidelberg* (1974).
9. H.-O. Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter (de Gruyter Studies in Mathematics, Vol. 9), Berlin–New York (1988).
10. T. E. Harris. On chains of infinite order. *Pacific J. Math.* **5**:707–724 (1955).
11. M. Iosifescu. A coupling method in the theory of dependence with complete connections according to Doeblin. *Rev. Roum. Math. Pures et Appl.* **37**:59–65 (1992).
12. M. Iosifescu and A. Spataru. On denumerable chains of infinite order. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **27**:195–214 (1973).
13. A. Johansson and A. Öberg. Square summability of variations of g-functions and uniqueness of g-measures. *Math. Res. Lett.* **10**:587–601 (2003).
14. S. Kalikow. Random Markov processes and uniform martingales. *Isr. J. Math.* **71**(1):33–54 (1990).
15. M. Keane. Strongly mixing g-measures. *Inventiones Math.* **16**:309–324 (1972).
16. Y. Lacroix. A note on weak-$*$ perturbations of g-measures. *The Indian Journal of Statistics* **62**:331–338 (2000).
17. S. P. Lalley. Regeneration representation for one-dimensional Gibbs states. *Ann. Prob.* **14**:1262–1271 (1986).
18. G. Maillard. *Chaînes à liaisons complètes et mesures de Gibbs unidimensionnelles*. PhD thesis, Université de Rouen (2003).
19. O. Onicescu and G. Mihoc. Sur les chaînes statistiques. *C. R. Acad. Sci. Paris* **200**:511–512 (1935).
20. D. Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Commun. Math. Phys.* **9**:267–278 (1968).
21. B. Simon. *The Statistical Mechanics of Lattice Gases*. Series in Physics. Princeton (1993).
22. O. Stenflo. Uniqueness in g-measures. *Institute of Physics Publishing* **16**:403–410 (2002). Nonlinearity.
23. 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).
24. P. Walters. Ruelle's operator theorem and g-measures. *Trans. Am. Math. Soc.* **214**:375–387 (1975).