

Chains and specifications

Roberto Fernández[†]
Grégory Maillard[‡]

July 17, 2004

Abstract: We review four types of results combining or relating the theories of discrete-time stochastic processes and of one-dimensional specifications. First we list some general properties of stochastic processes which are extremal among those consistent with a given set of transition probabilities. They include: triviality on the tail field, short-range correlations, realization via infinite-volume limits and ergodicity. Second we detail two new uniqueness criteria for stochastic processes and discuss corresponding mixing bounds. These criteria are analogous to those obtained by Dobrushin and Georgii for Gibbs measures. Third, we discuss conditions for a stochastic process to define a Gibbs measure and vice versa, that generalize well known equivalence results between ergodic Markov chains and fields. Finally we state a (re)construction theorem for specifications starting from single-site conditioning, which applies in a rather general setting.

Keywords: Discrete-time stochastic processes, chains with complete connections, Gibbs measure, Markov chains, ergodicity and rates of mixing.

AMS Subject Classification: 60G07; 82B05; 60G10; 60G60; 37A25; 60J05; 60J10.

[†]Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen, Site Colbert F-76821 Mont Saint Aignan, France; roberto.fernandez@univ-rouen.fr

[‡]EURANDOM, P.O. BOX 513, 5600 MB Eindhoven, The Netherlands; maillard@eurandom.tue.nl

1 Introduction

A major limitation of current non-Gibbsian theory is the lack of practical examples where the phenomenon play a conspicuous, if not crucial, role. A decade of studies has elucidated many subtle mathematical aspects of the Gibbs-vs-non-Gibbs issue. But a natural question remains unanswered: Can an instance be determined where something like a calculation of a physical quantity or a simulation algorithm is proven wrong or shown to fail because of non-Gibbsianness?

It occurred to us that stochastic processes could offer a simpler laboratory in this regard. On the one hand, the key aspects of Gibbsianness—non-nullness and quasi-locality—admit natural transcriptions in terms of one-sided conditional probabilities. On the other hand, many classes of processes have a well-developed theory that includes a wealth of construction and approximation schemes which can be used to estimate large-deviation rates, convergence rates of numerical algorithms, etc. A proven breakdown of some of these schemes due to the lack of continuity or positivity of the process could, perhaps, be shown to produce some observable consequence.

More generally, it appears that a convenient strategy towards these goals would be to *combine* the stochastic-process approach with the theory of one-dimensional specifications on which the notion of (non-)Gibbsianness depends. This triggered a preliminary program of recollection and development of links between both frameworks which, in fact, has led us to some new contributions to each of them (specially to the former). In this paper we present a panorama of the current status of this program, which was the subject of the Phd thesis of one of us (Maillard, 2003). It is a summary of two preprints (Fernández and Maillard 2003a and 2003b), to which we refer the reader for the proofs.

We have investigated three aspects of the relation between processes and specifications.

(I) Specification-theoretical techniques in the study of processes

While phase transitions are a well-known phenomenon in the theory of processes (Bramson and Kalikow, 1993), the statistical mechanical emphasis on *phase diagrams* seems to be lacking. Measures have been studied largely on an individual basis, without exploiting the fact of having the same transition probabilities. The results of Section 3—tail-field triviality, mixing properties and ergodicity of extremal processes—go in the direction of fill-

ing this gap. To tackle these issues we have found convenient to introduce the notion of *LIS* (left-interval specification), which is the one-sided analogue of a specification. While each LIS is uniquely determined by the single-site transitions traditionally used to define processes, it provides a more flexible tool to discuss asymptotic properties. Another by-product of the use of LIS is contained in Sections 4 and 5. We obtain there new uniqueness criteria and bounds on mixing rates by adapting well-known arguments for specifications.

(II) When is a Gibbs measure a process and viceversa?

The issue can be cast in more general terms. Given a process consistent with some nicely behaved transition probabilities (e.g. Markov or continuous), is it always consistent with some specification enjoying similar properties? An analogous question arises in the opposite direction. In Section 6 we address this issue by attempting to establish appropriate maps between specifications and LIS. For comparison, we mention the work discussed in Georgii (1988), Chapter 10, showing that general processes with (one-sided) Markovian transitions have two-sided Markovian *conditional probabilities*. These results, however, are not sufficient for our purposes. The passage from conditional probabilities to specifications, while always possible (see, e.g., Sokal, 1981), may not preserve the relevant properties.

(III) Process techniques in the study of specifications

As the only result in this direction we mention the algorithm of Section 7 to construct a specification from single-site kernels. This opens the possibility of studying specification-consistent measures on the basis of singleton conditional probabilities, as is usually done for process.

We have tried to include enough definitions to make the review self-contained. In particular, we start with a careful presentation of the process formalism. Such a presentation serves two purposes. First, we introduce the notation, *à la* Georgii (1988), needed to write reasonably compact formulas. Second, we take the opportunity to recall and compare a number of notions—transition kernels, transition densities, transition functions—which are closely related but whose identification may lead to confusion and to some degree of notational chaos.

2 Specification-framework for processes

2.1 Standard set up for processes

The objects fixed from the outset are a measurable space (E, \mathcal{E}) —called, for brevity, the *alphabet*— and a family $\Omega \subset E^{\mathbb{Z}}$ of (*admissible*) *configurations*. The rest of the construction is done on the measurable space (Ω, \mathcal{F}) , where \mathcal{F} is the restriction to Ω of $\mathcal{E}^{\mathbb{Z}}$. A *process* on (Ω, \mathcal{F}) is a probability measure on this space, or, equivalently, a measure on $(E^{\mathbb{Z}}, \mathcal{E}^{\mathbb{Z}})$ which has its support in Ω . They are canonically described by the resulting laws on the family of projections $X_n : \Omega \rightarrow E$, $X_n(\omega) = \omega_n$. Furthermore, in probability theory processes are classified according to their single-site conditional probabilities, interpreted as “transition probabilities”.

To describe these objects we need some notation. For $\Lambda \subset \mathbb{Z}$, let $X_\Lambda : \Omega \rightarrow E^\Lambda$; $\omega \mapsto \omega_\Lambda$, denote the projection $X_\Lambda = (X_n)_{n \in \Lambda}$ and let Ω_Λ be its inverse image and \mathcal{F}_Λ the induced σ -algebra on the latter, namely the one generated by cylinders with base in Ω_Λ . For interval-like Λ we shall prefer “range” notation: $\Omega_\ell^m \triangleq \Omega_{[\ell, m]}$, $\Omega_{\leq m} \triangleq \Omega_{] -\infty, m]}$ = $\Omega_{-\infty}^m$. This notation will be used for all relevant objects like in ω_ℓ^m , $\mathcal{F}_{\leq m}$, etc.

Definition 2.1 *A family of (singleton) transition probabilities is a set $(f_n)_{n \in \mathbb{Z}}$ of kernels $f_n(\cdot \mid \cdot) : \mathcal{F}_{\leq n} \times \Omega \rightarrow [0, 1]$ such that*

- (a) *each f_n is a probability measure with respect to the first coordinate: $f_n(\Omega_{\leq n} \mid \omega) = 1$ for each $n \in \mathbb{Z}$ and each $\omega \in \Omega$;*
- (b) *each f_n is a $\mathcal{F}_{\leq n-1}$ -measurable function with respect to the second coordinate (“past”), that is $f_n(A \mid \cdot)$ is $\mathcal{F}_{\leq n-1}$ -measurable for each $A \in \Omega_{\leq n}$, and*
- (c) *each probability measure f_n is deterministically frozen for each given past: $f_n(B \mid \omega) = \mathbf{1}_B(\omega)$ if $B \in \mathcal{F}_{\leq n-1}$.*

The kernels f_n model an stochastic evolution, where each past only determines a probability distribution for the next state of the evolution. While it is natural to think $f_n(\cdot \mid \omega)$ as a probability measure on Ω_n —the configuration space at time n —, it is mathematically more convenient to define them as measures on the increasing family of σ -algebras $\mathcal{F}_{\leq n}$.

Definition 2.2 A measure μ on (Ω, \mathcal{F}) is a **process consistent with the family** (f_n) if

$$\int f_n(B \mid \omega) \mu(d\omega) = \mu(B) \quad (2.3)$$

for each $n \in \mathbb{Z}$ and each $B \in \mathcal{F}_{\leq n}$.

If, to simplify, we denote $\nu(h) \triangleq E_\nu(h)$ for a measure ν and a ν -integrable function h , this identity becomes

$$\mu\left(f_n(h \mid \cdot)\right) = \mu(h) \quad (2.4)$$

for each $\mathcal{F}_{\leq n}$ -measurable h . This last equation can, and will, be written briefly as $\mu(f_n(h)) = \mu(h)$. From the point of view of probability-theory the requirements (2.3)/(2.4) are equivalent to demanding that for each $\mathcal{F}_{\leq n}$ -measurable h the function $f_n(h)$ be (a realization of) the conditional expectation $E_\mu(h \mid X_{\leq n})$. In the *translation-* or *shift-invariant* case, the different f_n are obtained as translations of a single kernel, for instance f_0 . This amounts to suppressing the subscript n of f_n .

Every \mathcal{F}_Λ -measurable function h is of the form $h(\omega) = \tilde{h}(X_\Lambda(\omega)) = \tilde{h}(\omega_\Lambda)$ for some function \tilde{h} . It is customary to identify h with \tilde{h} . Likewise, every measure ν on \mathcal{F}_Λ can be identified with a measure $\tilde{\nu} = \nu \circ X_\Lambda^{-1}$ on E^Λ . With these identifications (2.4) can be written in the form

$$\int \left[\int h(x_{\leq n}) f_n(d\omega_{\leq n} \mid \sigma_{-\infty}^{n-1}) \right] \mu(d\sigma) = \int h(\omega_{\leq n}) \mu(d\omega) \quad (2.5)$$

for all $h \in \mathcal{F}_{\leq n}$.

Objects like these have received a variety of names in the literature, like *chain with complete connections* (Onicescu–Mihoc, 1935), *chain of infinite order* (Harris, 1955) and *g-measure* (Keane, 1972). In the presence of continuity hypotheses of the type reviewed in Section 4, they correspond to what Kalikow (1990) calls *random Markov processes* or *uniform martingales*.

Usually, the alphabet space (E, \mathcal{E}) comes equipped with a natural measure λ and the transition probabilities are defined by density functions $p_n : \Omega_n \times \Omega_{\leq n-1} \rightarrow [0, \infty)$ in the form

$$f_n(d\omega_{\leq n} \mid \sigma_{-\infty}^{n-1}) = p_n(\omega_n \mid \sigma_{-\infty}^{n-1}) \left[\lambda(d\omega_n) \times \delta_{\sigma_{-\infty}^{n-1}}(d\omega_{\leq n-1}) \right]. \quad (2.6)$$

This is the standard situation when E is finite or countable, with λ equal to the counting measure. In the shift-invariant case,

$$p_n(\omega_n \mid \omega_{-\infty}^{n-1}) = g(\omega_{-\infty}^n) \quad (2.7)$$

for a single function g [in symbolic dynamical systems it is convenient to invert the sense of time and work with $g(z_n^\infty)$].

For completeness, let us mention some well-known examples of families of transition probabilities (we write them in the shift-invariant version for simplicity).

- *Markov of order k* : For each $B \in \mathcal{F}_0$ the function $f(B \mid \cdot)$ is measurable with respect to $\mathcal{F}_{[-k, -1]}$. With obvious identifications this can be codified through the identity

$$f(B \mid \sigma_{-\infty}^{-1}) = f(B \mid \sigma_{-k}^{-1}).$$

In the case (2.6), the densities p_n are of the form $p_n(\omega_n \mid \sigma_{n-k}^{n-1})$.

- *Mixture of Markov*:

$$f(d\omega_{\leq 0} \mid \sigma_{-\infty}^{-1}) = \sum_k a_k f^{(k)}(d\omega_{\leq 0} \mid \sigma_{-k}^{-1})$$

where each $f^{(k)}$ is a Markovian kernel of order k and $0 \leq a_k \leq 1$, $\sum_k a_k = 1$. The case of finite mixtures has been put forward by Raftery (1985). In fact, if the alphabet is countable every kernel of the form (2.6) with p_n continuous with respect to the second argument can be written as a (possibly infinite) mixture of Markov kernels.

- *Variable-length Markov*: For each $B \in \mathcal{F}_0$,

$$f(B \mid \sigma_{-\infty}^{-1}) = f(B \mid \sigma_{-k(\sigma)}^{-1})$$

where $k(\sigma)$ is a backwards stopping time, in the sense that the event $\{k = \ell\}$ is measurable with respect to $\mathcal{F}_{[-\ell, 0]}$. Transitions of this type, but with $k(x) \leq K$ for some finite $K \in \mathbb{N}$ have been studied by Bühlman and Wyner (1999). If $k(x)$ is unbounded, the kernels are generally discontinuous with respect to the past.

2.2 Standard set up for measures consistent with specifications

For comparison purposes let us now summarize the specification framework as used in mathematical statistical mechanics. We denote by \mathcal{S} the family of finite subsets of \mathbb{Z} .

Definition 2.8 A **specification** γ on (Ω, \mathcal{F}) is a family of probability kernels $\{\gamma_\Lambda\}_{\Lambda \in \mathcal{S}}$, $\gamma_\Lambda : \mathcal{F} \times \Omega \rightarrow [0, 1]$ such that for all Λ in \mathcal{S} ,

- (a) For each $A \in \mathcal{F}$, $\gamma_\Lambda(A | \cdot) \in \mathcal{F}_{\Lambda^c}$.
- (b) For each $B \in \mathcal{F}_{\Lambda^c}$ and $\omega \in \Omega$, $\gamma_\Lambda(B | \omega) = \mathbb{1}_B(\omega)$.
- (c) For each $\Delta \in \mathcal{S} : \Delta \supset \Lambda$,

$$\gamma_\Delta \gamma_\Lambda = \gamma_\Delta . \tag{2.9}$$

Property c) is called *consistency*, and it is written as a composition of probability kernels. With our notation for expectations, (2.9) is equivalently to

$$\gamma_\Delta \left(\gamma_\Lambda(h | \cdot) \middle| \sigma \right) = \gamma_\Delta(h | \omega) \tag{2.10}$$

for each \mathcal{F} -measurable function h and configuration $\sigma \in \Omega$.

A *Markov specification of range k* corresponds to the particular case in which for each $\Lambda \in \mathcal{S}$ the functions $\gamma_\Lambda(A | \cdot) \in \mathcal{F}_{\partial_k \Lambda}$ for all $A \in \mathcal{F}_\Lambda$. Here $\partial_k \Lambda = \{i \in \Lambda^c : |i - j| \leq k \text{ for some } j \in \Lambda\}$. Gibbsian specifications are defined in terms of interactions via the Boltzmann prescription. They can be characterized by suitable continuity and non-nullness properties which, at the present level of generality, involve some subtleties regarding spaces of observables and uniformity of the non-nullness. These have been discussed at some length in van Enter, Fernández and Sokal (1992), Section 2.3.3.

Definition 2.11 A probability measure μ on (Ω, \mathcal{F}) is said to be **consistent** with a specification γ if

$$\mu \gamma_\Lambda = \mu \quad \forall \Lambda \in \mathcal{S} . \tag{2.12}$$

The family of these measures will be denoted $\mathcal{G}(\gamma)$.

In comparing with the previous section, we see three major differences between the frameworks provided by transition probabilities and by specifications.

- (D1) Transition probabilities involve only singleton (single-site) kernels, while specifications require kernels defined for each finite region.
- (D2) Transition probabilities involve conditioning only with respect to the past, while each specification kernel corresponds to conditioning with respect to the whole complement of the relevant region, in particular past and future. This is the essential difference at the conceptual level.
- (D3) At the level of singletons, transition probabilities involve a further measurability issue. Indeed, in (2.3)/(2.4), only observables depending on the evolution up to time n are allowed, while no similar restriction is present in the corresponding expression (2.12) for specifications.

In order to exploit the complementarity of the two frameworks, and to establish relationships between them, we develop in the sequel approaches that reduce some of these differences. We work on issue (D1) from both sides: On the one hand (Section 2.3) we propose the systematic use of systems of multi-site transition probabilities that, as proven by the results of Sections 3, 4 and 5, lend themselves more easily to specification techniques. On the other hand, in Section 7 we discuss the possibility of characterizing specifications through their singletons, as is the case for transition probabilities. In Section 6 we address issue (D2) and we determine regimes where one- and two-sided conditionings are related.

2.3 LIS

A way to treat transition probabilities on a more similar footing to specifications is to consider transitions to multi-site regions. Due to the measurability constraint invoked in (D3) only interval-like regions can be considered (regions with holes would lead to partial integrals with forbidden dependences on *future* sites). Multi-site transitions are obtained from singletons in the obvious manner :

$$f_{[\ell,m]} \triangleq f_\ell f_{\ell+1} \cdots f_m \tag{2.13}$$

that is,

$$f_{[\ell,m]}(A|x_{-\infty}^{\ell-1}) = f_\ell \left[f_{\ell+1}(\cdots f_m(A|\cdot) \cdots |\cdot) \Big| x_{-\infty}^{\ell-1} \right]$$

for $A \in \mathcal{F}_{\leq m}$. The resulting family admits a formalization parallel to that of specifications, except that only intervals are allowed and that conditioning, and measurability, are only leftwards. That is why we call them *left interval-specifications*. In the sequel let us adopt the notation

$$\mathcal{S}_b = \text{finite intervals in } \mathbb{Z} \quad , \quad l_\Lambda \triangleq \min \Lambda \quad , \quad m_\Lambda \triangleq \max \Lambda$$

Definition 2.14 A **left interval-specification (LIS)** f on (Ω, \mathcal{F}) is a family of probability kernels $\{f_\Lambda\}_{\Lambda \in \mathcal{S}_b}$, $f_\Lambda : \mathcal{F}_{\leq m_\Lambda} \times \Omega \rightarrow [0, 1]$ such that for all Λ in \mathcal{S}_b ,

- (a) For each $A \in \mathcal{F}_{\leq m_\Lambda}$, $f_\Lambda(A | \cdot)$ is $\mathcal{F}_{\leq l_\Lambda - 1}$ -measurable.
- (b) For each $B \in \mathcal{F}_{\leq l_\Lambda - 1}$ and $\omega \in \Omega$, $f_\Lambda(B | \omega) = \mathbb{1}_B(\omega)$.
- (c) For each $\Delta \in \mathcal{S}_b : \Delta \supset \Lambda$,

$$f_\Delta f_\Lambda = f_\Delta \quad \text{on } \mathcal{F}_{\leq m_\Lambda} .$$

Definition 2.15 A probability measure μ on (Ω, \mathcal{F}) is said to be **consistent** with a LIS f (or a **chain** consistent with f) if for each $\Lambda \in \mathcal{S}_b$

$$\mu f_\Lambda = \mu \quad \text{on } \mathcal{F}_{\leq m_\Lambda}$$

The family of these measures will be denoted $\mathcal{G}(f)$.

Of course, the LIS and the standard set-up are totally equivalent:

Theorem 2.16 Let $(f_i)_{i \in \mathbb{Z}}$ be a family of probability kernels $f_i : \mathcal{F}_{\leq i} \times \Omega \rightarrow [0, 1]$ such that for each $i \in \mathbb{Z}$

- (a) $f_i(A | \cdot)$ is $\mathcal{F}_{\leq i-1}$ -measurable, for each $A \in \mathcal{F}_{\leq i}$.
- (b) $f_i(B | \omega) = \mathbb{1}_B(\omega)$ for each $B \in \mathcal{F}_{\leq i-1}$ and $\omega \in \Omega$,

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

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

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

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

This theorem, left as an exercise to the reader, shows that the notion of LIS does not, in itself, incorporate any new ingredient to the standard set-up. Nevertheless, LIS are useful objects to study the structure and properties of the set of processes consistent with a given family of transition probabilities. Results so obtained are reviewed in the next three Sections.

3 General results for LIS

The results grouped in this section do not impose *any* hypothesis on the LIS. They correspond to well-known properties of measures consistent with specifications (see, for instance, Chapters 7 and 14 in Georgii, 1988).

Theorem 3.1 (Extremality and triviality) *Let $f = (f_\Lambda)_{\Lambda \in \mathcal{S}_b}$ be a LIS on (Ω, \mathcal{F}) . Let $\mathcal{F}_{-\infty} \triangleq \bigcap_{k \in \mathbb{Z}} \mathcal{F}_{\leq k}$ (left 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 it 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 $h \geq 0$ such that $\nu = h\mu$.
- (d) Each $\mu \in \mathcal{G}(f)$ is uniquely determined (within $\mathcal{G}(f)$) by its restriction to the left 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 3.2 (Triviality and mixing) *For each probability measure μ on (Ω, \mathcal{F}) , the following statements are equivalent.*

- (a) μ is trivial on $\mathcal{F}_{-\infty}$.
- (b) $\lim_{\Lambda \uparrow \mathbb{Z}} \sup_{B \in \mathcal{F}_{\leq \Lambda-1}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0$, for all cylinder sets $A \in \mathcal{F}$.
- (c) $\lim_{\Lambda \uparrow \mathbb{Z}} \sup_{B \in \mathcal{F}_{\leq \Lambda-1}} |\mu(A \cap B) - \mu(A)\mu(B)| = 0$, for all $A \in \mathcal{F}$.

Theorem 3.3 (Infinite volume limits) *Let f be a LIS, μ an extreme point of $\mathcal{G}(f)$ and $(\Lambda_n)_{n \geq 1}$ a sequence of regions in \mathcal{S}_b such that $\Lambda_n \uparrow \mathbb{Z}$. Then*

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

The next theorem refers to ergodic properties. Its statement involves the (right) shift $\tau(i) = i + 1$ on \mathbb{Z} , which induces, in the usual fashion, shift operators (also denoted by τ) on subsets of \mathbb{Z} , configurations and families thereof, functions and measures. We assume that Ω is shift invariant and denote by \mathcal{I} the σ -algebra of shift-invariant measurable sets, $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ the set of shift-invariant probability measures on (Ω, \mathcal{F}) , and $\mathcal{G}_{\text{inv}}(f)$ the family of shift-invariant chains consistent with a LIS f . A measure in $\mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ is **ergodic** if it is trivial on \mathcal{I} .

Theorem 3.4 (Ergodic chains) *Let f be a shift-invariant or stationary LIS, that is such that*

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

for each $\Lambda \in \mathcal{S}_{\text{b}}$ and $\omega \in \Omega$. Then

- (a) A chain $\mu \in \mathcal{G}_{\text{inv}}(f)$ is extreme in $\mathcal{G}_{\text{inv}}(f)$ if and only if it 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 of chains compatibles with LIS

We turn now to criteria under which a LIS has at most one consistent chain.

4.1 Uniqueness from transition-probability theory

A number of complementary uniqueness criteria are available in the literature, mostly obtained using coupling techniques for singleton probabilities. In their original versions they suppose a finite or countable alphabet, $\Omega = E^{\mathbb{Z}}$ and shift-invariant singletons. In particular, the transition probabilities have the form (2.6)–(2.7) for a suitable function g defined on $E^{-\mathbb{N}}$. The different criteria involve two types of hypotheses:

- (i) *Continuity hypotheses* referring to the speed at which the k -variation

$$\text{var}_k(g) \triangleq \sup \left\{ |g(\omega) - g(\sigma)| : \omega_{-k}^0 = \sigma_k^0 \right\} \quad (4.1)$$

decreases with k .

(ii) One of the following *non-nullness hypotheses*. A function g is

- **(uniformly) non-null** if

$$\inf_{\omega \in \Omega_{\leq 0}} g(\omega) > 0$$

- **weakly (uniformly) non-null** if

$$\sum_{\omega_0} \inf_{\omega_{-\infty}^{-1}} g(\omega) > 0$$

Here is a list of conditions under which the transitions probabilities defined by g have at most one consistent chain. All but the last one involve a continuity condition of the form

$$\sum_{n \geq 1} \prod_{k=1}^n \Delta_k = +\infty \quad (4.2)$$

for a suitable sequence $\Delta_k \in [0, 1]$.

(a) *Harris (1955)*: g is weakly non-null and

$$\Delta_k = 1 - \frac{|E|}{2} \text{var}_k(g) . \quad (4.3)$$

(b) *Berbee (1987)*: g is non-null and

$$\Delta_k = \exp\left(-\text{var}_k(\log g)\right) . \quad (4.4)$$

(c) *Stenflo (2003)*: g is non-null and

$$\Delta_k(g) = \inf_{(\omega, \sigma) : \omega_{-k}^{-1} = \sigma_{-k}^{-1}} \sum_{\omega_0} g(\omega_{-\infty}^{-1} \omega_0) \wedge g(\sigma_{-\infty}^{-1} \omega_0) . \quad (4.5)$$

(d) *Johansson and Öberg (2002)*: g is non-null and

$$\sum_{k \geq 0} \text{var}_k^2(\log g) < +\infty . \quad (4.6)$$

These conditions are largely complementary. Harris' criterion demands a weaker continuity condition than Stenflo's if $|E| = 2$ but the opposite is true for the remaining cases. The variance of the log appears in Berbee's criterion because this author compares values of the g function through ratios rather than through differences. He and Lalley (1986, 2000) obtain in addition a regeneration scheme when $\sum \text{var}_k(\log g) < \infty$. This scheme has been extended by Comets, Fernández and Ferrari (2002) to weakly non-null g satisfying a continuity condition like (4.2) but with Δ_k given in (5.5) below.

We feel that in most of the preceding criteria little effort has been put into optimizing the non-nullness condition. Guided by our experience with non-Gibbsianness, we believe that a precise determination of the non-nullness hypothesis is as important as that of the continuity hypothesis.

4.2 Uniqueness from LIS theory

Resorting to specification-like techniques, we can obtain two new uniqueness criteria. The first one is the analogue of a condition found by Georgii (1974; see also Georgii, 1988, Section 8.3) —later rediscovered by Lebowitz, Bricmont and Pfister (1974).

Theorem 4.7 (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 $A \in \mathcal{F}_{\leq m}$ there exists an integer $n < m$ such that*

$$f_{[n,m]}(A \mid \xi) \geq c f_{[n,m]}(A \mid \eta) \quad \text{for all } \xi, \eta \in \Omega. \quad (4.8)$$

Then there exists at most one chain consistent with f .

Requirement (4.8) corresponds to a certain “lack of rigidity” with respect to the past. It combines non-nullness and continuity properties in a not very explicit way (for instance it is satisfied if g is non-null and has summable variations).

Our second criterion is a transcription of the well-known Dobrushin criterion in statistical mechanics (Georgii, 1988, Chapter 8 and Simon, 1993, Chapter V). No non-nullness hypothesis is involved, but topology plays a role. Indeed, the alphabet space (E, \mathcal{E}) is assumed to have a Borel measurable structure defined by some bounded metric d . The LIS is required to be **continuous** in the sense that the functions $\Omega \ni \omega \longrightarrow f_\Lambda(A \mid \omega)$ be continuous for all $\Lambda \in \mathcal{S}_b$ and all $A \in \mathcal{F}_\Lambda$.

For a given function h on configurations and $j \in \mathbb{Z}$ we define the d -**oscillation** of h at j :

$$\delta_j^d(h) \triangleq \sup \left\{ \frac{|h(\xi) - h(\eta)|}{d(\xi_j, \eta_j)} : \xi_j \neq \eta_j \text{ and } \xi = \eta \text{ off } j \right\} \quad (4.9)$$

and we call a d -**sensitivity matrix** for the LIS f a nonnegative matrix $\alpha = (\alpha_{ij})_{i,j \in \mathbb{Z}}$ such that $\alpha_{ij} = 0$ if $j \geq i$ and

$$\delta_j^d(f_i h) \leq \delta_i^d(h) \alpha_{ij} \quad (4.10)$$

if $j < i$ and $h \in \mathcal{F}_i$.

Theorem 4.11 (One-sided Dobrushin) *If f is a continuous LIS for which there exists a d -sensitivity matrix α with*

$$\sum_{j < i} \alpha_{ij} < 1 \quad \text{for all } i \in \mathbb{Z}, \quad (4.12)$$

then there exists at most one chain consistent with f .

The canonical choice for a d -sensitivity matrix is:

$$C_{ij}(f) \triangleq \sup_{\xi = \eta \text{ off } j} \frac{\left\| \overset{\circ}{f}_i(\cdot | \xi) - \overset{\circ}{f}_i(\cdot | \eta) \right\|_d}{d(\xi_j, \eta_j)}, \quad j < i \quad (4.13)$$

and $C_{ij}(f) = 0$ otherwise. Here $\|\cdot\|_d$ is the Vasserstein-Kantorovich-Rubinstein distance and $\overset{\circ}{f}_i(\cdot | \xi)$ is the restriction of each singleton measure to $(\Omega_i, \mathcal{F}_i)$.

If d is the discrete metric [$d(\xi_j, \eta_j) = 1$ if $\xi_j \neq \eta_j$ and 0 otherwise], $\|\cdot\|_d$ coincides with the variational norm. For a countable alphabet we therefore have

$$C_{ij} = \delta_j(g_i) \triangleq \sup \left\{ |g_i(\xi) - g_i(\eta)| : \xi = \eta \text{ off } j \right\} \quad (4.14)$$

with g_i is the function on $\Omega_{-\infty}^i$ defined by

$$g_i(\omega_{-\infty}^i) \triangleq p_i(\omega_i | \omega_{-\infty}^{i-1}) \quad (4.15)$$

[p_i defined in (2.6)]. In this way we obtain the following version of Dobrushin criterion.

Corollary 4.16 *A continuous LIS on a countable alphabet admits at most one consistent chain if*

$$\sum_{j < i} \delta_j(g_i) < 1 \quad \text{for all } i \in \mathbb{Z}. \quad (4.17)$$

5 Mixing properties for LIS

The following results refer to the speed at which the limits of Theorem 3.2 are achieved. They apply in the uniqueness regime, that is, for LIS having at most one consistent measure. Mixing results come in two flavors. On the one hand, they bound the *loss of memory*, namely the sensitivity of the expectations $f_{[-n,0]}(h_0 \mid \omega)$ on the past ω as n grows, for a function h_0 measurable with respect to \mathcal{F}_0 . On the other hand, they offer bounds on correlations $\text{Cor}_\mu(h_0, h_n) = |\mu(h_0 h_n) - \mu(h_0)\mu(h_n)|$.

5.1 Mixing from coupling

The mixing results in the literature are a corollary of the arguments leading to the uniqueness criteria reviewed in Section 4.1 above. They apply in the same setting (countable alphabet, shift-invariant transitions defined by a function g) and assume a condition of the form (4.2) for a certain sequence $\Delta_k \in [0, 1]$, $k \in \mathbb{N}$. The results rely on an auxiliary *house of cards process* defined by this sequence. It is a Markovian process with $E = \mathbb{N}$ and transitions

$$\begin{aligned} P(k+1 \mid k) &= \Delta_k \\ P(0 \mid k) &= 1 - \Delta_k \end{aligned} \tag{5.1}$$

and zero otherwise. Let

$$\Delta_n^* = P(W_n = 0) \tag{5.2}$$

where W_n is the chain starting at $W_0 = 0$ and evolving with rates (5.1).

If h_0 is a bounded $\mathcal{F}_{\{0\}}$ -measurable function, the available mixing results say that

$$\sup_{\omega, \sigma} \left| f_{[-n,0]}(h_0 \mid \omega) - f_{[-n,0]}(h_0 \mid \sigma) \right| \leq c_{h_0} \Delta_n^* \tag{5.3}$$

in the following situations:

(a) g non-null and

$$\Delta_k = \exp[-\text{var}_{k+1}(\log g)]. \tag{5.4}$$

(b) g weakly non-null and

$$\Delta_k = \inf_w \sum_{\omega_0} \inf_{\eta} g(\eta_{-\infty}^{-k-1} \omega_{-k}^{-1} \omega_0). \tag{5.5}$$

We see that, in both cases, the unique measure $\mu \in \mathcal{G}(f)$ satisfies

$$\text{Cor}_\mu(h_0, h_n) \leq c_{h_0 h_n} \Delta_n^* \quad (5.6)$$

for h_n bounded and $\mathcal{F}_{<n}$ -measurable.

Result (a) was obtained by Bressaud, Fernández and Galves (1999), it was preceded by a slightly weaker version by Iosifescu (1992). Result (b) appears explicitly in Comets, Fernández and Ferrari (2002). In the regime $\prod_{k=1}^{\infty} \Delta_k > 0$ it follows also from the regenerative constructions by Berbee (1987) and Lalley (1986-2001).

Inequalities (5.3) and (5.6) are to be combined with the following properties of Δ_n^* (Bressaud, Fernández and Galves, 1999, appendix):

$$\sum_m \prod_{k=0}^m \Delta_k = +\infty \iff \Delta_n^* \rightarrow 0; \quad (5.7)$$

$$\prod_{k=0}^{\infty} \Delta_k > 0 \iff \sum_k (1 - \Delta_k) < +\infty \iff \sum_n \Delta_n^* < +\infty; \quad (5.8)$$

$$1 - \Delta_k \leq \text{const } k^{-a} \implies \Delta_n^* \leq \text{const } n^{-a}, \quad (5.9)$$

and

$$1 - \Delta_k \leq \text{const } \exp(-ak) \implies \Delta_n^* \leq \text{const}(a^-) \exp[-(a^-)n] \quad (5.10)$$

where a^- stands for any number strictly less than a .

5.2 Mixing from Dobrushin theory

In the setting of Section 4.2 we introduce the space of functions of *bounded d -oscillations* in Λ :

$$\mathcal{B}_d(\Lambda) \triangleq \left\{ \mathcal{F}_\Lambda\text{-measurable } h : \sup_{j \in \Lambda} \delta_j^d(h) < \infty \right\}, \quad (5.11)$$

for each $\Lambda \subset \mathbb{Z}$. We also introduce the projection matrices

$$(P_\Lambda)_{kj} = \begin{cases} 1 & \text{if } k = j \text{ and } k \in \Lambda \\ 0 & \text{otherwise} \end{cases} \quad (5.12)$$

and we denote

$$\left[\frac{A}{1-A} \right]_{kj} \triangleq \sum_{n \geq 1} [A^n]_{kj} , \quad (5.13)$$

for a matrix $(A_{kj})_{k,j \in \mathbb{Z}}$ with nonnegative entries.

Dobrushin's "dusting" argument yields the following loss-of-memory bound.

Theorem 5.14 *Let f be a continuous LIS and (α_{ij}) a d -sensitivity matrix for f . Then, for every $\Lambda \in \mathcal{S}_b$, $j < l_\Lambda$ and $h \in \mathcal{B}_d(\Lambda)$,*

$$\delta_j^d(f_\Lambda h) \leq \sum_{k \in \Lambda} \delta_k^d(h) \left[\sum_{l=1}^{|\Lambda|} (P_\Lambda \alpha)^l \right]_{kj} = \sum_{k \in \Lambda} \delta_k^d(h) \left[\frac{P_\Lambda \alpha}{1 - P_\Lambda \alpha} \right]_{kj} . \quad (5.15)$$

This theorem yields obvious mixing results. For a compact alphabet, however, we get a more precise estimation of the mixing rates.

Theorem 5.16 *Assume Ω compact and let f be a continuous LIS satisfying the Dobrushin condition (4.12) for a d -sensitivity matrix (α_{ij}) . Let μ be the only chain consistent with f (it exists by compactness). Then if $h_1 \in \mathcal{B}_d(\Lambda)$ and $h_2 \in \mathcal{B}_d(\Delta)$, with $m_\Delta < l_\Lambda$,*

$$\text{Cor}_\mu(h_1, h_2) \leq \frac{D^2}{4} \sum_{l \in \Delta} \sum_{m \in \Lambda} \delta_m^d(h_1) A_{ml} \delta_l^d(h_2) \quad (5.17)$$

where

$$A_{ml} \triangleq \left[\frac{P_\Lambda \alpha}{1 - P_\Lambda \alpha} \right]_{ml} + \sum_{k \leq m_\Delta} \left[\frac{P_\Lambda \alpha}{1 - P_\Lambda \alpha} \right]_{mk} \left[\frac{P_{[k+1, m_\Delta]} \alpha}{1 - P_{[k+1, m_\Delta]} \alpha} \right]_{lk} . \quad (5.18)$$

To estimate the rates stated in the preceding theorems, the following argument is useful. Suppose that there exists a function $F : \mathbb{Z}^2 \rightarrow \mathbb{R}^+$ satisfying the triangle inequality $F(i, j) \leq F(i, k) + F(k, j)$, for all $i, j, k \in \mathbb{Z}$, such that

$$\gamma_i \triangleq \sum_{j < i} \alpha_{ij} e^{F(i,j)} < 1 , \quad (5.19)$$

for each $i \in \mathbb{Z}$. Then for each $\Lambda \in \mathcal{S}_b$

$$[(P_\Lambda \alpha)^n]_{kj} \leq \gamma_\Lambda e^{-F(k,j)} , \quad (5.20)$$

with $\gamma_\Lambda = \max_{i \in \Lambda} \gamma_i$. As a consequence,

$$\left[\frac{P_\Lambda \alpha}{1 - P_\Lambda \alpha} \right]_{kj} \leq \frac{\gamma_\Lambda}{1 - \gamma_\Lambda} e^{-F(k,j)}. \quad (5.21)$$

The bounds (5.20) and (5.21) can be used for exponential and power-law sensitivities, in the latter case through functions of the form $F(i, j) = c \log(1 + |i - j|)$.

It is easy to construct examples showing that these mixing results are complementary to those of Section 5.1.

6 Relations between LIS and specifications

6.1 The objective

Our goal, in this part of our program, is to establish maps between LIS and specifications that preserve important properties, in particular consistency of measures. At the very least we seek maps

$$b: f \longrightarrow \gamma^f \quad \text{such that} \quad \mathcal{G}(f) \subset \mathcal{G}(\gamma^f) \quad (6.1)$$

and

$$c: \gamma \longrightarrow f^\gamma \quad \text{such that} \quad \mathcal{G}(\gamma) \subset \mathcal{G}(f^\gamma). \quad (6.2)$$

In addition we impose two natural demands:

1. The applications should pass on as many properties as possible, such as continuity, non-nullness, loss of memory rates and validity of uniqueness criteria.
2. In favorable cases, the maps must be inverse of each other:

$$\gamma^{f^\gamma} = \gamma \quad , \quad f^{\gamma^f} = f \quad (6.3)$$

A program like this has been completely achieved for Markov processes with finite alphabet through matrix-theoretical arguments (see, for instance, Georgii, 1988, Chapter 3). Our results, besides generalizing the Markovian results in many directions, offer an alternative purely probabilistic treatment of them.

6.2 The spaces

In this section the alphabet is assumed to be finite. For simplicity we state the results in a shift-invariant setting. To be precise, we need two definitions and a collection of spaces.

Let us denote \mathcal{S}_\rightarrow the family of intervals of the form $V = [i, +\infty[$, $i \in \mathbb{Z}$, or $V = \mathbb{Z}$. We say that a LIS f satisfies a **hereditary uniqueness condition (HUC)** if for all $V \in \mathcal{S}_\rightarrow$ and all configurations $\omega \in \Omega$, the LIS $f^{(V,\omega)}$ defined by

$$f_\Lambda^{(V,\omega)}(\cdot | \xi) \triangleq f_\Lambda(\cdot | \omega_{V^c} \xi_V), \quad (6.4)$$

($\Lambda \in \mathcal{S}_b$, $\Lambda \subset V$, $\omega_{V^c} \xi_V \in \Omega$) admits a unique consistent chain. Likewise, a specification γ satisfies a HUC (for the family \mathcal{S}_\rightarrow) if for all sets $V \in \mathcal{S}_\rightarrow$ and all configurations $\omega \in \Omega$, the specification $\gamma^{(V,\omega)}$ defined by

$$\gamma_\Lambda^{(V,\omega)}(\cdot | \xi) \triangleq \gamma_\Lambda(\cdot | \omega_{V^c} \xi_V), \quad (6.5)$$

($\Lambda \in \mathcal{S}$, $\Lambda \subset V$, $\omega_{V^c} \xi_V \in \Omega$) admits a unique consistent measure.

All the LIS uniqueness conditions presented so far are in fact HUC. On the other hand, the Dobrushin and boundary-uniformity conditions for specifications are also HUC.

A specification γ is **continuous** on Ω if the functions $\Omega \ni \omega \longrightarrow \gamma_\Lambda(\sigma_\Lambda | \omega)$ are continuous. for all $\Lambda \in \mathcal{S}$ and all $\sigma_\Lambda \in \Omega_\Lambda$. It is **non-null** on Ω if $\gamma_\Lambda(\omega_\Lambda | \omega) > 0$ for each $\omega \in \Omega$ and $\Lambda \in \mathcal{S}$.

Here are the spaces we need.

LIS spaces:

$$\begin{aligned} \Theta &\triangleq \left\{ f \text{ continuous and non-null on } \Omega \right\} \\ \Theta_{\text{SUM}} &\triangleq \left\{ f \in \Theta : \sum_{j < 0} \delta_j(g) < +\infty \right\} \\ \Theta_{\text{HUC}} &\triangleq \left\{ f \in \Theta : f \text{ satisfies a HUC} \right\} \\ \Theta_{\text{EXP}} &\triangleq \left\{ f \in \Theta_{\text{SUM}} : \delta_j(g) \leq \text{const exp}(\alpha j) \right\} \end{aligned}$$

Specification spaces:

$$\begin{aligned}\Pi &\triangleq \left\{ \gamma \text{ continuous and non-null on } \Omega \right\} \\ \Pi_1 &\triangleq \left\{ \gamma \in \Pi : |\mathcal{G}(\gamma)| = 1 \right\} \\ \Pi_{\text{HUC}} &\triangleq \left\{ \gamma \in \Pi_1 : \gamma \text{ satisfies a HUC over all } [i, +\infty[, i \in \mathbb{Z} \right\} \\ \Pi_{\text{EXP}} &\triangleq \left\{ \gamma \in \Pi_{\text{HUC}} : \delta_j(\gamma) \leq \text{const exp}(-\alpha |j|) \right\}\end{aligned}$$

6.3 The results

Theorem 6.6 (LIS \rightsquigarrow specification) *There is a map*

$$\begin{aligned}b : \Theta_{\text{SUM}} &\rightarrow \Pi \\ f &\mapsto \gamma^f\end{aligned}$$

defined by

$$\gamma_{\Lambda}^f(\sigma_{\Lambda} \mid \omega) \triangleq \lim_{n \rightarrow +\infty} \frac{f_{[l_{\Lambda}, n]}(\sigma_{\Lambda} \omega_{m_{\Lambda}+1}^n \mid \omega_{< l_{\Lambda}})}{f_{[l_{\Lambda}, n]}(\omega_{\Lambda^c \cap [l_{\Lambda}, n]} \mid \omega_{< l_{\Lambda}})} \quad (6.7)$$

which satisfies

- (a) $\mathcal{G}(f) \subset \mathcal{G}(\gamma^f)$
- (b) b restricted to $b^{-1}(\Pi_1)$ is one-to-one.
- (c) If $f \in b^{-1}(\Pi_1)$, then $\mathcal{G}(f) = \mathcal{G}(\gamma^f) = \{\mu^f\}$, where μ^f is the only chain consistent with γ^f .

Theorem 6.8 (specification \rightsquigarrow LIS) *There is a map*

$$\begin{aligned}c : \Pi_{\text{HUC}} &\rightarrow \Theta_{\text{HUC}} \\ \gamma &\mapsto f^{\gamma}\end{aligned}$$

defined by

$$f_{\Lambda}^{\gamma}(A \mid \omega_{< l_{\Lambda}}) \triangleq \lim_{k \rightarrow +\infty} \gamma_{\Lambda \cup [m_{\Lambda}+1, m_{\Lambda}+k]}(A \mid \omega) \quad (6.9)$$

which enjoys the following properties:

- (a) $\mathcal{G}(f^{\gamma}) = \mathcal{G}(\gamma) = \{\mu^{\gamma}\}$, where μ^{γ} is the only Gibbs measure consistent with γ .

- (b) *The map c is one-to-one.*
- (c) *If γ satisfies the Dobrushin uniqueness condition, then so does f^γ*
- (d) *If γ satisfies the boundary-uniformity uniqueness condition, then so does f^γ .*

Finally we generalize the known correspondence between Markov transition probabilities and specifications.

Theorem 6.10 (LIS \rightsquigarrow specification)

- (a) *$b \circ c = \text{Id}$ over $c^{-1}(\Theta_{SUM})$, and $\mathcal{G}(f^\gamma) = \mathcal{G}(\gamma) = \{\mu^\gamma\}$*
- (b) *$c \circ b = \text{Id}$ over $b^{-1}(\Pi_{HUC})$ and $\mathcal{G}(\gamma^f) = \mathcal{G}(f) = \{\mu^f\}$*
- (c) *b and c establish a one-to-one correspondence between Θ_{EXP} and Π_{EXP} that preserves the consistent measure.*

7 Construction of a specification from singletons

To conclude we present a “(re)construction theorem” to build specifications starting from single-site kernels. This is the only instance in this work where we use the theory of LIS as an inspiration for the theory of specifications. The inspiration is, of course, Theorem 2.16 where LIS are constructed from singletons.

Our setting is general and not restricted to one dimension: There is an arbitrary alphabet space (E, \mathcal{E}) and a configuration space (Ω, \mathcal{F}) , where $\Omega \subset E^{\mathbb{Z}^d}$, for a given $d \geq 1$, and \mathcal{F} is the projection to Ω of $\mathcal{E}^{\mathbb{Z}^d}$. We are interested in specifications defined as densities with respect to natural alphabet measures [analogous to LIS of the form (2.6)]. That is, our initial objects are

- A family of *a priori* measures $(\lambda^i)_{i \in \mathbb{Z}^d}$,
- their products $\lambda^\Lambda \triangleq \bigotimes_{i \in \Lambda} \lambda^i$ for $\Lambda \subset \mathbb{Z}^d$, and

- the kernels over (Ω, \mathcal{F}) defined, for each $\Lambda \subset \mathbb{Z}^d$, by

$$\lambda_\Lambda(A \mid \omega) = (\lambda^\Lambda \otimes \delta_{\omega_{\Lambda^c}})(A) \quad (7.1)$$

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

The kernels λ_Λ do not form a specification. In particular the measures λ_i need not be normalized or even finite.

Theorem 7.2 *Let $(\gamma_i)_{i \in \mathbb{Z}^d}$ be a family of probability kernels on $\mathcal{F} \times \Omega_i$ such that*

1. *For each $i \in \mathbb{Z}^d$ and for some measurable function ρ_i ,*

$$\gamma_i = \rho_i \lambda_i$$

2. *The following properties hold:*

- (a) *Normalization: $(\lambda_i(\rho_i))(\omega) = 1$, for all $\omega \in \Omega$ and $i \in \mathbb{Z}^d$*
- (b) *Bounded-positivity: $\inf_\omega \lambda_j(\rho_j \rho_i^{-1})(\omega) > 0$ and $\sup_\omega \lambda_j(\rho_j \rho_i^{-1})(\omega) < +\infty$ for every $i, j \in \mathbb{Z}^d$*
- (c) *Order-consistency :*

$$\frac{\rho_i}{\lambda_i(\rho_i \rho_j^{-1})(\omega)} = \frac{\rho_j}{\lambda_j(\rho_j \rho_i^{-1})(\omega)} \quad (7.3)$$

for every $i, j \in \mathbb{Z}$ and every $\omega \in \Omega$

Then:

- (I) *There exists a unique family $\rho = \{\rho_\Lambda\}_{\Lambda \in \mathcal{S}}$ of positive measurable functions on Ω such that $\gamma \triangleq \{\rho_\Lambda \lambda_\Lambda\}_{\Lambda \in \mathcal{S}}$ is a specification with $\gamma_{\{i\}} = \gamma_i$ for each $i \in \mathbb{Z}^d$.*
- (II) *Such γ satisfies:*

- (i) $\mathcal{G}(\gamma) = \{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \mu \gamma_i = \mu \text{ for all } i \in \mathbb{Z}^d\}$
- (ii) *For all $\Lambda, \Gamma \in \mathcal{S}$ such that $\Gamma \subset \Lambda^c$.*

$$\rho_{\Lambda \cup \Gamma} = \frac{\rho_\Lambda}{\lambda_\Lambda(\rho_\Lambda \rho_\Gamma^{-1})}. \quad (7.4)$$

- (iii) For each $\Lambda \in \mathcal{S}$ there exist constants $C_\Lambda, D_\Lambda > 0$ such that $C_\Lambda \rho_k(\omega) \leq \rho_\Lambda(\omega) \leq D_\Lambda \rho_k(\omega)$ for all $k \in \Lambda$ and all $\omega \in \Omega$.
- (iv) If the functions ρ_i are continuous and $\int \sup_\omega (\rho_i \rho_j^{-1})(\sigma_i \omega_{\{i\}^c}) \lambda^i(d\sigma_i) < \infty$ for all $i, j \in \mathbb{Z}^d$, then the functions ρ_Λ , and thus the specification γ , are continuous.

The construction of the specification is done recursively through identity (7.4). The procedure recovers the corresponding generalizations of hypotheses (a)–(c) at each inductive step. This theorem is very similar to Theorem 1.33 in Georgii (1988). However, the latter is in fact a *reconstruction* result because singletons are supposed to come from a pre-existing specification. Our theorem shows that the strategy can be turned into a true construction algorithm, under the additional order-consistency condition (7.3). An alternative construction has been proposed by Dachian and Nahapetian (2001) for countable alphabets. Their algorithm relies on a purely pointwise order-consistency condition (ours is partially integrated on single sites).

8 Final remarks

The work reported here has touched only the very basic aspects of the link between chains and specifications. In particular, most of the results of Sections 4–6 either apply to the uniqueness regime, where not many surprises are expected, or involve Gibbsianness hypotheses. Therefore they are far from being useful in any non-Gibbsianness quest. We nevertheless believe that the issues raised are interesting *per se*, and justify the continuation of the program independently of any non-Gibbsian application. To conclude, we would like to list some (mostly obvious) questions suggested by the preceding results, which point directions for further work.

In Section 4 we used specification techniques to analyze LIS defining a unique chain. It would be natural to use similarly inspired tools to analyze coexistence regimes. In particular one should investigate a possible adaptation of arguments of the type used, for instance, by Fröhlich and Spencer (1982), and whether they are related to those employed to prove existence of multiple processes (Bramson and Kalikow, 1993; Berger, Hoffman and Sidravicius, 2003). In particular this could help to determine the optimality of various uniqueness criteria, complementing the result of the last reference.

The work of Section 6 leaves us with a number of tantalizing open questions. The main difficulty in establishing the domain and range of the natural maps proposed there is the worsening of the continuity bounds when passing from left- to right-conditioning, or from two-sided to left-only. Are these difficulties real, or only technical? Are there measures whose right continuity rates are very different from the left ones? An ultimate example would be a measure that is simultaneously a continuous process (i.e. with continuous past dependence) and non-Gibbsian (because of a discontinuous future dependence).

A possibly related issue is whether reasonably defined maps between LIS and specifications could end up connecting systems with different numbers of consistent measures. The possibility that γ^f have more consistent measures than f is left open in Theorem 6.6 above. Can this happen? If so, could one of the extra measures coincide on the left tail field with one of the common ones? Of course, for general cases with coexisting measures the very *definition* of maps like b or c of Section 6 is a delicate issue.

Finally, we mention two major aspects not treated at all in this work. On the one hand, there must be useful relations between large-deviation properties and variational approaches for processes and Gibbs measures. To our knowledge, the only example where relations of this sort have been exploited is the seminal work of Lebowitz, Maes and Speer (1990). A Gibbsian setting for probabilistic cellular automata is used in this reference to obtain a number of properties of these Markov processes with an uncountable alphabet. On the other hand, we have not explored ways in which process results and techniques could help understand measures defined via specifications. The pioneer work of Berbee (1989) remains a challenging benchmark.

Acknowledgements

The authors are grateful to the organizers of the workshop *Gibbs vs. non-Gibbs* for the opportunity to expose their work and for a very stimulating meeting.

References

- [Ber87] H. Berbee. Chains with infinite connections: Uniqueness and Markov representation. *Prob. Th. Rel. Fields*, 76:243–53, 1987.
- [Ber89] Henry Berbee. Uniqueness of Gibbs measures and absorption probabilities. *Ann. Probab.*, 17(4):1416–1431, 1989.
- [BFG9b] X. Bressaud, R. Fernández, and A. Galves. Decay of correlations for non Hölderian dynamics. a coupling approach. *Elect. J. Prob.*, 4, 1999b. (<http://www.math.washington.edu/~ejpecp/>).
- [BHS03] N. Berger, C. Hoffman, and V. Sidoravicius. Nonuniqueness for specifications in $\ell^{2+\epsilon}$. Preprint [arXiv:mat.PR/0312344](https://arxiv.org/abs/math.PR/0312344), 2003.
- [BK93] M. Bramson and S. A. Kalikow. Nonuniqueness in g -functions. *Israeli J. Math.*, 84:153–60, 1993.
- [BLP79] J. Bricmont, J. L. Lebowitz, and C.-E. Pfister. On the equivalence of boundary conditions. *J. Statist. Phys.*, 21(5):573–582, 1979.
- [BW99] P. Bühlmann and A. J. Wyner. Variable length Markov chains. *Ann. Statist.*, 27:480–513, 1999.
- [CFF02] F. Comets, R. Fernández, and P. A. Ferrari. Processes with long memory: regenerative construction and perfect simulation. *Ann. Appl. Probab.*, 12(3):921–943, 2002.
- [DN01] S. Dachian and B. S. Nahapetian. Description of random fields by means of one-point conditional distributions and some applications. *Markov Process. Related Fields*, 7(2):193–214, 2001.
- [FM3a] R. Fernández and G. Maillard. Chains with complete connections and one-dimensional Gibbs measures, 2003a. [math.PR/0305025](https://arxiv.org/abs/math.PR/0305025).
- [FM3b] R. Fernández and G. Maillard. Chains with complete connections. General theory, uniqueness, loss of memory and decay of correlations, 2003b. [math.PR/0305026](https://arxiv.org/abs/math.PR/0305026).
- [FS82] J. Fröhlich and T. Spencer. The Phase Transition in the One-Dimensional Ising Model with $\frac{1}{r^2}$ Interaction Energy. *Comm. Math. Phys.*, 84:87–101, 1982.

- [Geo74] H.-O. Georgii. Stochastische Felder und ihre Anwendung auf Interaktionssysteme. *Lecture Notes, Institut für Angewandte Mathematik, Universität Heidelberg*, 1974.
- [Geo88] H.-O. Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter (de Gruyter Studies in Mathematics, Vol. 9), Berlin–New York, 1988.
- [Har55] T. E. Harris. On chains of infinite order. *Pacific J. Math.*, 5:707–24, 1955.
- [Ios92] 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.
- [JO02] A. Johansson and A. Öberg. Square summability of variations of g -functions and uniqueness of g -measures. Preprint, 2002.
- [Kal90] S. Kalikow. Random Markov processes and uniform martingales. *Isr. J. Math.*, 71(1):33–54, 1990.
- [Kea72] M. Keane. Strongly mixing g -measures. *Inventiones Math.*, 16:309–24, 1972.
- [Lal86] S. P. Lalley. Regenerative representation for one-dimensional Gibbs states. *Ann. Prob.*, 14:1262–71, 1986.
- [Lal00] S. P. Lalley. Regeneration in one-dimensional Gibbs states and chains with complete connections. *Resenhas IME-USP*, 4:249–80, 2000.
- [LMS90] Joel L. Lebowitz, Christian Maes, and Eugene R. Speer. Statistical mechanics of probabilistic cellular automata. *J. Statist. Phys.*, 59(1-2):117–170, 1990.
- [Mai03] G. Maillard. *Chaînes à liaisons complètes et mesures de Gibbs unidimensionnelles*. PhD thesis, Université de Rouen, 2003.
- [OM35] O. Onicescu and G. Mihoc. Sur les chaînes statistiques. *C. R. Acad. Sci. Paris*, 200:511–12, 1935.

- [Raf85] A. Raftery. A model for high-order Markov chains. *J. R. Statist. Soc. B*, 47:528–539, 1985.
- [Sim93] B. Simon. *The Statistical Mechanics of Lattice Gases*. Series in physics. Princeton, 1993.
- [Sok81] A. D. Sokal. Existence of compatible families of proper regular conditional probabilities. *Z. Wahrsch. Verw. Gebiete*, 56(4):537–548, 1981.
- [Ste03] Ö Stenflo. Uniqueness in g -measures. *Nonlinearity*, 16(2):403–410, 2003.
- [vEFS93] 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.