Chains with complete connections and one-dimensinnal Gibbs measures

Gregory Maillard

THÈSE

en vue de l'obtention du titre de

Docteur de L'Université de Rouen

présentée par

GRÉGORY MAILLARD

Discipline	:	Mathématiques Appliquées
Spécialité	:	Probabilités

Chaînes à liaisons complètes et mesures de Gibbs unidimensionnelles

Date de soutenance : 26 juin 2003

Composition du Jury

Président	:	**C. Dellacherie,**	Directeur de Recherche au CNRS
Rapporteurs	:	**P. Collet,**	Professeur, École polytechnique
		A. Galves,	Professeur, Universidade de São Paulo
Examinateurs	:	**Y. Lacroix**,	Professeur, Université de Picardie
		J. de sam Lazaro,	Professeur, Université de Rouen
		B. Schmitt,	Professeur, Université de Dijon
Directeur de Thèse	:	**R. Fernández**,	Professeur, Université de Rouen

Thèse préparée à l'Université de Rouen

Laboratoire de Mathématiques Raphaël Salem, UMR-CNRS 6085

Abstract

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 Dobrushin criterion in statistical mechanics. (iii) Results on loss of memory and mixing properties for chains in the Dobrushin regime. These results are complementary of those existing in the literature, and generalize the Markovian results based on the Dobrushin ergodic coefficient. We discuss the relationship between discrete-time processes (chains) and one-dimensional Gibbs measures. On the other hand, we consider finite-alphabet (finite-spin) systems, possibly with a grammar (exclusion rule). We establish conditions for a stochastic process to define a Gibbs measure and vice versa. Our conditions generalize well known equivalence results between ergodic Markov chains and fields, as well as the known Gibbsian character of processes with exponential continuity rate. Our arguments are purely probabilistic; they are based on the study of regular systems of conditional probabilities (specifications). Furthermore, we discuss the equivalence of uniqueness criteria for chains and fields and we establish bounds for the continuity rates of the respective systems of finite-volume conditional probabilities. As an auxiliary result we prove a (re)construction theorem for specifications starting from single-site conditioning, which applies in a more general setting (general spin space, specifications not necessarily Gibbsian).

Contents

Introduction

Le terme “chaînes à liaisons complètes” a été créé par Onicescu et Mihoc (1935), puis exploité par Doeblin et Fortet (1937), pour les processus stochastiques discrets dont la dépendance sur le passé n’est pas nécessairement markovienne. La théorie de ces processus a beaucoup de points communs avec la théorie des mesures de Gibbs en mécanique statistique, particulièrement l’existence de transitions de phases. Néanmoins il y a une différence claire, au niveau formel, entre les deux théories. En effet, les processus stochastiques sont définis sur la base de transitions de probabilités sur un site. Une chaîne cohérente est celle pour laquelle sa probabilité est une réalisation des probabilités conditionnelles sur un site sachant le passé. Une mesure de Gibbs est définie en termes de spécifications qui déterminent ses probabilités conditionnelles sur un volume fini sachant l’extérieur du volume. En dimension un, ceci implique un conditionnement à la fois sur le passé et sur le futur. Dans ce travail, on propose une manière naturelle de réduire cette asymétrie, en introduisant un cadre issu de la mécanique statistique pour l’étude des processus. Ce cadre établit une relation plus directe entre les deux théories, ce qui nous permet d’une part de reproduire, pour les chaînes à liaisons complètes (appelées tout simplement chaînes), un certain nombre de résultats gibbsiens de référence (chapitres 4–5) et d’autre part d’établir des liens directs entre processus et mesures de Gibbs (chapitre 6).

Dans le chapitre 1, on introduit toutes les notations qui serviront dans la suite. Le chapitre 2 est consacré à une étude comparative des différentes approches existantes de la notion de processus non-markoviens (chaînes à liaisons complètes, chaînes d’ordre infini, g-mesures, processus de Markov aléatoires, martingales uniformes). On passe ensuite en revue les différentes conditions sur les transitions qui assurent l’unicité du processus cohérent avec ces transitions (Harris, 1955; Iosifescu et Spataru, 1973; Walters, 1975; Berbee, 1987; Hulse, 1997; Stenflo, 2002; Johansson et Öberg, 2002), ainsi que les conditions qui assurent une vitesse de mélange (Bowen 1975; Bressaud et al, 1999). Le chapitre 2, quant à lui, constitue un survol de la théorie générale des spécifications et des mesures de Gibbs : propriétés des mesures extrémales, corrélations à courtes portées, limites en volumes infinis et ergodicité. On s’intéresse également aux conditions sur les spécifications assurant l’unicité de la mesure cohérente, ainsi qu'une vitesse de perte de mémoire et de mélange. Enfin, on montre qu'un système de noyaux sur un site satisfaisant des propriétés de normalisation, de bornitude et de cohérence d'ordre par rapport à une mesure a priori, peuvent être étendus, de facon unique, à une spécification toute entière (Théorème 3.1.8). Ce théorème est établi dans un cadre général, pour un espace de spins arbitraires et pour des noyaux non nécessairement gibbsiens. Il généralise le théorème de reconstruction (Théorème 1.33 dans Georgii (1988)).

Afin d'adapter le cadre de mécanique statistique à la théorie des processus, notre approche est basée sur une notion analogue à celle de spécification (en mécanique statistique), que nous appelons *left interval-specification* (LIS). Ce sont des noyaux, pour des régions ayant la forme d'intervalles, qui dépendent de l'histoire précédente du processus. En contraste, les spécifications nécessitent des régions finies arbitraires et dépendent de la configuration sur l'extérieur tout entier de la région. Ceci revient, en dimension un, à une dépendance à la fois sur le passé et sur le futur. La différence est, bien sûr, une conséquence du caractère "un côté" associé à l'évolution (du temps) stochastique, par rapport au manque de direction favorite dans la description spatiale fournie par une mesure de Gibbs.

La description en termes de LIS est totalement équivalente à la description traditionnelle en termes de probabilités de transition (Théorème 4.2.1). Mais comme les chapitres 4 et 5 l'illustrent, notre approche a l'avantage de nous permettre "d'importer", d'une manière naturelle, des notions, des techniques et des arguments provenant de mécanique statistique. Il peut alors être utile dans la direction opposée, d'explorer les conséquences des propriétés connues des chaînes sur la théorie des mesures de Gibbs. Afin d'aller dans cette direction, nous étudions, au cours du chapitre 6, les conditions sous lesquelles les chaînes et les mesures de Gibbs peuvent être identifiées.

Les résultats des chapitres 4–5 sont regroupés dans Fernández et Maillard (2003). On y présente trois types de résultats. Premièrement, on obtient les propriétés générales des chaînes extrémales pour tout type d'alphabet, c'est-à-dire la trivialité sur la tribu queue (Théorème 4.2.6), les corrélations à courte portée (Théorème 4.2.7), la réalisation du processus via des limites en volumes infinis (Théorème 4.2.8) et l'ergodicité (Théorème 4.2.9) . Deuxièmement, on produit de nouvelles conditions suffisantes pour l'unicité de la chaîne cohérente. D'une part, on obtient une transcription d'un critère donné par Georgii (1974) pour les champs gibbsiens de dimension un (Théorème 5.1.1). Pour ces derniers, ce critère est connu comme étant optimal, dans le sens qu'il souligne l'absence de transition de phase pour les modèles de spin à deux corps avec une interaction de $1/r^{2+\varepsilon}$, quelque soit $\varepsilon > 0$. Le critère n'impose pas de restriction sur le type d'alphabet. D'autre part, on prouve un critère de Dobrushin "un côté" (Théorème 5.1.19), qui cor-respond à un résultat d'unicité bien connu en mécanique statistique (voir Théorème 3.3.6, ou pour plus de détails, Simon, 1993, chapitre V). Ce critère est valide pour des systèmes avec un alphabet métrique. On présente de simples exemples de chaînes pour lesquels le critère de Dobrushin s'applique mais qui sortent du champ d'application de la plupart des autres critères d'unicité connus (Harris, 1955; Iosifescu et Spataru, 1973; Walters, 1975; Berbee, 1987; Stenflo, 2002; Johansson et Öberg, 2002). On discute également d'une légère amélioration d'un critère dû à Hulse (1997) (Théorème 5.1.15), qui est plus fort que Dobrushin pour un alphabet fini et est également valide pour des alphabets dénombrables.

Notre troisième type de résultat concerne la perte de mémoire et les propriétés de mélange des chaînes dans le régime de Dobrushin. Nos résultats, obtenus sur les bases d'une théorie gibbsienne similaire (voir Théorème 3.4, et à nouveau Simon, 1993), sont complémentaires, à la fois dans leur précision et dans leur champ d'applicabilité, d'autres résultats disponibles dans la littérature (Iosifescu, 1992; Bressaud, Fernández et Galves, 1999 et les références qu'ils contiennent). Les résultats dépendent d'une matrice de sensibilité qui généralise le coefficient d'ergodicité de Dobrushin pour les chaînes de Markov.

Les systèmes de dimension un sont simultanément l'objet de la théorie des processus stochastiques et de la théorie des mesures de Gibbs. La complémentarité des deux approches n'a pas encore été pleinement exploitée. Dans le chapitre 6, nous étudions les conditions sous lesquelles un processus stochastique définit en fait une mesure de Gibbs, et dans la direction opposée, lorsqu'une mesure de Gibbs peut être vue comme un processus.

Ce type de questions a été complètement élucidé pour les processus et les champs markoviens. Voir, par exemple, le chapitre 11 du livre de Georgii (1988). L'équivalence, cependant, est obtenue par des considérations de valeurs propres et de vecteurs propres qui ne sont pas facilement applicables aux processus non markoviens. Le caractère gibbsien des processus avec une décroissance exponentielle du taux de continuité est également connu. Il découle de la caractérisation de Bowen des mesures de Gibbs (Théorème 5.2.4. dans Keller, 1998, par exemple). Aucun résultat ne semble être établi dans la direction opposée, c'est-à-dire sur la caractérisation d'une mesure de Gibbs unidimensionnelle, ayant une interaction exponentiellement sommable, en tant que processus stochastique.

Les résultats du chapitre 6 sont rassemblés dans Fernández et Maillard (2003). On y présente à la fois une généralisation et une alternative de ce travail précédent. On établit directement une correspondance entre les spécifications et les LIS. On considère des systèmes avec un alphabet fini, pouvant avoir une grammaire, c'est-à-dire avec des règles d'exclusion telles que l'ensemble des configurations non exclues forme un ensemble compact. Nous ne supposons pas d'invariance translationnelle que ce soit pour les noyaux ou bien pour les mesures cohérentes. La principale limitation dans nos résultats est que, afin d'assurer que les limites nécessaires soient définies de manière unique, les spécifications et les processus doivent satisfaire une condition d'unicité forte appelée *condition d'unicité héréditaire* (HUC). Une seconde propriété, appelée *bon futur* (GF) est imposée aux processus stochastiques afin de garantir un contrôle sur le conditionnement par rapport au futur. Pour des noyaux de probabilités invariants par le shift, cette seconde condition est plus forte que la première. HUC est vérifiée, par exemple, pour les spécifications satisfaisant Dobrushin et le critère d'uniforme bornitude (donné par le Corollaire 5.1.28). Les propriétés GF et HUC sont toutes deux satisfaites par une grande famille de processus, par exemple par les chaînes ayant des variations sommables étudiées par Harris (1955), Ledrappier (1974), Walters (1975), Lalley (1986), Berbee (1987), Bressaud et al (1999),...

Nos résultats montrent que sous ces hypothèses il existe : (i) une application qui pour chaque LIS associe une spécification telle que le processus consistant avec la première est une mesure de Gibbs cohérente avec la seconde (Théorème 6.3.12), et (ii) une application qui à chaque spécification associe une LIS telle que la mesure de Gibbs cohérente avec la première est un processus consistant avec seconde (Théorème 6.3.17). Si le domaine de définition et le domaine image coïncident, ces applications sont les inverses l'une de l'autre. Ceci a lieu en particulier, dans le cas d'un taux de décroissance de continuité exponentiel (Théorème 6.3.20). Comme partie des preuves, on obtient des estimateurs liant les taux de continuité pour les LIS et les spécifications associées par ces applications (Théorème 6.3.19). On montre aussi que la validité des critères de Dobrushin, respectivement de bornitude uniforme, pour une spécification entraîne la validité de critères analogues pour la LIS associée (Théorème 6.3.18).

L'approche des chaînes par l'intermédiaire des LIS offre de nouvelles perspectives non seulement sur les relations entre les processus stochastiques et les champs aléatoires, mais aussi sur l'étude des propriétés des chaînes. Notre travail a fourni quelques réponses, mais de nombreuses questions restent en suspens.

Sur les chaînes :

- Liens manquant entre les différents critères d'unicité
- Optimalité des différents critéres d'unicité
- Propriétés de mélange hors cadres connus
- Théorèmes limites et grandes déviations

Sur la relation chaînes-champs :

- Lorsque l'on a des transitions de phases
- Lorsque l'on a un alphabet infini...

Introduction

Le terme “chaînes à liaisons complètes” a été créé par Onicescu et Mihoc (1935), puis exploité par Doeblin et Fortet (1937), pour les processus stochastiques discrets dont la dépendance sur le passé n’est pas nécessairement markovienne. La théorie de ces processus a beaucoup de points communs avec la théorie des mesures de Gibbs en mécanique statistique, particulièrement l’existence de transitions de phases. Néanmoins il y a une différence claire, au niveau formel, entre les deux théories. En effet, les processus stochastiques sont définis sur la base de transitions de probabilités sur un site. Une chaîne cohérente est celle pour laquelle sa probabilité est une réalisation des probabilités conditionnelles sur un site sachant le passé. Une mesure de Gibbs est définie en termes de spécifications qui déterminent ses probabilités conditionnelles sur un volume fini sachant l’extérieur du volume. En dimension un, ceci implique un conditionnement à la fois sur le passé et sur le futur. Dans ce travail, on propose une manière naturelle de réduire cette asymétrie, en introduisant un cadre issu de la mécanique statistique pour l’étude des processus. Ce cadre établit une relation plus directe entre les deux théories, ce qui nous permet d’une part de reproduire, pour les chaînes à liaisons complètes (appelées tout simplement chaînes), un certain nombre de résultats gibbsiens de référence (chapitres 4–5) et d’autre part d’établir des liens directs entre processus et mesures de Gibbs (chapitre 6).

Dans le chapitre 1, on introduit toutes les notations qui serviront dans la suite. Le chapitre 2 est consacré à une étude comparative des différentes approches existantes de la notion de processus non-markoviens (chaînes à liaisons complètes, chaînes d’ordre infini, g-mesures, processus de Markov aléatoires, martingales uniformes). On passe ensuite en revue les différentes conditions sur les transitions qui assurent l’unicité du processus cohérent avec ces transitions (Harris, 1955; Iosifescu et Spataru, 1973; Walters, 1975; Berbee, 1987; Hulse, 1997; Stenflo, 2002; Johansson et Öberg, 2002), ainsi que les conditions qui assurent une vitesse de mélange (Bowen 1975; Bressaud et al, 1999). Le chapitre 2, quant à lui, constitue un survol de la théorie générale des spécifications et des mesures de Gibbs : propriétés des mesures extrémales, corrélations à courtes portées, limites en volumes infinis et ergodicité. On s’intéresse également aux conditions sur les spécifications assurant l’unicité de la mesure cohérente, ainsi qu'une vitesse de perte de mémoire et de mélange. Enfin, on montre qu'un système de noyaux sur un site satisfaisant des propriétés de normalisation, de bornitude et de cohérence d'ordre par rapport à une mesure a priori, peuvent être étendus, de facon unique, à une spécification toute entière (Théorème 3.1.8). Ce théorème est établi dans un cadre général, pour un espace de spins arbitraires et pour des noyaux non nécessairement gibbsiens. Il généralise le théorème de reconstruction (Théorème 1.33 dans Georgii (1988)).

Afin d'adapter le cadre de mécanique statistique à la théorie des processus, notre approche est basée sur une notion analogue à celle de spécification (en mécanique statistique), que nous appelons *left interval-specification* (LIS). Ce sont des noyaux, pour des régions ayant la forme d'intervalles, qui dépendent de l'histoire précédente du processus. En contraste, les spécifications nécessitent des régions finies arbitraires et dépendent de la configuration sur l'extérieur tout entier de la région. Ceci revient, en dimension un, à une dépendance à la fois sur le passé et sur le futur. La différence est, bien sûr, une conséquence du caractère "un côté" associé à l'évolution (du temps) stochastique, par rapport au manque de direction favorite dans la description spatiale fournie par une mesure de Gibbs.

La description en termes de LIS est totalement équivalente à la description traditionnelle en termes de probabilités de transition (Théorème 4.2.1). Mais comme les chapitres 4 et 5 l'illustrent, notre approche a l'avantage de nous permettre "d'importer", d'une manière naturelle, des notions, des techniques et des arguments provenant de mécanique statistique. Il peut alors être utile dans la direction opposée, d'explorer les conséquences des propriétés connues des chaînes sur la théorie des mesures de Gibbs. Afin d'aller dans cette direction, nous étudions, au cours du chapitre 6, les conditions sous lesquelles les chaînes et les mesures de Gibbs peuvent être identifiées.

Les résultats des chapitres 4–5 sont regroupés dans Fernández et Maillard (2003). On y présente trois types de résultats. Premièrement, on obtient les propriétés générales des chaînes extrémales pour tout type d'alphabet, c'est-à-dire la trivialité sur la tribu queue (Théorème 4.2.6), les corrélations à courte portée (Théorème 4.2.7), la réalisation du processus via des limites en volumes infinis (Théorème 4.2.8) et l'ergodicité (Théorème 4.2.9) . Deuxièmement, on produit de nouvelles conditions suffisantes pour l'unicité de la chaîne cohérente. D'une part, on obtient une transcription d'un critère donné par Georgii (1974) pour les champs gibbsiens de dimension un (Théorème 5.1.1). Pour ces derniers, ce critère est connu comme étant optimal, dans le sens qu'il souligne l'absence de transition de phase pour les modèles de spin à deux corps avec une interaction de $1/r^{2+\varepsilon}$, quelque soit $\varepsilon > 0$. Le critère n'impose pas de restriction sur le type d'alphabet. D'autre part, on prouve un critère de Dobrushin "un côté" (Théorème 5.1.19), qui cor-respond à un résultat d'unicité bien connu en mécanique statistique (voir Théorème 3.3.6, ou pour plus de détails, Simon, 1993, chapitre V). Ce critère est valide pour des systèmes avec un alphabet métrique. On présente de simples exemples de chaînes pour lesquels le critère de Dobrushin s'applique mais qui sortent du champ d'application de la plupart des autres critères d'unicité connus (Harris, 1955; Iosifescu et Spataru, 1973; Walters, 1975; Berbee, 1987; Stenflo, 2002; Johansson et Öberg, 2002). On discute également d'une légère amélioration d'un critère dû à Hulse (1997) (Théorème 5.1.15), qui est plus fort que Dobrushin pour un alphabet fini et est également valide pour des alphabets dénombrables.

Notre troisième type de résultat concerne la perte de mémoire et les propriétés de mélange des chaînes dans le régime de Dobrushin. Nos résultats, obtenus sur les bases d'une théorie gibbsienne similaire (voir Théorème 3.4, et à nouveau Simon, 1993), sont complémentaires, à la fois dans leur précision et dans leur champ d'applicabilité, d'autres résultats disponibles dans la littérature (Iosifescu, 1992; Bressaud, Fernández et Galves, 1999 et les références qu'ils contiennent). Les résultats dépendent d'une matrice de sensibilité qui généralise le coefficient d'ergodicité de Dobrushin pour les chaînes de Markov.

Les systèmes de dimension un sont simultanément l'objet de la théorie des processus stochastiques et de la théorie des mesures de Gibbs. La complémentarité des deux approches n'a pas encore été pleinement exploitée. Dans le chapitre 6, nous étudions les conditions sous lesquelles un processus stochastique définit en fait une mesure de Gibbs, et dans la direction opposée, lorsqu'une mesure de Gibbs peut être vue comme un processus.

Ce type de questions a été complètement élucidé pour les processus et les champs markoviens. Voir, par exemple, le chapitre 11 du livre de Georgii (1988). L'équivalence, cependant, est obtenue par des considérations de valeurs propres et de vecteurs propres qui ne sont pas facilement applicables aux processus non markoviens. Le caractère gibbsien des processus avec une décroissance exponentielle du taux de continuité est également connu. Il découle de la caractérisation de Bowen des mesures de Gibbs (Théorème 5.2.4. dans Keller, 1998, par exemple). Aucun résultat ne semble être établi dans la direction opposée, c'est-à-dire sur la caractérisation d'une mesure de Gibbs unidimensionnelle, ayant une interaction exponentiellement sommable, en tant que processus stochastique.

Les résultats du chapitre 6 sont rassemblés dans Fernández et Maillard (2003). On y présente à la fois une généralisation et une alternative de ce travail précédent. On établit directement une correspondance entre les spécifications et les LIS. On considère des systèmes avec un alphabet fini, pouvant avoir une grammaire, c'est-à-dire avec des règles d'exclusion telles que l'ensemble des configurations non exclues forme un ensemble compact. Nous ne supposons pas d'invariance translationnelle que ce soit pour les noyaux ou bien pour les mesures cohérentes. La principale limitation dans nos résultats est que, afin d'assurer que les limites nécessaires soient définies de manière unique, les spécifications et les processus doivent satisfaire une condition d'unicité forte appelée *condition d'unicité héréditaire* (HUC). Une seconde propriété, appelée *bon futur* (GF) est imposée aux processus stochastiques afin de garantir un contrôle sur le conditionnement par rapport au futur. Pour des noyaux de probabilités invariants par le shift, cette seconde condition est plus forte que la première. HUC est vérifiée, par exemple, pour les spécifications satisfaisant Dobrushin et le critère d'uniforme bornitude (donné par le Corollaire 5.1.28). Les propriétés GF et HUC sont toutes deux satisfaites par une grande famille de processus, par exemple par les chaînes ayant des variations sommables étudiées par Harris (1955), Ledrappier (1974), Walters (1975), Lalley (1986), Berbee (1987), Bressaud et al (1999),...

Nos résultats montrent que sous ces hypothèses il existe : (i) une application qui pour chaque LIS associe une spécification telle que le processus consistant avec la première est une mesure de Gibbs cohérente avec la seconde (Théorème 6.3.12), et (ii) une application qui à chaque spécification associe une LIS telle que la mesure de Gibbs cohérente avec la première est un processus consistant avec seconde (Théorème 6.3.17). Si le domaine de définition et le domaine image coïncident, ces applications sont les inverses l'une de l'autre. Ceci a lieu en particulier, dans le cas d'un taux de décroissance de continuité exponentiel (Théorème 6.3.20). Comme partie des preuves, on obtient des estimateurs liant les taux de continuité pour les LIS et les spécifications associées par ces applications (Théorème 6.3.19). On montre aussi que la validité des critères de Dobrushin, respectivement de bornitude uniforme, pour une spécification entraîne la validité de critères analogues pour la LIS associée (Théorème 6.3.18).

L'approche des chaînes par l'intermédiaire des LIS offre de nouvelles perspectives non seulement sur les relations entre les processus stochastiques et les champs aléatoires, mais aussi sur l'étude des propriétés des chaînes. Notre travail a fourni quelques réponses, mais de nombreuses questions restent en suspens.

Sur les chaînes :

- Liens manquant entre les différents critères d'unicité
- Optimalité des différents critéres d'unicité
- Propriétés de mélange hors cadres connus
- Théorèmes limites et grandes déviations

Sur la relation chaînes-champs :

- Lorsque l'on a des transitions de phases
- Lorsque l'on a un alphabet infini...

Chapter 1

General notation

1.1 Measure-theoretical notions

Let us introduce some general notations for measures and measure kernels. Let $(X, \mathcal{X})$ be a measurable space. We denote respectively by $\mathcal{M}(X, \mathcal{X})$ and $\mathcal{P}(X, \mathcal{X})$ the sets of positive measures and of probability measures on $(X, \mathcal{X})$. Let μ be a measure on $(X, \mathcal{X})$, h, $h_1 \geq 0$, h_2 $\mathcal{X}$-measurable functions and $\mathcal{B}$ a sub-σ-algebra of $\mathcal{X}$. We denote (provided that the integrals exist) the integral of h with respect to μ by

$$\mu(h) = \int h \, d\mu, \tag{1.1.1}$$

the conditional μ-expectation of h given $\mathcal{B}$

$$\mu \left(h \mid \mathcal{B} \right) = \mathrm{E}_{\mu} \left(h \mid \mathcal{B} \right), \tag{1.1.2}$$

and the measure $h_1 \, \mu$ defined on $(X, \mathcal{X})$ by

$$(h_1 \, \mu) \, (h_2) = \mu \, (h_1 \, h_2) \, . \tag{1.1.3}$$

Let us consider a second measurable space $(Y, \mathcal{Y})$. A function $\pi : \mathcal{Y} \times X \to \mathbb{R}^+$ is a **kernel** on $\mathcal{Y} \times X$ if

(a) For each $x \in X$, $\pi \left(\cdot \mid x \right)$ is a measure on $(X, \mathcal{X})$.

(b) For each $A \in \mathcal{Y}$, $\pi \left(A \mid \cdot \right)$ is $\mathcal{X}$-measurable.

If moreover $\pi\left(Y \mid \cdot\right) = 1$ then π is a **probability kernel** on $\mathcal{Y} \times X$. For π a kernel on $\mathcal{Y} \times X$, h a $\mathcal{Y}$-measurable function and μ a measure on $(X, \mathcal{X})$, we denote (provided that the integrals exist) $\pi\, h$ the $\mathcal{X}$-measurable function defined by

$$\pi\, h = \pi\left(h \mid \cdot\right) = \int h(y)\, \pi\left(dy \mid \cdot\right) \tag{1.1.4}$$

and $\mu\, \pi$ the measure on $(Y, \mathcal{Y})$ defined by

$$\left(\mu\, \pi\right)(h) = \int \pi\left(h \mid x\right)\, \mu(dx) = \iint h(y)\, \pi\left(dy \mid x\right)\, \mu(dx). \tag{1.1.5}$$

For h_1, h_2 $\mathcal{X}$-measurable functions we denote $h_1\, \pi$ the kernel on $\mathcal{X} \times Y$ defined by

$$\left(h_1\, \pi\right)\left(h_2 \mid y\right) = \pi\left(h_1\, h_2 \mid y\right). \tag{1.1.6}$$

Let us consider a third measurable space $(Z, \mathcal{Z})$. The composition $\pi_1\, \pi_2$ of π_1, a kernel on $\mathcal{Y} \times X$, and π_2, a kernel on $\mathcal{Z} \times Y$, is a kernel on $\mathcal{Z} \times X$ defined by

$$\pi_1\, \pi_2\left(A \mid x\right) = \int \pi_2\left(A \mid y\right)\, \pi_1\left(dy \mid x\right) \tag{1.1.7}$$

for all $A \in \mathcal{Z}$ and all $x \in X$.

Let P and Q be two probability measures on $(X, \mathcal{X})$. The **variational norm** between P and Q is defined by one of three following equalities

$$\begin{aligned} \|P - Q\| &= \max\left\{\, |P(A) - Q(A)| : A \in \mathcal{X} \,\right\} \\ &= \max\left\{ \frac{|P(h) - Q(h)|}{\mathrm{Spr}(h)} : h \text{ bounded and } \mathcal{X} - \text{measurable} \right\} \\ &= \frac{R\left(|h_1 - h_2|\right)}{2}, \end{aligned} \tag{1.1.8}$$

where $\mathrm{Spr}(h) = \sup_{x \in X} h(x) - \inf_{x \in X} h(x)$ is the **spread** of h and R is an arbitrary measure such that $dP/dR = h_1$ and $dQ/dR = h_2$. In the particular case where X is countable, the variational norm between P and Q can be defined by

$$\|P - Q\| = \frac{1}{2} \sum_{x \in X} |P(x) - Q(x)|. \tag{1.1.9}$$

A proof of these results can be found in Georgii (1988) [section 8.1].

For each metric space (X, d), the **Vaserstein-Kantorovich-Rubinstein norm (VKR)** relative to the space (X, d) is defined by

$$\begin{aligned} \|P - Q\|_d &= \sup\left\{ \left| \int h\, d(P - Q) \right| : \|h\|_L \leq 1 \right\} \qquad (1.1.10) \\ &= \inf\left\{ \int d(x, y)\, d\mu(x, y) : \mu \in \mathcal{P}(X \times X) \text{ with marginals } P \text{ and } Q \right\}, \end{aligned}$$

where $\|h\|_L \triangleq \sup\{|h(x)-h(y)|/d(x,y) : x \neq y \text{ in } X\}$ is the Lipschitz seminorm for suitable real-valued functions h on X.
A general discussion of these metrics can be found in Dudley (2002)[section 11.8].

1.2 Configuration spaces

We will consider fields on a lattice, that is a measure space, called **the configuration space**, of the form $E^{\mathbb{L}}$, where $(E, \mathcal{E})$ is a measurable space and $\mathbb{L}$ a countable set (the "lattice"), whose elements are called **sites**. An element $\omega \in E^{\mathbb{L}}$ is called a **configuration**, and each value $\omega_j \in E$, the **spin at site** j. The space E is sometimes called *alphabet*. Furthermore, we will consider a subset $\Omega \subset E^{\mathbb{L}}$, of **allowed configurations**. The space Ω is endowed with the projection $\mathcal{F}$ of the product σ-algebra associated to $E^{\mathbb{L}}$. 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

- We denote $\mathcal{S}$ the set of finite subsets of $\mathbb{L}$. For each $\Lambda \in \mathcal{S}$ and each configuration $\sigma \in E^{\mathbb{L}}$ we denote σ_Λ its projection on Λ, namely the family $(\sigma_i)_{i\in\Lambda} \in E^\Lambda$. The set of admissible configurations in $\Lambda \in \mathcal{S}$ is

$$\Omega_\Lambda \triangleq \{\sigma_\Lambda \in E^\Lambda : \exists\, \omega \in \Omega \text{ with } \omega_\Lambda = \sigma_\Lambda\}, \tag{1.2.1}$$

  while $\mathcal{F}_\Lambda$ is the sub-σ-algebra of $\mathcal{F}$ generated by the cylinders with base in Ω_Λ. If $\Lambda,\, \Delta \in \mathcal{S}$ 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 $i \in \Delta$.

- For kernels associated to a specification, for V an infinite subset of $\mathbb{L}$, $\lim_{\Lambda\uparrow V} \gamma_\Lambda$ is the limit of the net $\{\gamma_\Lambda, \{\Lambda\}_{\Lambda\in\mathcal{S},\, \Lambda\subset V}, \subset\}$ directed by inclusion. To lighten up formulas involving probability kernels, we will freely use $\rho(\sigma_\Lambda)$ instead of $\rho(\{\omega \in \Omega : \omega_\Lambda = \sigma_\Lambda\})$ for ρ a measure on Ω, $\Lambda \subset \mathbb{Z}$ and $\sigma_\Lambda \in \Omega_\Lambda$.

When $\mathbb{L}$ is a countable set with total order, we can identify $\mathbb{L}$ with $\mathbb{Z}$.

- 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 $n \in \mathbb{Z}$, $\mathcal{F}_{\leq n} \triangleq \mathcal{F}_{]-\infty,n]}$. For every $\Lambda \in \mathcal{S}_b$ we denote $l_\Lambda \triangleq \min\Lambda$ and $m_\Lambda \triangleq \max\Lambda$, $\Lambda_- =]-\infty, l_\Lambda - 1]$, $\Lambda_+ = [m_\Lambda + 1, +\infty[$ and $\Lambda_+^{(k)} = [m_\Lambda + 1, m_\Lambda + k]$ for all $k \in \mathbb{N}^*$.

- To measure the dependence of a function with respect to a part of its support, let us denote for all $\omega, \sigma \in \mathcal{A}_{-\infty}^{i}$,

$$\omega \stackrel{k}{=} \sigma \quad \Longleftrightarrow \quad \omega_j = \sigma_j, \ \forall \ j \in \{-k+i, \ldots, i\} \tag{1.2.2}$$

for sequences which have in common the k last letters and

$$\omega \stackrel{\neq k}{=} \sigma \quad \Longleftrightarrow \quad \omega_j = \sigma_j, \ \forall \ j \neq k \tag{1.2.3}$$

for sequences which are equal everywhere off k. Then for every $\mathcal{F}_{\leq i}$-measurable function h, we define the **variation** of h by

$$\mathrm{var}_k(h) \triangleq \sup \left\{ |h(\omega) - h(\sigma)| : \omega, \ \sigma \in \mathcal{A}_{-\infty}^{i}, \ \omega \stackrel{k}{=} \sigma \right\} \tag{1.2.4}$$

and the **oscillation** of h by

$$\delta_k(h) \triangleq \sup \left\{ |h(\omega) - h(\sigma)| : \omega, \ \sigma \in \mathcal{A}_{-\infty}^{i}, \ \omega \stackrel{\neq k}{=} \sigma \right\}. \tag{1.2.5}$$

- For kernels associated to a LIS (defined in Chapter 4), $\lim_{\Lambda \uparrow V} f_\Lambda$ is the limit of the net $\{f_\Lambda, \{\Lambda\}_{\Lambda \in \mathcal{S}_b}, \ \Lambda \subset V, \subset\}$, for V an infinite interval of $\mathbb{Z}$.

At the beginning of each chapter we will always clarify the particular setting in which we will work.

Chapter 2

Chains with complete connections

We consider non-Markovian processes, that is processes whose transition probabilities depend on the whole past history. These processes can be found in the literature under different appellations. They were first introduced by Onicescu and Mihoc (1935) under the name **chains with complete connections (chaînes à liaisons complètes)** to study urn models. Doeblin and Fortet (1937) proved the first results on the existence of the invariant measure. Harris (1955) extended existence results and proved one of the weakest uniqueness condition available. He called these processes **chains of infinite order**. Non-Markovian processes appeared also in the formalism introduced by Keane (1972) to study covering transformations. In his theory, transition probabilities are called **g-functions** and invariant measures **g-measures**. Another approach is due to Kalikow (1990) who introduced **random Markov processes** as generalizations of n-step Markov chains. He also defined the concept of bounded **uniform martingale** and studied its ergodic properties. In the following we will see that, at least in the translation invariant setting, all these notions in fact coincide.

In this chapter $E = \mathcal{A}$ is a finite alphabet. We will relax this hypothesis in chapters 3–5. Here we assume that $\Omega = \mathcal{A}^{\mathbb{Z}}$.

Transition probabilities, g-functions, n-step Markov chains and uniform martingales are functions measurable only with respect to the "past" and the "present". To state the necessary definitions we need to introduce the following notation: Each configuration of the past $\omega_{-\infty}^{-1} \in \mathcal{A}_{-\infty}^{-1}$ is denoted by $\underline{\omega}$. Double-infinite sequences are denoted without sub or superscripts, $\omega \in \mathcal{A}^{\mathbb{Z}}$.

2.1 Chains with complete connections

Chains with complete connections are induced by conditional probabilities of the form

$$P\left(X_n = \omega_n \;\middle|\; X_{-\infty}^{n-1} = \omega_{-\infty}^{n-1}\right). \tag{2.1.1}$$

These transition probabilities appear to be a extension of the notion of k-step Markov chain with an infinite k. These objects must be taken with some precautions because, in the non-Markovian case, the conditioning is always on an event of probability zero.

Definition 2.1.2

A **system of transition probabilities** *is a family* $\{P_n\}_{n\in\mathbb{Z}}$ *of functions with* $P_n(\,\cdot|\cdot\,):\mathcal{A}\times\mathcal{A}_{-\infty}^{n-1}\longrightarrow[0,1]$, *such that the following conditions hold for every* $n\in\mathbb{Z}$

(a) *For every* $\omega_n\in\mathcal{A}$ *the function* $P_n(\omega_n \mid \cdot\,)$ *is measurable with respect to* $\mathcal{F}_{\leq n-1}$.

(b) *For every* $\omega_{-\infty}^{n-1}\in\mathcal{A}_{-\infty}^{n-1}$,

$$\sum_{\omega_n\in\mathcal{A}} P_n\left(\omega_n \;\middle|\; \omega_{-\infty}^{n-1}\right) = 1. \tag{2.1.3}$$

Using these transition probabilities, we can construct the l**-move transition probabilities** $(l\geq 0)$

$$P_{[n,n+l]}\left(\omega_n^{n+l} \;\middle|\; \omega_{-\infty}^{n-1}\right) \triangleq \prod_{k=0}^{l} P_{n+k}\left(\omega_{n+k} \;\middle|\; \omega_{-\infty}^{n+k-1}\right). \tag{2.1.4}$$

A probability measure μ *in* $\mathcal{P}\left(\mathcal{A}^{\mathbb{Z}},\mathcal{F}\right)$ *is* **consistent** *with a system of transition probabilities* $\{P_n\}_{n\in\mathbb{Z}}$ *if*

$$\int h\left(\omega_{-\infty}^{n}\right)\mu(d\omega) = \int \sum_{\sigma_n\in\mathcal{A}} h\left(\omega_{-\infty}^{n-1}\,\sigma_n\right) P_n\left(\sigma_n \;\middle|\; \omega_{-\infty}^{n-1}\right)\mu(d\omega) \tag{2.1.5}$$

for all $n\in\mathbb{Z}$ *and all* $\mathcal{F}_{\leq n}$*-measurable functions* h. *Such a measure* μ *is called a* **chain with complete connections** *consistent with the system of transition probabilities* $\{P_n\}_{n\in\mathbb{Z}}$.

Definition 2.1.6

A system of transition probabilities is **weakly non-null** *on* $\mathcal{A}^{\mathbb{Z}}$ *if for all* $n\in\mathbb{Z}$, *there exists* $\omega_n\in\mathcal{A}$ *such that for all* $\omega_{-\infty}^{n-1}\in\mathcal{A}_{-\infty}^{n-1}$

$$P_n\left(\omega_n \;\middle|\; \omega_{-\infty}^{n-1}\right) > 0. \tag{2.1.7}$$

Definition 2.1.8

A system of transition probabilities is **non-null** *on* $\mathcal{A}^{\mathbb{Z}}$ *for all* $n\in\mathbb{Z}$ *and all* $\omega_{-\infty}^{n}\in\mathcal{A}_{-\infty}^{n}$,

$$P_n\left(\omega_n \;\middle|\; \omega_{-\infty}^{n-1}\right) > 0. \tag{2.1.9}$$

Definition 2.1.10

A system of transition probabilities is **continuous** *if the functions $P_n(\omega_n \mid \cdot)$ are continuous for each $n \in \mathbb{Z}$ and each $\omega_n \in \mathcal{A}$, that is*

$$\sup\left\{\left|P_n\left(\omega_n \mid \xi_{-\infty}^{n-1}\right) - P_n\left(\omega_n \mid \sigma_{-\infty}^{n-1}\right)\right| : \xi,\, \sigma \in \mathcal{A}_{-\infty}^{n-1},\ \xi \overset{k}{=} \sigma\right\} \xrightarrow{k\to+\infty} 0. \tag{2.1.11}$$

Remark 2.1.12

Different types of continuity can be considered. For example a system of transition probabilities is said to be **log-continuous** if for each $n \in \mathbb{Z}$ and each $\omega_n \in \mathcal{A}$,

$$\sup\left\{\left|\frac{P_n\left(\omega_n \mid \xi_{-\infty}^{n-1}\right)}{P_n\left(\omega_n \mid \sigma_{-\infty}^{n-1}\right)} - 1\right| : \xi,\, \sigma \in \mathcal{A}_{-\infty}^{n-1},\ \xi \overset{k}{=} \sigma\right\} \xrightarrow{k\to+\infty} 0. \tag{2.1.13}$$

This notion of continuity which has been introduced by Berbee (1987) coincides with (2.1.11) in the non-nullness case. Another type of continuity which has been introduced by Lalley (1986) is the **multiple-move log-continuous** and is defined by

$$\sup\left\{\left|\frac{P_{[n,n+l]}\left(\omega_n^{n+l} \mid \xi_{-\infty}^{n-1}\right)}{P_{[n,n+l]}\left(\omega_n^{n+l} \mid \sigma_{-\infty}^{n-1}\right)} - 1\right| : \xi,\, \sigma \in \mathcal{A}_{-\infty}^{n-1},\ \xi \overset{k}{=} \sigma\right\} \xrightarrow{k\to+\infty} 0, \tag{2.1.14}$$

for each $n \in \mathbb{Z},\ l \in \mathbb{N}$ and each $\omega_n^{n+l} \in \mathcal{A}_n^{n+l}$.

Proposition 2.1.15

A system of continuous transition probabilities has at least one probability measure consistent with it.

This result is a consequence of the fact that $\mathcal{A}^{\mathbb{Z}}$ is compact and $P_n(\omega_n \mid \cdot)$ are continuous.

2.2 Examples of chains with complete connections

This section is based on Fernández, Ferrari and Galves (2001).

Markov

A probability measure is a **Markov chain of order** n if its transition probabilities depend only on the n preceding states in the past history. More precisely, the transition probabilities satisfy

$$P\left(\omega_0 \mid \underline{\omega}\right) = P\left(\omega_0 \mid \omega_{-n}^{-1}\right) \tag{2.2.1}$$

for every $\omega_{-\infty}^{0} \in \mathcal{A}_{-\infty}^{0}$.

Finite mixtures of Markov chains

These processes were first introduce by Raftery (1985) (see also Raftery and Tavaré, 1994) under the name of **mixture transition distribution** (MTD) model. A probability measure is said to be a **mixture transition distribution** if its transition probabilities are finite combinations of transition probabilities of the form

$$P\left(\omega_0 \mid \underline{\omega}\right) = \sum_{k=1}^{n} \lambda_k P^{(k)}\left(\omega_0 \mid \omega_{-k}\right), \tag{2.2.2}$$

where $n \geq 0$ is fixed, $\{\lambda_k\}$ is a sequence of non-negative numbers such that $\sum_{k=1}^{n} \lambda_k = 1$ and each $P^{(k)}\left(\omega_0 \mid \omega_{-k}\right)$ is a k-order Markov transition probability depending only on the site k in the past history.

Bounded variable length Markov chains

These processes were introduced by Bühlman and Wyner (1999). A probability measure is a **variable-length Markov chain of order** n if its transition probabilities depend on a finite number of preceding states which are bounded functions of the past history. More precisely, there exists a lag function $l : \mathcal{A}_{-\infty}^{-1} \longrightarrow [0, n]$ such that

$$P\left(\omega_0 \mid \underline{\omega}\right) = P\left(\omega_0 \mid \omega_{-l(\underline{\omega})}^{-1}\right), \tag{2.2.3}$$

with the convention that if $\underline{\omega}$ is such that $l(\underline{\omega}) = 0$, then $P\left(\omega_0 \mid \underline{\omega}\right)$ is independent of $\underline{\omega}$.

Infinite mixtures of Markov chains

It is possible to extend the notion of MTD model with k ranging over the entire past and with transitions $P^{(k)}(\,\cdot \mid \cdot\,)$ which are k-order Markov transitions. Thus a probability measure is said to be a **countable mixture of Markov chains** if its transition probabilities are countable combinations of Markov transitions of increasing order. More precisely

$$P\left(\omega_0 \mid \underline{\omega}\right) = \lambda_0 P^{(0)}(\omega_0) + \sum_{k \geq 1} \lambda_k P^{(k)}\left(\omega_0 \mid \omega_{-k}^{-1}\right), \tag{2.2.4}$$

where $\{\lambda_k\}$ is a sequence of non-negative numbers such that $\sum_{k \geq 0} \lambda_k = 1$,
$P^{(0)}$ is a probability measure over $\mathcal{A}$ and each $P^{(k)}\left(\omega_0 \mid \omega_{-\infty}^{-1}\right)$ is a k-order Markov transition. The transition probabilities (2.2.4) can be interpreted as the result of two independent random steps. First, an integer $k \geq 0$ is chosen with probability λ_k, and second a letter of the alphabet is chosen with respect to the probability $P^{(k)}$. Thus each transition probability depends on a finite, but random, number of preceding states.

Variable length Markov chains

A **variable length Markov chain** is a variable length Markov chain of order n with

$n = +\infty$. Therefore a probability measure is a variable-length Markov chain if its transitions depend of a finite number of preceding states which is a function of the past history. More precisely, there exists a lag function $l : \mathcal{A}_{-\infty}^{-1} \longrightarrow \mathbb{N}$ such that

$$P\left(\omega_0 \,\middle|\, \underline{\omega}\right) = P\left(\omega_0 \,\middle|\, \omega_{-l(\underline{\omega})}^{-1}\right), \tag{2.2.5}$$

with the convention that if $\underline{\omega}$ is such that $l(\underline{\omega}) = 0$, then $P\left(\omega_0 \,\middle|\, \underline{\omega}\right)$ is independent of $\underline{\omega}$.

2.3 Related (in fact identical) notions

2.3.1 g-measures

The terms g-function and g-measure were introduce by Keane (1972), in the setting of covering transformations (see Definition 2.5.11 below). He studied the problem of existence and uniqueness of a g-measure with a given non-null and continuous g-function. Here we give definitions in our setting.

Definition 2.3.1
*A g-**function** is a $\mathcal{F}_{\leq 0}$-measurable function $g : \mathcal{A}_{-\infty}^{0} \longrightarrow [0,\ 1]$ which satisfies*

$$\sum_{a_0 \in \mathcal{A}} g\left(a_{-\infty}^{0}\right) = 1. \tag{2.3.2}$$

Remark 2.3.3
A g-function can be thought as a transition probability and can be written in the following form

$$g\left(a_{-\infty}^{0}\right) \triangleq g\left(a_0 \,\middle|\, a_{-\infty}^{-1}\right) \tag{2.3.4}$$

for all $a_{-\infty}^{0}$ in $\mathcal{A}_{-\infty}^{0}$.

Definition 2.3.5
*Let g be a g-function. A probability measure μ in $\mathcal{P}\left(\mathcal{A}_{-\infty}^{0}, \mathcal{F}_{\leq 0}\right)$ is a g-**measure** if μ is stationary and consistent with g, that is if*

$$\int h(\omega) g(\omega\, a) \mu(d\omega) = \int_{\{\omega_0 = a\}} h(\omega) \mu(d\omega) \tag{2.3.6}$$

for all a in $\mathcal{A}$ and all h in $\mathcal{F}_{\leq -1}$.

Definition 2.3.7
*A g-function is **continuous** if*

$$\operatorname{var}_k(g) \xrightarrow{k \to +\infty} 0 \tag{2.3.8}$$

*where the variation is defined by (1.2.4) and g is **non-null** if g is strictly positive.*

Remark 2.3.9
We can define a metric over $\mathcal{A}_{-\infty}^{0}$ by

$$d(\omega,\sigma) \triangleq \begin{cases} 2^{-n} & \text{if } \omega \text{ and } \sigma \text{ differ for the first time in the } n^{\text{th}} \text{ digit.} \\ 0 & \text{if } \omega = \sigma \end{cases} \tag{2.3.10}$$

The space $\left(\mathcal{A}_{-\infty}^{0}, d\right)$ is then a compact metric space and the continuity over this space coincide with (2.3.8).

Proposition 2.3.11 (Relation between Hölder and exponential variation)
Let g be a g-function. g is Hölderian if and only if g has exponential variations.

Proof Let g be a g-function defined over the space $\left(\mathcal{A}_{-\infty}^{0}, d\right)$, with d defined by (2.3.10). If g is Hölderian, then there exists and $\alpha > 0$ and $0 < \beta \leq 1$ such that

$$|g(\omega) - g(\sigma)| \leq \alpha \left(d(\omega,\sigma)\right)^{\beta}, \tag{2.3.12}$$

then for every $\omega, \sigma \in \mathcal{A}_{-\infty}^{0}$ such that $\omega \overset{k}{=} \sigma$, we have

$$|g(\omega) - g(\sigma)| \leq \alpha \left(2^{-k}\right)^{\beta}, \tag{2.3.13}$$

that is

$$\text{var}_k(g) \leq \alpha \left(2^{\beta}\right)^{-k}. \tag{2.3.14}$$

Thus if g is Hölderian, then $\text{var}_k(g)$ is exponential. Conversely, if $\text{var}_k(g)$ is exponential there exist $\alpha > 0$ and $a > 1$ such that for every $k \geq 0$

$$\text{var}_k(g) \leq \alpha\, a^{-k}. \tag{2.3.15}$$

Therefore, for every $\omega, \sigma \in \mathcal{A}_{-\infty}^{0}$ such that $\omega \overset{k}{=} \sigma$,

$$|g(\omega) - g(\sigma)| \leq \alpha \times \left(2^{-k}\right)^{\log_2 a}, \tag{2.3.16}$$

that is g is Hölderian. □

Remark 2.3.17
The transition probabilities of a g-measure are necessarily stationary so it suffices to define the transition probability (the g-function) in one site of $\mathbb{Z}$, for example in 0. Without the stationarity condition, we need to consider the entire family $\{g_i\}_{i\in\mathbb{Z}}$ of functions.

Definition 2.3.18
A sequence of functions $\{g_i\}_{i\in\mathbb{Z}}$, is a **family of g-functions** *if for each $i \in \mathbb{Z}$ the function $\tau_{-i}\, g_i$ is a g-function, with τ_i the shift by i acting on $\mathbb{Z}$.*

In fact "family of g-functions" is a synonymous with "system of transition probabilities".

Definition 2.3.19
Let $\{g_i\}_{i\in\mathbb{Z}}$ *a family of* g*-functions. A probability measure* μ *in* $\mathcal{P}\left(\mathcal{A}^{\mathbb{Z}},\mathcal{F}\right)$ *is* **consistent** *with* $\{g_i\}$ *if*

$$\int h(\omega)\, g_i(a\,\omega)\, \mu(d\omega) = \int_{\{\omega_i=a\}} h(\omega)\, \mu(d\omega) \tag{2.3.20}$$

for all $i \in \mathbb{Z}$*, all* a *in* $\mathcal{A}$ *and all* h *in* $\mathcal{F}_{\leq i-1}$*.*

2.3.2 Random Markov processes and uniform martingales

Random Markov processes and uniform martingales were introduce by Kalikow (1990) in order to exploit ergodic properties of g-measures.

Definition 2.3.21
A **complete random Markov process** *is a stationary process* $\{X_i, N_i\}_{i\in\mathbb{Z}}$ *where*

(a) *each* $X_i \in \mathcal{A}$ *and each* $N_i \in \mathbb{N}$*.*

(b) N_0 *is independent of* $\{X_i, N_i\}_{i<0}$*.*

(c) *For every* $j \in \mathbb{N}$*,*

$$P\left(X_0 = a_0 \mid X_{-\infty}^{-1}\,,\ N_0 = j\right) = P\left(X_0 = a_0 \mid X_{-j}^{-1}\,,\ N_0 = j\right). \tag{2.3.22}$$

Definition 2.3.23
A **random Markov process** *is the first coordinate of a complete random Markov process, that is, if* $\{X_i, N_i\}_{i\in\mathbb{Z}}$ *is a complete random Markov process then* $\{X_i\}_{i\in\mathbb{Z}}$ *is a random Markov process.*

If the N_i are bounded above by n, then the associated random Markov process is an k-order Markov chain. In this way, a random Markov process appears to be a generalization of a Markov process. The values $P\left(X_0 = \omega_0 \mid X_{-j}^{-1} = \omega_{-j}^{-1}\,,\ N_0 = j\right)$ for all $j \in \mathbb{N}$, $\omega_{-j}^{0} \in \mathcal{A}_{-j}^{0}$ and $P(N_0 = j)$ for all $j \in \mathbb{N}$ determine in a canonical way a random Markov process.

Definition 2.3.24
An **uniform martingale** *is a stationary process* $\{X_i\}_{i\in\mathbb{Z}}$ *such that for all* $\varepsilon > 0$*, there exists* $N \in \mathbb{N}$ *such that for all* $n > N$ *and all* $a_{-\infty}^{0} \in \mathcal{A}_{-\infty}^{0}$

$$\left|P\left(X_0 = a_0 \mid X_{-n}^{-1} = a_{-n}^{-1}\right) - P\left(X_0 = a_0 \mid X_{-\infty}^{-1} = a_{-\infty}^{-1}\right)\right| < \varepsilon. \tag{2.3.25}$$

Remark 2.3.26
A process is an uniform martingale if and only if the martingale convergence theorem holds uniformly for the martingale

$$\left\{P\left(X_0 = a_0 \mid X_{-n}^{-1} = a_{-n}^{-1}\right)\right\}_{n>1}.$$

This explains the name.

In the shift-invariant case, we can compare the notions of continuous g-measures, random Markov processes and uniform martingales. In fact a compactness argument shows that a process is a uniform martingale if and only if it is a g-measure consistent with a continuous g-function. Kalikow (1990) has proven the equivalence between the notions of uniform martingale and random Markov process.

2.4 Transfer operator

Let $t : \mathcal{A}_{-\infty}^{0} \to \mathcal{A}_{-\infty}^{0}$, $x \mapsto \tau\, x$ a (right) shift over $\mathcal{A}_{-\infty}^{0}$ defined by

$$\left(t\, x\right)_n = x_{n-1} \tag{2.4.1}$$

for each $n \leq 0$ and X be a subset of $\mathcal{A}_{-\infty}^{0}$. If X is t-invariant, we can then define the map T as the restriction of t to the subset X. This map T is called a **subshift**.

In this setting, we can define the notions of g-function and g-measure as follows: A **g-function over** (X, T) is a measurable function $g : X \longrightarrow [0, 1]$ which satisfies

$$\sum_{y \in T^{-1} x} g\left(y\right) = 1 \tag{2.4.2}$$

for all $x \in X$. In particular if $X = \mathcal{A}^{\mathbb{Z}}$ this definition coincides with Definition 2.3.1. For each g-function, there is an associated **transfer operator** defined by

$$\left(L_g\, h\right)(x) = \sum_{y \in T^{-1} x} g\left(y\right)\, h\left(y\right), \tag{2.4.3}$$

for all continuous functions h on X and all $x \in X$. The dual of L_g, denoted L_g^*, maps the set of probability measures on X on itself:

$$L_g^*\, \mu(h) = \mu\left(L_g\, h\right), \tag{2.4.4}$$

for all probability measures on X and all continuous functions h on X. The following result gives a characterization of the notion of g-measure via the transfer operator.

Theorem 2.4.5 (Ledrappier (1974))
*A probability measure μ on X is a g-***measure over (X,T)** *if one of the two following equivalent assertions holds*

(a) $L_g^* \mu = \mu$.

(b) *μ is T-invariant and*

$$\int h(x)\, \mu(dx) = \int \sum_{y \in T^{-1}T\,x} g(y)\, h(y)\, \mu(dx) \tag{2.4.6}$$

for all integrable functions h on X.

In particular if $X = \mathcal{A}^{\mathbb{Z}}$, this definition of g-measure coincides with Definition 2.3.5. To prove Theorem 2.4.5, it suffices to remark that

$$\mathrm{E}_\mu \left(h \;\middle|\; T^{-1}\,\mathcal{B} \right) = (L_g\, h) \circ T, \;\; \mu\text{-a.s.} \tag{2.4.7}$$

where $\mathcal{B}$ is the Borel σ-algebra associated to X.

We will complete the links with dynamical systems in sections 2.6 and 4.2.10.

2.5 Existence and uniqueness

For a finite alphabet and continuous transitions, there is no existence problem. The compactness of the configurations space implies that there always exists at less one chain of infinite order consistent with a given family of transition probabilities. We will now focus on the uniqueness problem. All results have been established in the setting of g-functions.

2.5.1 Uniqueness for a finite alphabet

Theorem 2.5.1 (Harris (1955))
Let g be a continuous, weakly non-null g-function on $\mathcal{A}_{-\infty}^0$ satisfying

$$\sum_{n \geq 1} \prod_{k=1}^{n} \left(1 - \frac{|E|}{2} \mathrm{var}_k(g) \right) = +\infty. \tag{2.5.2}$$

Then there exists an unique g-measure.

This old result gives one of the sharpest bound of uniqueness but curiously, it remained forgotten (except by the Romanian school, see Iosifescu and Theodorescu (1990) and all references therein) during a long time.

Berbee (1987) and Lalley (1986) independently addressed the issue. Their results are almost identical, through slightly different formalizations. Lalley relied on the notion of multiple-move log-continuity (Remark 2.1.12) and, for accidental reasons, published his full result only in 2000. We state Berbee's version because it is closer to our approach.

Theorem 2.5.3 (Berbee (1987), Lalley (1986,2000))
Let g be a non-null g-function such that

$$\sum_{n\geq 1} \exp\left(-\sum_{k=1}^{n} \mathrm{var}_k(\log g)\right) = +\infty, \tag{2.5.4}$$

then there exists an unique g-measure.

In the case $|E| = 2$, Harris' condition is weaker than that of Berbee and Lalley. For the case $|E| > 2$, these results are complementary. The proof given by Harris only works for $|E| = 2$. A complete proof has been given by Coelho and Quas (1998) and more recently by Stenflo (2002). Moreover Berbee gave a regeneration scheme (via a Markov process) under the stronger condition $\sum_{k\geq 1} \mathrm{var}_k(\log g) < \infty$.

Theorem 2.5.5 (Stenflo (2002))
Let g be a non-null g-function such that

$$\sum_{n\geq 1} \prod_{k=1}^{n} \Delta_k(g) = +\infty \tag{2.5.6}$$

where $\Delta_k(g) \triangleq \inf\{\sum_{\omega_0\in\mathcal{A}} \min\left(g\left(\underline{\omega}\,\omega_0\right), g\left(\underline{\sigma}\,\omega_0\right)\right) : \underline{\omega} \overset{k}{=} \underline{\sigma}\}$, then there exists an unique g-measure.

Condition (2.5.6) is weaker than conditions (2.5.2) and (2.5.4) in the sense that $1 - \frac{|E|}{2}\,\mathrm{var}_k(g) \leq \Delta_k(g)$ and $\exp\left(-\mathrm{var}_k(\log g)\right) \leq \Delta_k(g)$ for all $k \geq 1$. For a proof of this result see Stenflo (2002). On the other hand, Harris' result requires only weak non-nullness rather than (strong) non-nullness.

Theorem 2.5.7 (Johansson and Öberg (2002))
Let g be a non-null g-function such that

$$\sum_{k\geq 0} \mathrm{var}_k^2(\log g) < +\infty, \tag{2.5.8}$$

then there exists an unique g-measure.

In fact Johanson and Öberg's result is a little more general: They only require that the variation $\mathrm{var}_k\, g(\omega) \triangleq \sup_{\xi,\,\eta \in \mathcal{A}^0_{-k+1}} \left| g\left(\omega^{-k}_{-\infty}\, \xi^0_{-k+1}\right) - g\left(\omega^{-k}_{-\infty}\, \eta^0_{-k+1}\right)\right|$ (which depends of what happened before time $-k$) is square summable μ-always surely, where μ is a probability measure consistent with g. Moreover Johanson and Öberg have construct some examples which illustrate the fact that condition (2.5.8) is complementary to (2.5.4). A different approach due to Hulse is based on single-site oscillations. It yields yet another class of g-functions for which an uniqueness criterion holds.

Theorem 2.5.9 (Hulse (1997))
Let g be a non-null g-function such that

$$\sum_{k \le 0} \delta_k(g) < +\infty \tag{2.5.10}$$

where the oscillation δ_k is defined by (1.2.5), then there exists an unique g-measure.

In chapter 5, we will generalize this result for countable alphabets and non shift-invariant transition probabilities satisfying only weak non-nullness assumptions.

2.5.2 Uniqueness for countable or compact alphabet

Other results have been given in a more general setting. A few years after Harris, Keane readdressed the problem of uniqueness.

Definition 2.5.11
Let (X, d) be a compact metric space and T a homomorphism of the space X. T is said to be a **covering transformation** *of X if there exists an integer $n \ge 2$ and a real $\rho > 1$ such that*

(a) *For every $\omega \in X$, $|T^{-1}(\omega)| = n$.*

(b) *T is a local homeomorphism.*

(c) *For sufficiently small $\delta > 0$, $d(\omega, \sigma) = \delta$ implies $d(T\omega, T\sigma) \ge \rho\, \delta$.*

(d) *For each $\varepsilon > 0$ there exists N such that if $n \ge N$ and $\omega \in X$, then $T^{-n}(\omega)$ is ε-dense in X.*

Theorem 2.5.12 (Keane (1972))
Let (X, d) be a compact metric space, T a covering transformation of X and g a g-function over (X, T). If g is non-null and Lipschitz, then there exists exactly one g-measure over (X, T).

Petit (1975) has extended this result to Hölderian functions.

Iosifescu and Spataru (1973) have extended Harris' result in the setting of a countable alphabet and a non shift-invariant system of transition probabilities. Here, we state the result in the stationary case.

Theorem 2.5.13 (Iosifescu and Spataru (1973))
Assume that E is a countable alphabet (for example $\mathbb{N}^*$) and let g be a g-function over $\mathcal{A}_{-\infty}^0$. Assume that g is weakly non-null, continuous and satisfies

$$\sum_{n\geq 0}\prod_{k=0}^{n}(1-\alpha_k)=+\infty, \tag{2.5.14}$$

where $\alpha_k \triangleq \sup\left\{\sum_{j\geq 1}\left(s_j(\omega)-s_{j-1}(\omega)\vee s_j(\sigma)\right)^+ + \left(s_j(\sigma)-s_{j-1}(\sigma)\vee s_j(\omega)\right)^+ : \omega \overset{\neq k}{=} \sigma\right\}$,
with $s_j(\omega)\triangleq\sum_{i=1}^{j} g(\omega\, i)$. *Then there exists an unique g-measure.*

Walters (1975) gave a uniqueness condition in a more general setting than Harris, but with a less sharp condition over the variation (See also Ledrappier (1974) for another proof in the setting of finite alphabet).

Definition 2.5.15
The transformation $T : X \to X$ is said to be **topologically mixing** *if for all non-empty sets U, $V \subset X$, there exists $N > 0$ such that $T^{-n}U \cap V \neq \emptyset$ for all $n \geq N$.*

Theorem 2.5.16 (Walters (1975))
Let X be a compact metrizable space and $T : X \to X$ be a topologically mixing subshift. If g is a non-null g-function over (X, T) such that

$$\sum_{k\geq 0}\operatorname{var}_k(\log g)<+\infty, \tag{2.5.17}$$

then there exists an unique g-measure over (X, T).

More recently Comets, Fernández and Ferrari obtained the following uniqueness result.

Theorem 2.5.18 (Comets et al. (2002))
Let E be a possibly countable alphabet and g be a continuous, weakly non-null g-function such that

$$\sum_{n\geq 1}\prod_{k=1}^{n}\beta_k=+\infty, \tag{2.5.19}$$

where

$$\beta_k \triangleq \min_{b_{-k}^{-1}}\sum_{a_0\in\mathcal{A}}\inf_{c_{-\infty}^{-k-1}} g\left(c_{-\infty}^{-k-1}\, b_{-k}^{-1}\, a_0\right). \tag{2.5.20}$$

Then there exists an unique probability measure consistent with g.

This result is not new. For example it coincides with Harris' result for a two-symbol alphabet and is weaker for larger alphabets. The novelty in this work is a regenerating scheme which generalizes Berbee's work and yields a perfect simulation algorithm.

2.6 Decay of correlations

Fix $n \in \mathbb{N}$. We will use the following notation for correlations

$$\mathrm{Cor}_\mu(h_1, h_2) \triangleq |\mu(h_1 h_2) - \mu(h_1)\mu(h_2)|, \tag{2.6.1}$$

where $h_1 \in \mathcal{F}_{\leq 0}$ and $h_2 \in \mathcal{F}_{\leq -n}$.

Remark 2.6.2
Let t be the shift over $\mathcal{A}_{-\infty}^0$ defined by (2.4.1), μ a g-measure and $h_3,\ h_4$ two $\mathcal{F}_{\leq 0}$-measurable functions . Then, by (2.4.7), the correlations can be expressed by

$$\begin{aligned}
\mathrm{Cor}_\mu\left(h_3, h_4 \circ t^n\right) &= \int h_3(\omega)\, h_4 \circ t^n(\omega)\, \mu(d\omega) - \int h_3(\omega)\, \mu(d\omega) \int h_4(\omega)\, \mu(d\omega) \\
&= \int h_3(\omega)\, L_g^n h_4(\omega)\, \mu(d\omega) - \int h_3(\omega) \left(\int L_g^n\, h_4(\sigma)\, \mu(d\sigma) \right) \mu(d\omega) \\
&= \int h_3(\omega) \int \left(L_g^n h_4(\omega) - L_g^n\, h_4(\sigma) \right) \mu(d\sigma)\, \mu(d\omega).
\end{aligned} \tag{2.6.3}$$

Therefore the mixing rate can be bounded by the speed of the loss of memory of the chain.

For Lipschitzian g-functions the decay of correlations is exponential with respect to n (see for example Bowen (1975), Schmitt (2001)). In the non-Hölderian case, let $\beta = (\beta_n)_{n\in\mathbb{N}}$ be a real-valued sequence. We define a Markov chain $\left(S_n^\beta\right)_{n\in\mathbb{N}}$ starting from the origin

$$P\left(S_0^\beta = 0\right) = 1, \tag{2.6.4}$$

and having the following transition probabilities

$$\begin{aligned}
p_{i,\,i+1} &= 1 - \beta_i \\
p_{i,\,0} &= \beta_i,
\end{aligned} \tag{2.6.5}$$

for every i in $\mathbb{N}$. For every $n \geq 1$ we define

$$\beta_n^* = P\left(S_n^\beta = 0\right). \tag{2.6.6}$$

Theorem 2.6.7 (Bressaud, Fernández and Galves (1999))
Let g be a non-null g-function satisfying $\sum_{k\geq 1} \mathrm{var}_k(\log g) < +\infty$ and μ the unique probability measure consistent with g then for each integrable function $h_1 \in \mathcal{F}_{\leq 0}$ and each continuous fonction $h_2 \in \mathcal{F}_{\leq n}$, $n \geq 0$, such that $\sup_{k\geq 0} \frac{\mathrm{var}_k(h_2)}{\mathrm{var}_k(\log h_2)} < +\infty$

$$Cor_\mu(h_1, h_2) \leq C\,(h_1, h_2) \times \beta_n^*, \tag{2.6.8}$$

where $C\,(h_1, h_2)$ is a constant which depends on the choice of h_1 and h_2.

The mixing rate is driven by β_n^* which is determined by the sequence $\left\{\beta_k \triangleq 1 - e^{\mathrm{var}_k(\log g)}\right\}$. In particular, if $\mathrm{var}_k(\log g)$ is exponential in the sense $\mathrm{var}_k(\log g) \leq a^{-k}$ for each $k \geq 0$, with $a > 1$, then the decay of correlations is exponential with a rate b strictly greater than a. In chapter 5, we will give a sharper upper bound on the decay of correlations of a g-measure when the sum of the oscillations of the g-function is less than 1. This permits to obtain the same exponential rate for the decays and to extend the class of g-measures for which it is possible to estimate the mixing rate.

2.7 Non-uniqueness

The following result illustrates the complementary issue, namely the existence of g-functions with several consistent measures.

Theorem 2.7.1 (Bramson and Kalikow (1993))
There exist non-null continuous g-functions that do not satisfy uniqueness.

A simple proof of this result has been given by Lacroix (2000). Let us recall the original construction of the g-function introduced by Bramson and Kalikow but defined here over the alphabet $E = \{-1, 1\}$: let $\{m_j\}_{j\in\mathbb{N}^*}$ be an increasing sequence of odd strictly positive integers and $\{p_j\}_{j\in\mathbb{N}^*}$ be a decreasing sequence of strictly positive real numbers satisfying

$$\sum_{j=1}^{+\infty} p_j = 1 \tag{2.7.2}$$

and

$$p_k \leq \frac{1}{2} \sum_{j>k} p_j \tag{2.7.3}$$

for all $k \geq 1$. Then we define N a random variable supported by $\{m_j\}$ and admitting $\{p_j\}$ as distribution, that is

$$P(N = m_j) = p_j \tag{2.7.4}$$

and

$$W(\alpha, m_j) \triangleq \begin{cases} 1-\varepsilon & \text{if the majority of } \alpha_{-1}, ..., \alpha_{-m_j} \text{ are 1's} \\ \epsilon & \text{otherwise} \end{cases} \tag{2.7.5}$$

for all $\alpha \in A^{-\mathbb{N}^*}$, $j \in \mathbb{N}^*$ and where $\varepsilon \in \left]0, \frac{1}{2}\right[$ is fixed. The values $W(\alpha, m_j)$ and $P(N = m_j)$ define a g-function in the following way

$$g(\alpha) \triangleq \sum_{j=1}^{+\infty} W(\alpha, m_j) \times p_j, \tag{2.7.6}$$

for all $\alpha \in E^{-\mathbb{N}^*}$.

Main ideas of the proof:

In order to have a phase transition, our aim is to make $\text{var}_k(g)$ decrease slowly to zero. The dependence of the g-function with respect the past is driven by the sequences $\{m_j\}$ and $\{p_j\}$. For a fixed j, one chooses a dependence of size m_j with probability p_j. We will fix $\{p_j\}$ and $\{m_j\}$ so that the remote past does influence the present: $\{p_j\}$ will decrease slowly and $\{m_j\}$ will increase fast. First, two stationary processes (X^+, μ^+) and (X^-, μ^-) are constructed as the (shift-invariant) limits of the processes obtained by fixing respectively a past identically equal to $+1$ and a past identically equal to -1. The monotonicity property of g guarantees the existence of these limits and the consistency follows from the continuity of g. In the second part, the two processes X^+, X^- are showed to be distinct. For that, we remark that the g-function satisfies the symmetry property

$$P(-1 \mid -\alpha) \triangleq 1 - f(-\alpha) = f(\alpha) \triangleq P(1 \mid \alpha). \tag{2.7.7}$$

Let's consider the flipping transformation T

$$(T(\alpha_i))_{i<0} = (-\alpha_i)_{i<0}\,. \tag{2.7.8}$$

We have $\mu^- = T\mu^+$ and $\mu^+ = 1 - \mu^-$. Hence to show that $\mu^- \neq \mu^+$, it suffices to prove that

$$P(X_0^+ = 1) > \frac{1}{2}, \tag{2.7.9}$$

that is

$$E[X_0^+] > 0. \tag{2.7.10}$$

Bramson and Kalikow prove this by showing, through coupling techniques, that X^+ stochastically dominates an asymmetric Markovian process. This asymmetry implies (2.7.9).

Later Quas (1996) adapted the work of Bramson and Kalikow to obtain a non-uniqueness result for C^1 expanding maps.

Chapter 3

Gibbs measures

In this chapter, let us consider a configuration space of the form $E^{\mathbb{L}}$, where $(E, \mathcal{E})$ is a measurable space and $\mathbb{L}$ is a countable set. Let Ω be a subset of $E^{\mathbb{L}}$. For kernels associated to a specification (defined below), $\lim_{\Lambda \uparrow V} \gamma_\Lambda$ is the limit of the net $\{\gamma_\Lambda, \{\Lambda\}_{\Lambda \in \mathcal{S},\, \Lambda \subset \mathbb{L}}, \subset\}$ directed by inclusion.

The proper formalization of the notion of Gibbs measure is due to Dobrushin (1968) and Lanford and Ruelle (1969). In words, a measure is a Gibbs measure if for each finite Λ, its conditional probability, conditioned on the configuration outside Λ, is given by the Boltzmann-Gibbs formula based on a Hamiltonian H_Λ^Φ. We start by formalizing the notion of "conditioning on the exterior of a volume Λ". This corresponds to the concept of specification.

3.1 Specifications, consistency and extremality

3.1.1 Specifications

Definition 3.1.1

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

(a) *For each* $A \in \mathcal{F}$, $\gamma_\Lambda(A \mid \cdot) \in \mathcal{F}_{\Lambda^c}$.

(b) *For each* $B \in \mathcal{F}_{\Lambda^c}$ *and* $\omega \in \Omega$, $\gamma_\Lambda(B \mid \omega) = \mathbb{1}_B(\omega)$.

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

$$\gamma_\Delta \gamma_\Lambda = \gamma_\Delta \tag{3.1.2}$$

i.e.

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

for all measurable functions h and configurations $\omega \in \Omega$.

Definition 3.1.3
A probability measure μ is said to be **consistent** *with a specification $\{\gamma_\Lambda\}_{\Lambda\in\mathcal{S}}$ if*

$$\mu\gamma_\Lambda = \mu, \quad \forall\, \Lambda \in \mathcal{S}, \tag{3.1.4}$$

i.e.

$$\iint h(\sigma)\gamma_\Lambda(d\sigma \mid \omega)\mu(d\omega) = \int h(\omega)\mu(d\omega)$$

for all $\Lambda \in \mathcal{S}$, and all measurable functions h. The family of these measures will be denoted by $\mathcal{G}(\gamma)$.

3.1.2 Singleton consistency

We prove a (re)construction theorem for specifications starting from single-site conditioning, which applies in a general setting (general spin space, specifications not necessarily Gibbsian). We consider a family of **a priori** measures $\lambda = (\lambda^i)_{i\in\mathbb{L}}$ in $\mathcal{M}(E,\mathcal{E})$ and their products $\lambda^\Lambda \triangleq \bigotimes_{i\in\Lambda}\lambda^i$ for $\Lambda \subset \mathbb{L}$. We denote by $(\lambda_\Lambda)_{\Lambda\in\mathcal{S}}$ the family of measure kernels defined over $(\Omega, \mathcal{F})$ by

$$\lambda_\Lambda(h \mid \omega) = \left(\lambda^\Lambda \otimes \delta_{\omega_{\Lambda^c}}\right)(h) \tag{3.1.5}$$

for every measurable function h and configuration ω. These kernels satisfy the following identities for every $\Lambda \in \mathcal{S}$:

$$\lambda_\Lambda(B \mid \cdot\,) \;=\; \mathbb{1}_B(\,\cdot\,),\; \forall\, B \in \mathcal{F}_{\Lambda^c} \tag{3.1.6}$$

and

$$\lambda_{\Lambda\cup\Delta} \;=\; \lambda_\Lambda\lambda_\Delta\,, \quad \forall\, \Delta \in \mathcal{S} : \Lambda \cap \Delta = \emptyset. \tag{3.1.7}$$

Theorem 3.1.8
Let λ be as above and $(\gamma_i)_{i\in\mathbb{L}}$ be a family of probability kernels on $\mathcal{F}_i \times \Omega$ such that

1) *For each $i \in \mathbb{L}$ and for some measurable function ρ_i,*

$$\gamma_i \;=\; \rho_i\lambda_i\,. \tag{3.1.9}$$

2) *The following properties hold:*

(a) *Normalization on Ω: for every i in $\mathbb{L}$,*

$$\left(\lambda_i\left(\rho_i\right)\right)(\omega) \;=\; 1\,, \quad \forall\, \omega \in \Omega\,. \tag{3.1.10}$$

(b) *Bounded-positivity on* Ω*: for every* $i, j \in \mathbb{L}$,

$$\inf_{\omega \in \Omega} \lambda_j \left(\rho_j \, \rho_i^{-1}\right)(\omega) \;>\; 0 \tag{3.1.11}$$

and

$$\sup_{\omega \in \Omega} \lambda_j \left(\rho_j \, \rho_i^{-1}\right)(\omega) \;<\; +\infty \;. \tag{3.1.12}$$

(c) *Order-consistency on* Ω*: for every* i, j *in* $\mathbb{L}$ *and every* $\omega \in \Omega$,

$$\rho_{ij}(\omega) \;=\; \frac{\rho_i}{\lambda_i \left(\rho_i \, \rho_j^{-1}\right)}(\omega) \;=\; \frac{\rho_j}{\lambda_j \left(\rho_j \, \rho_i^{-1}\right)}(\omega) \;. \tag{3.1.13}$$

Then there exists a unique family $\rho = \{\rho_\Lambda\}_{\Lambda \in \mathcal{S}}$ *of positive measurable functions on* $(\Omega, \mathcal{F})$ *such that*

(i) $\gamma \triangleq \{\rho_\Lambda \lambda_\Lambda\}_{\Lambda \in \mathcal{S}}$ *is a specification on* $(\Omega, \mathcal{F})$ *with* $\gamma_{\{i\}} = \gamma_i$ *for each* $i \in \mathbb{Z}^d$.

(ii) $\rho_{\Lambda \cup \Gamma} = \dfrac{\rho_\Lambda}{\lambda_\Lambda \left(\rho_\Lambda \, \rho_\Gamma^{-1}\right)}$, *for all* $\Lambda, \Gamma \in \mathcal{S}$ *such that* $\Gamma \subset \Lambda^c$.

(iii) $\mathcal{G}(\gamma) = \{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \mu\gamma_i = \mu$ *for all* $i \in \mathbb{L}\}$.

(iv) *For each* $\Lambda \in \mathcal{S}$ *there exist constants* $C_\Lambda, D_\Lambda > 0$ *such that* $C_\Lambda \, \rho_k(\omega) \le \rho_\Lambda(\omega) \le D_\Lambda \, \rho_k(\omega)$ *for all* $k \in \Lambda$ *and all* $\omega \in \Omega$.

Remarks

3.1.14 This theorem is a strengthening of the reconstruction result given by Theorem 1.33 in Georgii (1988). In the latter, the order-consistency condition (3.1.13) is replaced by the requirement that the singletons come from a pre-existing specification (which the prescription reconstructs).

3.1.15 Identity (ii) can be used, in fact, to inductively define the family ρ by adding one site at a time. In fact, this is what is done in the proof below. The inequalities (iv) relate the non-nullness properties of ρ to those of the original family $\{\rho_i\}_{i \in \mathbb{Z}^d}$.

3.1.16 In the case E denumerable, λ_i=counting measure, the order-consistency requirement (3.1.13) is automatically verified if the singletons are defined through a measure μ on $\mathcal{F}$ of the form

$$\rho_i(\omega) \;=\; \lim_{n \to \infty} \frac{\mu(\omega_{V_n})}{\mu(\omega_{V_n \setminus \{i\}})}$$

for an exhausting sequence of volumes $\{V_n\}$. Indeed, a simple computation shows that the last two terms in (3.1.13) coincide with

$$\lim_{n \to \infty} \frac{\mu(\omega_{V_n})}{\mu(\omega_{V_n \setminus \{i,j\}})} \;.$$

Proof In the following all functions are defined on Ω or on a projection of Ω over a subset of $\mathbb{L}$.

Initially we define ρ by choosing a total order for $\mathbb{L}$ and prescribing, inductively, that for each $\Lambda \in \mathcal{S}$ with $|\Lambda| \geq 2$ and each $\omega \in \Omega$

$$\rho_\Lambda(\omega) \;=\; \frac{\rho_k}{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)}(\omega)\;, \tag{3.1.17}$$

where $k = \max \Lambda$ and $\Lambda_k^* = \Lambda \setminus \{k\}$. For each $\Lambda, \Gamma \in \mathcal{S}$ such that $\Gamma \subset \Lambda^c$, we will prove, by induction over $|\Lambda \cup \Gamma|$, that the functions so defined satisfy the following properties:

(I1) $\inf_{\omega\in\Omega} \lambda_\Lambda\left(\rho_\Lambda\,\rho_\Gamma^{-1}\right)(\omega) > 0$ and $\sup_{\omega\in\Omega} \lambda_\Lambda\left(\rho_\Lambda\,\rho_\Gamma^{-1}\right)(\omega) < +\infty$.

(I2) $\rho_{\Lambda\cup\Gamma} = \dfrac{\rho_\Lambda}{\lambda_\Lambda\left(\rho_\Lambda\,\rho_\Gamma^{-1}\right)}$.

(I3) $\lambda_\Lambda\left(\rho_\Lambda\right) = 1$.

(I4) If μ is a probability measure on $(\Omega, \mathcal{F})$ such that $\mu\left(\rho_i\,\lambda_i\right) = \mu$, $\forall\, i \in \Lambda$, then $\mu\left(\rho_\Lambda\,\lambda_\Lambda\right) = \mu$.

(I5) $\left(\rho_{\Lambda\cup\Gamma}\,\lambda_{\Lambda\cup\Gamma}\right)\left(\rho_i\,\lambda_i\right) = \rho_{\Lambda\cup\Gamma}\,\lambda_{\Lambda\cup\Gamma}$, $\forall\, i \in \Lambda\cup\Gamma$.

Let us first comment why these properties imply the theorem. It is clear that properties (I3)–(I5), together with the deterministic character of (λ_Λ) on $\mathcal{F}_{\Lambda^c}$ [property (3.1.6)], imply that $\left(\rho_\Lambda\lambda_\Lambda\right)_{\Lambda\in\mathcal{S}}$ verifies assertions (i)–(iv). Furthermore, if $\widetilde{\gamma}$ is a specification such that $\widetilde{\gamma}_{\{i\}} = \gamma_i$ for all $i \in \mathbb{L}$ then, by consistency, $\widetilde{\gamma}_\Lambda(\,\cdot \mid \omega)\,\gamma_i = \widetilde{\gamma}_\Lambda(\,\cdot \mid \omega)$ for every $\Lambda \in \mathcal{S}$, $i \in \Lambda$ and $\omega \in \Omega$. Therefore property (I5) implies that $\widetilde{\gamma}_\Lambda(\,\cdot \mid \omega)\;(\rho_\Lambda\lambda_\Lambda) = \widetilde{\gamma}_\Lambda(\,\cdot \mid \omega)$, that is $\rho_\Lambda\lambda_\Lambda(\,\cdot \mid \omega) = \widetilde{\gamma}_\Lambda(\,\cdot \mid \omega)$. So the construction is unique.

Initial inductive step The first non-trivial case is when $|\Lambda \cup \Gamma| = 2$. This implies that $|\Lambda| = |\Gamma| = 1$ and hence (I1)–(I3) coincide with hypotheses (3.1.10)–(3.1.13) while (I4) is trivially true. To prove (I5), assume that $\Lambda = \{i\}$ and $\Gamma = \{j\}$. By (3.1.7) and (3.1.13), we have

$$\left(\rho_{ij}\,\lambda_{ij}\right)\left(\left(\rho_i\,\lambda_i\right)(h)\right) \;=\; \lambda_j\left[\left(\frac{\rho_i\,\lambda_i}{\lambda_i\left(\rho_i\,\rho_j^{-1}\right)}\right)\left(\left(\rho_i\,\lambda_i\right)(h)\right)\right]\;.$$

As the factor $\left(\rho_i\,\lambda_i\right)(h)/\lambda_i\left(\rho_i\,\rho_j^{-1}\right)$ is independent of the configuration at $\{i\}$, the remaining integration with respect to the measure $\rho_i\lambda_i$ disappears due to the normalization condition (3.1.10). We obtain

$$\left(\rho_{ij}\,\lambda_{ij}\right)\left(\left(\rho_i\,\lambda_i\right)(h)\right) \;=\; \lambda_j\left[\left(\rho_i\,\lambda_i\right)\left(\frac{h}{\lambda_i\left(\rho_i\,\rho_j^{-1}\right)}\right)\right] \;=\; \left(\rho_{ij}\,\lambda_{ij}\right)(h)$$

Inductive step We suppose the assertions true for $|\Lambda \cup \Gamma| = n, (n \geq 2)$, and consider Λ, Γ such that $\Gamma \subset \Lambda^c$ and $|\Lambda \cup \Gamma| = n + 1$.

(I1) Assume first that $|\Gamma| = 1$ and let $k = \max \Lambda$. Combining the definition (3.1.17) and the property (3.1.7), we obtain

$$\lambda_\Lambda \left(\rho_\Lambda \, \rho_\Gamma^{-1}\right) \;=\; \lambda_{\Lambda_k^*} \left(\frac{\lambda_k \left(\rho_k \, \rho_\Gamma^{-1}\right)}{\lambda_k \left(\rho_k \, \rho_{\Lambda_k^*}^{-1}\right)} \right) . \tag{3.1.18}$$

If $|\Gamma| \geq 2$ we consider $l \triangleq \max \Gamma$ and apply the definition (3.1.17) to obtain

$$\lambda_\Lambda \left(\rho_\Lambda \, \rho_\Gamma^{-1}\right) \;=\; \lambda_\Lambda \left(\rho_\Lambda \, \rho_l^{-1} \, \lambda_l \left(\rho_l \, \rho_{\Gamma_l^*}^{-1}\right)\right) . \tag{3.1.19}$$

We can now apply the inductive hypothesis (I1) to the right-hand side of (3.1.18) and (3.1.19) to prove (I1) at the next inductive level.

(I2) The argument is symmetric in Λ and Γ, so we can assume without loss that $k = \max(\Lambda \cup \Gamma)$ belongs to Λ. If $|\Lambda| = 1$ (I2) is just the definition (3.1.17) applied to $\Lambda \cup \Gamma$. We assume, hence, that $|\Lambda| \geq 2$ and consider $j \in \Lambda$ such that $j \neq k$. By the inductive assumption (I2) we have

$$\rho_\Lambda \;=\; \frac{\rho_{\Lambda_j^*}}{\lambda_{\Lambda_j^*} \left(\rho_{\Lambda_j^*} \, \rho_j^{-1}\right)} \;=\; \frac{\rho_j}{\lambda_j \left(\rho_j \, \rho_{\Lambda_j^*}^{-1}\right)} . \tag{3.1.20}$$

We first combine the rightmost preceding expression with the factorization property (3.1.7) to write

$$\lambda_\Lambda \left(\rho_\Lambda \, \rho_\Gamma^{-1}\right) \;=\; \lambda_{\Lambda_j^*} \left(\frac{\lambda_j \left(\rho_j \, \rho_\Gamma^{-1}\right)}{\lambda_j \left(\rho_j \, \rho_{\Lambda_j^*}^{-1}\right)} \right) . \tag{3.1.21}$$

We now apply once more the inductive assumption (I2) in the form

$$\lambda_j \left(\rho_j \, \rho_\Gamma^{-1}\right) \;=\; \rho_{\Gamma \cup \{j\}}^{-1} \, \rho_j \tag{3.1.22}$$

in combination with the rightmost identity in (3.1.20), to obtain

$$\lambda_j \left(\rho_j \, \rho_{\Lambda_j^*}^{-1}\right) \;=\; \rho_{\Lambda_j^*}^{-1} \, \rho_j \, \lambda_{\Lambda_j^*} \left(\rho_{\Lambda_j^*} \, \rho_j^{-1}\right) . \tag{3.1.23}$$

From (3.1.21)–(3.1.23) we get

$$\lambda_\Lambda \left(\rho_\Lambda \, \rho_\Gamma^{-1}\right) \;=\; \frac{\lambda_{\Lambda_j^*} \left(\rho_{\Lambda_j^*} \, \rho_{\Gamma \cup \{j\}}^{-1}\right)}{\lambda_{\Lambda_j^*} \left(\rho_{\Lambda_j^*} \, \rho_j^{-1}\right)} .$$

We now use this relation together with the first identity in (3.1.20) to conclude that

$$\frac{\rho_\Lambda}{\lambda_\Lambda\left(\rho_\Lambda\,\rho_\Gamma^{-1}\right)} \;=\; \frac{\rho_{\Lambda_j^*}}{\lambda_{\Lambda_j^*}\left(\rho_{\Lambda_j^*}\,\rho_{\Gamma\cup\{j\}}^{-1}\right)}\;.$$

We iterate this formula $\left|\Lambda_j^*\right| - 1$ times and we arrive to

$$\frac{\rho_\Lambda}{\lambda_\Lambda\left(\rho_\Lambda\,\rho_\Gamma^{-1}\right)} \;=\; \frac{\rho_k}{\lambda_k\left(\rho_k\,\rho_{(\Lambda\cup\Gamma)_k^*}^{-1}\right)}$$

which is precisely $\rho_{\Lambda\cup\Gamma}$ according to our definition (3.1.17).

(I3) We assume that $|\Lambda| \geq 2$, otherwise (I3) is just the normalization hypothesis (3.1.10). Let $k = \max\Lambda$. Definition 3.1.17 and property 3.1.7 yield

$$\lambda_\Lambda\left(\rho_\Lambda\right) \;=\; \lambda_k\left(\frac{\lambda_{\Lambda_k^*}\left(\rho_{\Lambda_k^*}\right)}{\lambda_{\Lambda_k^*}\left(\rho_{\Lambda_k^*}\,\rho_k^{-1}\right)}\right) \;=\; \lambda_k\left(\frac{1}{\lambda_{\Lambda_k^*}\left(\rho_{\Lambda_k^*}\,\rho_k^{-1}\right)}\right)$$

where the last identity follows from the inductive hypothesis (I3). But, as in (3.1.22),

$$\lambda_{\Lambda_k^*}\left(\rho_{\Lambda_k^*}\,\rho_k^{-1}\right) \;=\; \lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)\,\rho_{\Lambda_k^*}\,\rho_k^{-1}\;,$$

therefore

$$\lambda_\Lambda\left(\rho_\Lambda\right) \;=\; \frac{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)}{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)} \;=\; 1\;.$$

(I4) To avoid a triviality we assume that $|\Lambda| \geq 2$. Let μ be a probability measure on $(\Omega, \mathcal{F})$ such that $\mu\left(\rho_i\,\lambda_i\right) = \mu$ for all $i \in \Lambda$. Consider $k = \max\Lambda$ and a measurable function h. By the factorization property (3.1.7) of λ_Λ and the definition (3.1.17) of ρ_Λ, we have

$$\mu\Big(\left(\rho_\Lambda\lambda_\Lambda\right)(h)\Big) \;=\; \mu\left[\lambda_{\Lambda_k^*}\left(\left(\frac{\rho_k}{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)}\,\lambda_k\right)(h)\right)\right]\;.$$

By the inductive hypothesis (I4) μ is consistent with $\rho_{\Lambda_k^*}\lambda_{\Lambda_k^*}$ and with $\rho_k\lambda_k$, thus

$$\mu\Big(\left(\rho_\Lambda\lambda_\Lambda\right)(h)\Big) \;=\; \mu\left[\left(\rho_k\,\lambda_k\right)\left(\rho_{\Lambda_k^*}^{-1}\left(\frac{\rho_k}{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)}\,\lambda_k\right)(h)\right)\right]\;.$$

But, in the right-hand side, the two innermost integrals with respect to λ_k commute with the external one, so we have

$$\begin{aligned}
\mu\Big(\left(\rho_\Lambda\lambda_\Lambda\right)(h)\Big) \;&=\; \mu\left[\left(\rho_k\,\lambda_k\right)\left(\frac{h}{\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)}\quad\lambda_k\left(\rho_k\,\rho_{\Lambda_k^*}^{-1}\right)\right)\right]\\
&=\; \mu\Big(\left(\rho_k\,\lambda_k\right)(h)\Big)\;,
\end{aligned}$$

which proves (I4).

(I5) Denote $\Delta = \Lambda \cup \Gamma$ and pick $i, j \in \Delta$, $i \neq j$ and a measurable function h. By (3.1.7) and (I2) we have

$$\Big(\rho_\Delta \, \lambda_\Delta\Big)\Big((\rho_i \, \lambda_i)\,(h)\Big) \;=\; \lambda_j \left[\frac{\Big(\rho_{\Delta_j^*} \, \lambda_{\Delta_j^*}\Big)\Big((\rho_i \, \lambda_i)\,(h)\Big)}{\lambda_{\Delta_j^*}\Big(\rho_{\Delta_j^*} \, \rho_j^{-1}\Big)} \right] .$$

Therefore, applying inductive assumption (I5) we obtain

$$\begin{aligned} \Big(\rho_\Delta \, \lambda_\Delta\Big)\Big((\rho_i \, \lambda_i)\,(h)\Big) \;&=\; \lambda_j \left[\frac{\Big(\rho_{\Delta_j^*} \, \lambda_{\Delta_j^*}\Big)\,(h)}{\lambda_{\Delta_j^*}\Big(\rho_{\Delta_j^*} \, \rho_j^{-1}\Big)} \right] \\ &=\; (\rho_\Delta \, \lambda_\Delta)\,(h) \; . \quad \square \end{aligned} \tag{3.1.25}$$

3.1.3 Triviality

Theorem 3.1.26
Let $\gamma = (\gamma_\Lambda)_{\Lambda \in \mathcal{S}}$ be a specification defined over $(\Omega, \mathcal{F})$ and

$$\mathcal{G}(\gamma) \triangleq \{\mu \in \mathcal{P}(\Omega, \mathcal{F}) : \mu\gamma_\Lambda = \mu, \quad \forall \, \Lambda \in \mathcal{S}\} ,$$

the convex set of all probability measure consistent with γ and $\mathcal{F}_\infty \triangleq \bigcap_{\Lambda \in \mathcal{S}} \mathcal{F}_{\Lambda^c}$, the **tail σ-field**. *Then*

(a) *μ is extreme in $\mathcal{G}(\gamma)$ if and only if μ is trivial over $\mathcal{F}_\infty$.*

(b) *Let $\mu \in \mathcal{G}(\gamma)$ and $\nu \in \mathcal{P}(\Omega, \mathcal{F})$ such that $\nu \ll \mu$,*
$\nu \in \mathcal{G}(\gamma)$ if and only if $\exists h \geq 0$, $\mathcal{F}_\infty$-measurable $: \nu = h\mu$.

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

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

3.1.4 Short range correlations

Proposition 3.1.27
For each probability measure over $(\Omega, \mathcal{F})$, the following statements are equivalent

(a) *μ is trivial over $\mathcal{F}_\infty$*

(b) *For all cylinder sets A in $\mathcal{F}$,*

$$\lim_{\Lambda\uparrow\mathbb{L}} \sup_{B\in\mathcal{F}_{\Lambda^c}} \Big|\mu(A\cap B)-\mu(A)\,\mu(B)\Big| = 0. \tag{3.1.28}$$

(c) *For all $A\in\mathcal{F}$,*

$$\lim_{\Lambda\uparrow\mathbb{L}} \sup_{B\in\mathcal{F}_{\Lambda^c}} \mid \mu(A\cap B)-\mu(A)\,\mu(B) \mid = 0. \tag{3.1.29}$$

3.1.5 Ergodic measures

Let $\mathbb{L}=\mathbb{Z}^d$ for some $d\geq 1$. For each $j\in\mathbb{Z}^d$, the transformation

$$\tau_j : \mathbb{Z}^d \to \mathbb{Z}^d,\ i\mapsto i+j \tag{3.1.30}$$

is called the **shift** by j. The shift induces actions on configurations, measurable sets, measurable functions and measures that we denote with the same symbol: for each $j\in\mathbb{Z}^d$; for $\omega\in\Omega$ $\tau_j(\omega)=(\omega_{i-j})_{i\in\mathbb{Z}}$, for $A\in\mathcal{F}$, $\tau_j A=\{\tau_j\omega : \omega\in A\}$; for h $\mathcal{F}$-measurable, $(\tau_j h)(\omega)=h(\tau_j^{-1}\omega)$, and for a measure μ on $(\Omega,\mathcal{F})$, $(\tau_j\mu)(h)=\mu(\tau_j^{-1}h)$. Objects invariant under the action of the shift by j for all $j\in\mathbb{Z}^d$ are called **shift-invariant**. A specification γ is **shift-invariant** if

$$\gamma_{\Lambda+j}\left(\tau_j A \mid \tau_j\omega\right) = \gamma_\Lambda\left(A \mid \omega\right)$$

for each $\Lambda\in\mathcal{S}$, $j\in\mathbb{Z}^d$, $\omega\in\Omega$ and with $\Lambda+j=\{i+j : i\in\Lambda\}$. We denote by $\mathcal{P}_{\rm inv}(\Omega,\mathcal{F})$ the set of shift-invariant probability measures on $(\Omega,\mathcal{F})$ and by $\mathcal{G}_{\rm inv}(\gamma)$ the set shift-invariant Gibbs measure consistent with the specification γ. A probability measure $\mu\in\mathcal{P}_{\rm inv}(\Omega,\mathcal{F})$ is **ergodic** if μ is trivial on $\mathcal{I}\triangleq\left\{A\in\mathcal{F} : \tau_j A=A,\ \forall\, j\in\mathbb{Z}^d\right\}$ the σ-algebra of all shift-invariant events.

Theorem 3.1.31 (Ergodic Gibbs measures)
Let γ be a shift-invariant specification.

(a) *A measure $\mu\in\mathcal{G}_{\rm inv}(\gamma)$ is extreme in $\mathcal{G}_{\rm inv}(\gamma)$ if and only if μ is ergodic.*

(b) *If $\mu\in\mathcal{G}_{\rm inv}(\gamma)$ and $\nu\in\mathcal{P}_{\rm inv}(\Omega,\mathcal{F})$ satisfy $\nu\ll\mu$ then $\nu\in\mathcal{G}_{\rm inv}(\gamma)$.*

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

3.2 Gibbs measures

Gibbs measure theory requires a topological structure for the configuration space $E^{\mathbb{L}}$. We assume that E has a topology and consider the associated Borel σ-algebra. We endow Ω with the restriction of the corresponding product structures.

3.2.1 Definitions

Definition 3.2.1

1) *For every* $\Lambda \in \mathcal{S}$, *let* $\mathcal{L}_\Lambda$ *be the* **space of all bounded** $\mathcal{F}_\Lambda$**-measurable functions**. *The* **space of all bounded local functions** *is then defined as* $\mathcal{L} \triangleq \bigcup_{\Lambda \in \mathcal{S}} \mathcal{L}_\Lambda$.

2) *A function is* **quasilocal** *if it is the uniformly convergent limit of some sequence of local functions or equivalently if*

$$\lim_{\Lambda \uparrow \mathbb{L}} \sup \left\{ \left| h(\omega) - h(\sigma) \right| : \omega,\, \sigma \in \Omega,\ \omega_\Lambda = \sigma_\Lambda \right\} = 0. \tag{3.2.2}$$

We denote by $\overline{\mathcal{L}}$ *the space of all bounded quasilocal functions.*

Definition 3.2.3

A specification γ *is said to be:*

1) **Quasilocal** *if, for each* $\Lambda \in \mathcal{S}$, $h \in \overline{\mathcal{L}}$ *implies* $\gamma_\Lambda\, h \in \overline{\mathcal{L}}$.

2) **Continuous** *if the functions* $\Omega \ni \omega \to \gamma_\Lambda \left(A \,\middle|\, \omega \right)$ *are continuous for all* $\Lambda \in \mathcal{S}$ *and all* $A \in \mathcal{F}$.

If E is finite then continuity coincides with quasilocality.

Definition 3.2.4

Let $\mu^0 = \prod_{x \in \mathbb{L}} \mu^0_x$ *be an a priori probability measure.*

1) *A specification* γ *is said to be* **non-null** *in* Ω *with respect to* μ^0 *if*

$$\gamma_\Lambda(A \mid \omega) > 0, \tag{3.2.5}$$

for all $\omega \in \Omega$, $\Lambda \in \mathcal{S}$ *and* $A \in \mathcal{F}_\Lambda$ *such that* $\mu^0(A) > 0$.

2) *A specification* γ *is said to be* **uniformly non-null** *in* Ω *with respect to* μ^0 *if, for each* $\Lambda \in \mathcal{S}$, *there exist constants* $0 < \alpha_\Lambda \le \beta_\Lambda < \infty$ *such that*

$$\alpha_\Lambda \mu^0(A) \le \gamma_\Lambda(A \mid \omega) \le \beta_\Lambda \mu^0(A) \tag{3.2.6}$$

for all $\omega \in \Omega$ *and* $A \in \mathcal{F}_\Lambda$.

Definition 3.2.7

A probability measure is said to be a **Gibbs measure** *on* Ω *is there exists a continuous and non-null specification* γ *on* Ω *such that* $\mu \in \mathcal{G}(\gamma)$.

3.2.2 Gibbsian specifications

The traditional way to define Gibbs measures is through Boltzmann-Gibbs weights. These are proportional to $e^{-\beta H_\Lambda}$, where β is the inverse temperature and H_Λ is the Hamiltonian for the region Λ. For completeness we discuss here the relation between our Definition 3.2.7 and the Hamiltonian representation.
Two important requirements must be considered at this point.

1) The Hamiltonians must be sums of terms depending on spins at finite sets of sites. Hamiltonian for larger regions are obtained simply by adding extra terms to Hamiltonian for smaller regions.

2) There must exist a well defined Hamiltonian describing the effect of freezing exterior spins in some configuration. This requires suitable summability conditions.

The first requirement implies that the basic objects in the construction of Boltzmann-Gibbs weights are not the Hamiltonians but the interactions.

Definition 3.2.8
An **interaction** *(or* **interaction potential** *or* **potential***) is a family* $\Phi = \{\Phi_A\}_{A\in\mathcal{S}}$*, where each* Φ_A *is a real valued function of the configurations such that for each* $A \in \mathcal{S}$*,* Φ_A *is* $\mathcal{F}_A$*-measurable.*

Definition 3.2.9
Let Φ *be an interaction. Then for all* $\Lambda \in \mathcal{S}$*, the* **Hamiltonian** H_Λ^Φ *for volume* Λ *is the function*

$$H_\Lambda^\Phi(\omega) \triangleq \sum_{\substack{A\in\mathcal{S}\\ A\cap\Lambda\neq\emptyset}} \Phi_A(\omega) \tag{3.2.10}$$

provided that this sum converges to a finite limit for all $\omega \in \Omega$*.*

Remark 3.2.11
For our purposes it is convenient to think of the configuration outside Λ as be fixed and the configuration inside Λ as variable. Therefore, for any $\sigma \in \Omega$, we define the Hamiltonian on Λ for a fixed external condition σ as

$$H_{\Lambda,\sigma}^\Phi(\omega) = H_\Lambda^\Phi\left(\omega_\Lambda\,\sigma_{\Lambda^c}\right). \tag{3.2.12}$$

The summability properties of the interaction (3.2.10) determine the scope of the corresponding statistical mechanical theory. In addition to the notion of convergence introduced in Definition 3.2.9 above, we wish to distinguish two stronger notions of summability.

Definition 3.2.13
We say that the interaction Φ *is*

1) **Uniformly convergent** *if, for every* $\Lambda \in \mathcal{S}$, $\sum_{\substack{A\in\mathcal{S}\\ A\cap\Lambda\neq\emptyset}} \Phi_A(\omega)$ *converges uniformly in* ω.

2) **Absolutely summable** *if, for every* $i \in \mathbb{L}$, $\sum_{\substack{A\in\mathcal{S}\\ A\ni i}} \|\Phi_A\|_\infty < \infty$.

Absolutely summability implies uniform convergence, which in turns implies the convergence of the interaction.

Definition 3.2.14
Let Φ *be a convergent interaction, so that we can define the Hamiltonians* H_Λ^Φ. *Let* $\mu^0 \triangleq \prod_{x\in\mathbb{L}} \mu_x^0$ *be a probability measure, called the a priori measure. We define the* **partition function** *by*

$$Z_\Lambda^\Phi\left(\omega_{\Lambda^c}\right) \triangleq \int \exp\left[-H_\Lambda^\Phi(\omega)\right] \prod_{x\in\mathbb{L}} \mu_x^0\left(\omega_x\right). \tag{3.2.15}$$

Moreover, if $Z_\Lambda^\Phi\left(\omega_{\Lambda^c}\right) < +\infty$ *for all* $\Lambda \in \mathcal{S}$ *and all* $\omega \in \Omega$, *we say that the interaction is* μ^0*-admissible.*

Definition 3.2.16
Let $\mu^0 = \prod_{x\in\mathbb{L}} \mu_x^0$ *be a probability measure, and let* Φ *be a convergent,* μ^0*-admissible interaction. Then the probability measure* $\gamma_\Lambda^\Phi(\omega \mid \cdot)$ *on* $\mathcal{F}$ *defined by*

$$\gamma_\Lambda^\Phi(A \mid \omega) \triangleq \frac{\displaystyle\int \mathbb{1}_A(\omega) \exp\left[-H_\Lambda^\Phi(\omega)\right] \prod_{x\in\Lambda} \mu_x^0(\omega_x)}{Z_\Lambda^\Phi\left(\omega_{\Lambda^c}\right)} \tag{3.2.17}$$

for all $A \in \mathcal{F}$ *and all* $\omega \in \Omega$ *is called the* **Gibbs distribution** *on* Ω *in volume* Λ *with boundary condition* ω_{Λ^c} *corresponding to the interaction* Φ *and the a priori measure* μ^0. *The specification* γ^Φ *is called the* **Gibbsian specification** *for* Φ *and* μ^0. *A probability measure consistent with* γ^Φ *is called a* **Gibbs measure** *for* Φ *and* μ^0.

Remark 3.2.18
A measure μ is a Gibbs measure for γ^Φ if and only if $\mu\gamma_\Lambda^\Phi = \mu$ for all $\Lambda \in \mathcal{S}$, that is

$$\iint \frac{\mathbb{1}_A(\omega_\Lambda\, \sigma_{\Lambda^c}) \exp\left[-H_\Lambda^\Phi(\omega_\Lambda\, \sigma_{\Lambda^c})\right]}{Z_\Lambda^\Phi\left(\sigma_{\Lambda^c}\right)} \prod_{x\in\Lambda} \mu_x^0(\omega_x)\, \mu(d\sigma) = \mu(A) \tag{3.2.19}$$

for all $A \in \mathcal{F}$ and $\Lambda \in \mathcal{S}$. In honor of Dobrushin Lanford and Ruelle, these equations are called DLR equations.

Theorem 3.2.20 (Gibbs representation) (Sullivan 1973, Kozlov 1974)
Assume that $\Omega = E^{\mathbb{L}}$ *or that for every* $\Lambda \in \mathcal{S}$, *if* $\omega_\Lambda \in \Omega_\Lambda$ *and* $\sigma_{\Lambda^c} \in \Omega_{\Lambda^c}$, *then* $\omega_\Lambda\, \sigma_{\Lambda^c} \in \Omega$. *Let* γ *be a specification and let* μ^0 *be a product measure. Then the following are equivalent*

(a) *There exists an absolutely summable interaction* Φ *such that* γ *is the Gibbsian specification for* Φ *and* μ^0.

(b) γ *is continuous and is uniformly nonnull with respect to* μ^0.

Moreover if E *is finite then these are also equivalent to the following*

(c) γ *is continuous and non-null with respect to* μ^0.

3.3 Uniqueness

If the specification γ is continuous and the space Ω is compact, then there always exists at least one compatible Gibbs measure. Indeed, the probability measures on a compact space form a (weakly) compact set. Hence, if $(\Lambda_n)_{n\in\mathbb{N}} \subset \mathcal{S}$ is any exhausting sequence of regions and $(\omega^{(n)})_{n\in\mathbb{N}} \subset \Omega$ any sequence of boundary conditions, the sequence of measures $\gamma_{\Lambda_n}(\,\cdot \mid \omega^{(n)})$, $n \in \mathbb{N}$, has some accumulation point. Continuity ensures that such a limit belongs to $\mathcal{G}(\gamma)$.

3.3.1 Dobrushin uniqueness

Let γ be a specification and fix $i \in \mathbb{L}$. By assertion (b) of Definition 3.1.1, $\gamma_{\{i\}}$ is uniquely determined by is projection over $\{i\}$ that is

$$\overset{\circ}{\gamma}_{\{i\}} \left(A \mid \omega\right) \triangleq \gamma_{\{i\}} \left(\{\sigma_i \in A\} \mid \omega\right), \quad \forall\, A \in \mathcal{E},\ \forall\, \omega \in \Omega. \tag{3.3.1}$$

So $\overset{\circ}{\gamma}_{\{i\}}$ is a probability kernel over $(E, \mathcal{E}) \times \left(\Omega, \mathcal{F}_{\{i\}^c}\right)$.
Using notation (1.1.9), the ω_j-dependency of $\overset{\circ}{\gamma}_{\{i\}} \left(\,\cdot \mid \omega_{-\infty}^{i-1}\right)$ for all $j \in \mathbb{L}$ can be estimated by

$$C_{ij}(\gamma) \triangleq \sup\left\{ \left\| \overset{\circ}{\gamma}_{\{i\}} \left(\,\cdot \mid \xi\right) - \overset{\circ}{\gamma}_{\{i\}} \left(\,\cdot \mid \eta\right) \right\| : \xi,\, \eta \in \Omega,\ \xi \overset{\neq j}{=} \eta \right\}. \tag{3.3.2}$$

We can define the **interdependence matrix of Dobrushin** by

$$C(\gamma) \triangleq \left(C_{ij}(\gamma)\right)_{i,j\in\mathbb{L}}. \tag{3.3.3}$$

Definition 3.3.4
A specification γ will be said to satisfy **Dobrushin's condition** *if γ is continuous and satisfies*

$$c(\gamma) \triangleq \sup_{i\in\mathbb{L}} \sum_{j\in\mathbb{L}} C_{ij}(\gamma) < 1. \tag{3.3.5}$$

Theorem 3.3.6 (Dubrushin 68, Lanford 73)
If γ satisfies Dobrushin's condition, then there exists at most one Gibbs measure consistent with γ.

Remark 3.3.7
Dobrushin's condition requires the continuity of the specification but not its non-nullness. If the configuration space is compact and γ is continuous, then (3.3.5) implies the existence of **exactly one** consistent (Gibbs) measure.

Dobrushin's uniqueness result has been extended to general metric spaces by Dobrushin and Shlosman (1985). Let us consider a bounded distance d on E (there is always one, for instance the discrete distance $d_{\text{disc}}(a,b) = 1$ if $a \neq b$ and $d_{\text{disc}}(a,b) = 0$ if $a = b$). For every $i \in \mathbb{L}$ and every function $h : \Omega \to \mathbb{R}$, $\mathcal{F}_{\leq i}$-measurable, we define the d**-oscillation** of h with respect to the site $j \leq i$ by

$$\delta_j^d(h) \triangleq \sup\left\{ \frac{|h(\xi) - h(\eta)|}{d\left(\xi_j, \eta_j\right)} : \xi, \eta \in \Omega_{-\infty}^i,\ \xi \overset{\neq j}{=} \eta \right\}, \tag{3.3.8}$$

with the convention $0/0 = 0$ (and where $\xi \overset{\neq j}{=} \eta$ means that ξ equals η off j). In particular if E is finite, then it will be equipped with the discrete metric and the oscillation with respect to the site j will be denoted by δ_j.

Using notation (5.1.21), for all $i \in \mathbb{L}$ and for all $j \neq i$, the ω_j-dependency of $\overset{\circ}{\gamma}_{\{i\}} \left(\cdot \,\middle|\, \omega_{-\infty}^{i-1} \right)$ can be estimated by the coefficients

$$C_{ij}^d(\gamma) \triangleq \sup\left\{ \frac{\left\| \overset{\circ}{\gamma}_{\{i\}}\left(\cdot \mid \xi\right) - \overset{\circ}{\gamma}_{\{i\}}\left(\cdot \mid \eta\right) \right\|_d}{d\left(\xi_j, \eta_j\right)} : \xi,\, \eta \in \Omega,\ \xi \overset{\neq j}{=} \eta \right\}. \tag{3.3.9}$$

Theorem 3.3.10 (Dobrushin and Shlosman (1985))
Let d be as above. If γ is continuous and satisfies

$$\sup_{i\in\mathbb{L}} \sum_{j\in\mathbb{L}} C_{ij}^d(\gamma) < 1, \tag{3.3.11}$$

then there exists at most one Gibbs measure consistent with γ.

Remark 3.3.12
In particular if d is the discrete metric on E, the VKR-norm coincides with the variational norm and conditions (3.3.5) and (3.3.11) are identical.

3.3.2 Uniform uniqueness condition

Theorem 3.3.13

Let γ be any specification. Assume there exists a constant $c > 0$ satisfying the following property: For every cylinder set $A \in \mathcal{F}$ there exists $\Lambda \in \mathcal{S}_b$ such that

$$\gamma_\Lambda(A \mid \xi) \geq c\, \gamma_\Lambda(A \mid \eta) \quad \text{for all } \xi, \eta \in \Omega, \tag{3.3.14}$$

then there exists at most one Gibbs measure consistent with γ.

Remarks

3.3.15 No assumption over the continuity or the non-nullness of the specification is made.

3.3.16 In practice the uniform uniqueness criterion is only useful in one-dimension.

3.3.17 Under a compactness assumption, there exists exactly one Gibbs measure consistent with γ.

3.3.18 The uniform uniqueness criterion is optimal. Indeed, let $\mathbb{L} = \mathbb{Z}$ and Φ be a shift-invariant potential defined by

$$\Phi_A(\omega) = \begin{cases} |i-j|^{-p}\ \varphi(\omega_i, \omega_j) & \text{if } A = \{i,j\}\ i \neq j \\ 0 & \text{otherwise,} \end{cases} \tag{3.3.19}$$

where φ is a measurable, symmetric and bounded function and $p > 1$ (Φ is absolutely summable if and only if $p > 1$). If $p > 2$ then γ^Φ satisfies uniform uniqueness condition. On the other hand, if $1 < p \leq 2$ a phase transition occurs for a suitable choice of E and φ (see Fröhlich and Spencer (1982) and references therein).

3.4 Decay of correlations

Theorem 3.4.1 (Föllmer (1982))

Suppose Ω is compact and γ satisfies Dobrushin's uniqueness condition (3.3.5), and let μ the unique probability measure consistent with γ. Then for all continuous functions h_1 and h_2 we have

$$|\mu(h_1\, h_2)| = \frac{1}{4} \sum_{i,j \in \mathbb{L}} \delta_i(h_1)\, \delta_j(h_2)\ \left[\frac{1}{1 - C(\gamma)}\right]_{ij}, \tag{3.4.2}$$

where the matrix $\left[\frac{1}{1-C(\gamma)}\right]$ is defined componentwisely by the corresponding series:

$$\left[\frac{1}{1 - C(\gamma)}\right]_{ij} \triangleq \sum_{k \geq 0} \left[(C(\gamma))^k\right]_{ij}. \tag{3.4.3}$$

A similar statement involving δ^d, C_{ij}^d and (3.3.11) is also valid, but there seems to be no published reference.

In chapter 5, we will adapt uniqueness and decay of correlations results to the theory of chains. For that we will need to extend the notion of system of transition probabilities to probability kernels for which the conditioning depends only on the left side. This is what we will do in the next chapter.

Chapter 4

Left interval-specifications and chains

We introduce an statistical mechanical formalism for the study of discrete-time stochastic processes with which we prove general properties of extremal chains, including triviality on the tail σ-algebra, short-range correlations, realization via infinite-volume limits and ergodicity.

The theory of chains with complete connections 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 chapter we propose a natural way to 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.

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 provides 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 chapter 5 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 chapter 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 (Bramson and Kalikow (1993), or Lacroix (2000)). Statistical mechanics is the framework developed, precisely, to study this phenomenon which corresponds to the appearance of (first-order) phase transitions.

In this chapter, $(E, \mathcal{E})$ is a measurable space and Ω a subset of $E^{\mathbb{Z}}$. 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 4.2.9 where $\mathbb{Z}$ acts by isomorphisms.

4.1 Definitions

Definition 4.1.1 (LIS)
A **left interval-specification** *f on $(\Omega, \mathcal{F})$ is a family of probability kernels $\{f_\Lambda\}_{\Lambda \in \mathcal{S}_b}$, $f_\Lambda : \mathcal{F}_{\leq m_\Lambda} \times \Omega \longrightarrow [0,1]$ such that for all Λ in $\mathcal{S}_b$,*

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

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

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

$$f_\Delta f_\Lambda \;=\; f_\Delta \quad \text{on } \mathcal{F}_{\leq m_\Lambda}\,, \tag{4.1.2}$$

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

These conditions are analogous to those defining a *specification* in the theory of Gibbs measures (see Georgii, 1988, 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*.

Definition 4.1.3 (left interval-consistency)
A probability measure μ on $(\Omega,\ \mathcal{F})$ is said to be **consistent** *with a LIS f if for each $\Lambda \in \mathcal{S}_b$*

$$\mu f_\Lambda \ = \mu \quad \text{on } \mathcal{F}_{\leq m_\Lambda}\ . \tag{4.1.4}$$

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 $\mathcal{G}(f)$.*

Remarks

4.1.5 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*.

4.1.6 *Chains with complete connections* is the original nomenclature introduced by Onicescu and Mihoc (1935) . 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* (Harris, 1955), *g-measures* (Keane, 1972), list processes (Lalley, 1986), *uniform martingales* or *random Markov processes* (Kalikow, 1990).

4.2 Main results

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

Theorem 4.2.1 (Singleton consistency for chains)
Let $(f_i)_{i\in\mathbb{Z}}$ be a family of probability kernels $f_i : \mathcal{F}_{\leq i} \times \Omega \to [0,1]$ such that for each $i \in \mathbb{Z}$

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

(b) *For each $B \in \mathcal{F}_{\leq i-1}$ 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

$$f_\Lambda \ = \ f_{l_\Lambda}\, f_{l_\Lambda+1} \cdots f_{m_\Lambda} \tag{4.2.2}$$

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

$$\mathcal{G}(f) \ = \ \Big\{\mu : \mu f_i = \mu,\ \text{for all } i \text{ in } \mathbb{Z}\Big\}\ . \tag{4.2.3}$$

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

$$f_\Lambda \;=\; f_{\{l_\Lambda\}}\, f_{\{l_\Lambda+1\}} \cdots f_{\{m_\Lambda\}} \tag{4.2.4}$$

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

$$f_{[l,m]} \;=\; f_{[l,n]}\, f_{[n+1,m]} \tag{4.2.5}$$

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

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 4.3, are patterned on the Gibbsian proofs, taking care of the one-sided measurability of the LIS kernels.

Theorem 4.2.6 (Extremality and triviality)
Let $f = (f_\Lambda)_{\Lambda\in\mathcal{S}_b}$ be a left interval-specification on $(\Omega, \mathcal{F})$. Denote by $\mathcal{F}_{-\infty} \triangleq \bigcap_{k\in\mathbb{Z}} \mathcal{F}_{\le k}$ 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 $h \ge 0$ 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 4.2.7 (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) $\lim_{\Lambda\uparrow\mathbb{Z}} \sup_{B\in\mathcal{F}_{\Lambda_-}} \mid \mu(A\cap B) - \mu(A)\mu(B) \mid = 0$, *for all cylinder sets A in $\mathcal{F}$.*

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

Theorem 4.2.8 (Infinite volume limits)
Let f be a LIS, μ an extreme point of $\mathcal{G}(f)$ and $(\Lambda_n)_{n\ge1}$ 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.*

The following theorem is the only result of our theory of LIS 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 $\omega \in \Omega$ $\tau(\omega) = (\omega_{i-1})_{i\in\mathbb{Z}}$, 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 4.2.9 (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 μ, $\nu \in \mathcal{P}_{\text{inv}}(\Omega, \mathcal{F})$ and $0 < s < 1$ are such that $s\,\mu + (1-s)\,\nu \in \mathcal{G}_{\text{inv}}(f)$ then μ, $\nu \in \mathcal{G}_{\text{inv}}(f)$.*

Remark 4.2.10
If the LIS is shift-invariant we obtain a dynamical system setting by considering $g(x) = f_{\{0\}}\left(x_0 \mid x_{-\infty}^{-1}\right)$, $x \in \Omega_{-\infty}^0$ and T equal to the right-shift over $\Omega_{-\infty}^0$. In this setting the "singleton" LIS in 0, $f_{\{0\}}$ coincides with the **transfer operator** that is for each $x \in \Omega_{-\infty}^0$ and $\mathcal{F}_{\leq 0}$-measurable function h

$$\left(L_g h\right)(x) \triangleq \sum_{y \in T^{-1}x} g(y)\, h(y) = f_{\{0\}}\left(h \mid x\right)$$

and the n-iterates of the transfer operator L_g^n satisfies

$$\left(L_{f_0}^n h\right)(x) \triangleq \sum_{y \in T^{-n}x} f_{[-n+1,0]}\left(y_{-n+1}^0 \mid y_{-\infty}^{-n}\right) h(y) = f_{[-n+1,0]}\left(h \mid x\right).$$

4.3 Proofs

4.3.1 Singleton consistency for chains

The fact that the objects defined by (4.2.2) are kernels from $\mathcal{F}_{\le m_\Lambda} \times \Omega$ 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

$$f_{\{i\}}(1 \mid \cdot\,) \;=\; f_{\{i\}}\big(\Omega_{\le i} \bigm| \cdot\,\big) \;=\; 1$$

and the inductive step

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

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

Properties (a) and (b) of the definition 4.1.1 of LIS are an immediate consequence of similar properties of the kernels f_i. To prove consistency, we first remark that for $l \le m \le p$, $\omega \in \Omega$ and any $\mathcal{F}_{\le p}$-measurable function h,

$$\begin{aligned}\big(f_{[l,m]}\, f_{[l,p]}\big)(h \mid \omega) \;&=\; f_{[l,m]}\Big(f_{[l,p]}(h) \Bigm| \omega\Big)\\ &=\; f_{[l,p]}(h \mid \omega)\; f_{[l,m]}(1 \mid \omega)\\ &=\; f_{[l,p]}(h \mid \omega)\,. \end{aligned} \tag{4.3.1}$$

The second equality is due to the proven property (b) of Definition 4.1.1 plus the fact that $f_{[l,p]}\,(h \mid \cdot\,)$ is $\mathcal{F}_{\le l-1}$-measurable. The last equality is the just proven normalization. Identity (4.3.1) justifies the last equality in the following string of identities, valid for $l \le m < p$,

$$f_{[l,p]}\, f_{[l,m]} \;=\; f_{[l,m]}\, f_{[m+1,p]}\, f_{[l,m]} \;=\; f_{[l,m]}\, f_{[l,p]} \;=\; f_{[l,p]}\,. \tag{4.3.2}$$

The other equalities are simply due to definition 4.2.2. A similar identity is trivially true for $l \le m = p$. 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 (4.3.2) in the middle identity and we assumed $l_\Delta < l_\Lambda$, otherwise we revert to (4.3.2).

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

$$\mu\, f_i = \mu\,,\ \forall i \in \Lambda \quad \Longrightarrow \quad \mu\, f_\Lambda \;=\; \mu\,. \tag{4.3.3}$$

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 (4.3.3) directly proves the non-trivial inclusion in (4.2.3). 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 (4.3.3) g_Λ must be consistent with f_Λ for each $\Lambda \in \mathcal{S}_b$. But then, if $\omega \in \Omega$ and h is $\mathcal{F}_{\leq m_\Lambda}$-measurable

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

The second identity is a consequence of the $\mathcal{F}_{l_\Lambda-1}$-measurability of $f_\Lambda(h|\cdot\,)$ plus property (b) of Definition 4.1.1. The last equality is the normalization of g_Λ. □

4.3.2 Extreme chains

We start with general results on probability kernels.

Proposition 4.3.4
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*

$$\mathcal{I}^{\mathcal{B}}_\pi(\mu) \;\triangleq\; \Big\{A \in \mathcal{B} : \pi(A \mid \cdot\,) = \mathbb{1}_A(\,\cdot\,)\ \mu\text{-a.s.}\Big\} \tag{4.3.5}$$

is a σ-algebra.

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

$$(h\mu)\,\pi = h\mu \text{ on } \mathcal{B} \quad \text{if and only if} \quad h \text{ is } \mathcal{I}^{\mathcal{B}}_\pi(\mu)\text{-measurable}\,. \tag{4.3.6}$$

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

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

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

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

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

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

and, by the consistency of μ with π,

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

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 $\{h\ge c\}\in\mathcal{I}^{\mathcal{B}}_\pi(\mu)$, for all $c>0$. Let us fix some $c>0$ and denote $g = \mathbb{1}_{h\ge c}$. We have

$$\begin{aligned}\mu\Big((1-g)\,h\,\pi(g)\Big) \;&=\; (h\,\mu)\big(\pi(g)\big) - \mu\big(g\,h\,\pi(g)\big) \;=\; (h\,\mu)(g) - \mu\big(g\,h\,\pi(g)\big)\\ &=\; \mu\Big(g\,h\,\big(1-\pi(g)\big)\Big)\,.\end{aligned}$$

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

$$\begin{aligned}\mu\Big((1-g)\,h\,\pi(g)\Big) \;&\ge\; c\,\mu\Big(g\big(1-\pi(g)\big)\Big) \;=\; c\,\mu\big(\pi(g)\big) - c\,\mu\big(g\,\pi(g)\big)\\ &=\; c\,\mu\Big((1-g)\,\pi(g)\Big)\,.\end{aligned}$$

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

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

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

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

$$\begin{aligned}(\mathbb{1}_A\,\mu)\,\pi(B) \;&=\; (\mathbb{1}_A\,\mu)\,\pi(A\cap B) + (\mathbb{1}_A\,\mu)\,\pi(B\setminus A)\\ &\le\; \mu\pi(A\cap B) + (\mathbb{1}_A\,\mu)\,\pi\left(A^c\right)\,.\end{aligned}$$

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

$$(\mathbb{1}_A\mu)\,\pi(B) \;\le\; (\mathbb{1}_A\mu)(B)\,. \tag{4.3.7}$$

By the same token,

$$(\mathbb{1}_A\mu)\,\pi\left(B^c\right) \;\le\; (\mathbb{1}_A\mu)\left(B^c\right)\,. \tag{4.3.8}$$

But the consistency of μ with π implies that the sum of the LHS of (4.3.7) and (4.3.8) equals the sum of the corresponding RHS, namely $\mu(A)$. We conclude that $(\mathbb{1}_A\mu)\,\pi(B) = (\mathbb{1}_A\mu)(B)$. $\square$

Corollary 4.3.9
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

$$\mathcal{G}(\Pi) \;=\; \Big\{\mu \in \mathcal{P}(\Omega,\mathcal{F}) : \mu\,\pi = \mu \text{ on } \mathcal{F}_\pi \text{ for all } \pi \in \Pi\Big\} \tag{4.3.10}$$

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

$$\mathcal{I}_\Pi(\mu) \;=\; \bigcap_{\pi\in\Pi} \mathcal{I}_\pi^{\mathcal{F}_\pi}(\mu) \tag{4.3.11}$$

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

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

and

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

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 4.3.4 implies that ν, $\nu' \in \mathcal{G}(\Pi)$, a fact that contradicts the extremality of μ. $\square$

Lemma 4.3.12
Let f be a LIS defined on $(\Omega,\mathcal{F})$ and $\mu \in \mathcal{G}(f)$. Let us denote $\mathcal{F}_{-\infty}^\mu$ the μ-completion of $\mathcal{F}_{-\infty}$. Then

$$\bigcap_{n\geq 0} \mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leq k}}(\mu) \;=\; \mathcal{F}_{-\infty}^\mu \tag{4.3.13}$$

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

$$\bigcap_{\Lambda\in\mathcal{S}_b} \mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\leq m_\Lambda}}(\mu) \;=\; \mathcal{F}_{-\infty}^\mu\,. \tag{4.3.14}$$

Proof
Identity (4.3.13) follows from the observation that for each $B \in \bigcap_n \mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\leq k}}(\mu)$ the set $A \triangleq \bigcap_n \big\{f_{[k-n,k]}(B \mid \cdot\,) = 1\big\}$ satisfies $A = B$ μ-a.s. and $A \in \mathcal{F}_{-\infty}$. Equality (4.3.14) is a consequence of (4.3.13) because

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

Proof of Theorem 4.2.6

(a) It is immediate.

(b) ($\Rightarrow$) The implication follows readily from Corollary 4.3.9 and the fact that, by (4.3.14), $\bigcap_{\Lambda\in\mathcal{S}_b}\mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\le m_\Lambda}}(\mu)$ 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 $k\in\mathbb{Z}$ $\mu_k\triangleq\mu\big|_{\mathcal{F}_{\le k}}$ and $\nu_k\triangleq\nu\big|_{\mathcal{F}_{\le k}}$. As in particular $\nu_k\ll\mu_k$ on $\mathcal{F}_{\le k}$, there exists $g_k\ge 0$, $\mathcal{F}_{\le k}$-measurable, satisfying $\nu_k=g_k\,\mu_k$ on $\mathcal{F}_{\le k}$). All we have to prove is that

$$g_k \text{ is } \mathcal{F}_{-\infty}^{\mu}\text{-measurable}\quad\forall\,k\in\mathbb{Z}\,. \tag{4.3.15}$$

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

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

on $\mathcal{F}_{\le k}$ for all $n\in\mathbb{N}$. As g_k is $\mathcal{F}_{\le k}$-measurable, we conclude from Proposition 4.3.4 that g_k is $\bigcap_n\mathcal{I}_{f_{[k-n,k]}}^{\mathcal{F}_{\le k}}(\mu)$-measurable. Its, $\mathcal{F}_{-\infty}^{\mu}$-measurability follows, hence, from (4.3.13).

(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 $h,h'\ge 0$ 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 4.3.4 plus the fact that h is $\mathcal{I}_{f_\Lambda}^{\mathcal{F}_{\le m_\Lambda}}(\mu)$-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). □

4.3.3 Triviality and short-range correlations

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

Proof of Proposition 4.2.7

(a) $\Rightarrow$ **(c)** Let $A\in\mathcal{F}$ and $k\in\mathbb{Z}$. Since $\mathcal{F}_{-\infty}=\bigcap_{n\ge1}\mathcal{F}_{\le k-n}$, the reverse martingale theorem yields

$$\mu\big(A\bigm|\mathcal{F}_{\le k-n}\big)\ \xrightarrow[n\to+\infty]{L^1(\mu)}\ \mu\big(A\bigm|\mathcal{F}_{-\infty}\big)\,. \tag{4.3.16}$$

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

$$\mu\Big(\left|\mu\left(A \mid \mathcal{F}_{\Delta_-}\right) - \mu(A)\right|\Big) \; < \; \varepsilon\,. \tag{4.3.17}$$

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

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

(b) $\Rightarrow$ (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 4.2.8

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

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

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. □

4.3.4 Ergodicity

We need a well known result of ergodic theory. See, for instance, Georgii (1988), Theorem 14.5, for a proof.

Theorem 4.3.18

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

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

Lemma 4.3.19
Let $\mu \in \mathcal{P}_{inv}(\Omega, \mathcal{F})$, 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 $(B_n)_{n\geq 1}$ be a sequence of cylinder sets such that $\mu(A\Delta B_n) \leq 2^{-n}$ for all $n \geq 1$. Since $\mu \in \mathcal{P}_{\mathrm{inv}}(\Omega, \mathcal{F})$, we have that

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

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 $n \geq 1$ we choose $i(n) \geq 0$ 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 $\mu\left(A \Delta C_n\right) \leq 2^{-n}$. Therefore, the set $C \triangleq \bigcap_{m\geq 1} \bigcup_{n \geq m} C_n$ belongs to $\mathcal{F}_{-\infty}$ and satisfies

$$\mu(A\Delta C) \;\leq\; \mu\left(\bigcap_{m\geq 1} \bigcup_{n\geq m} A \Delta C_n\right) \;\leq\; \lim_{m\to +\infty} \sum_{n\geq m} 2^{-n} \;=\; 0\,. \quad \square$$

Proof of Theorem 4.2.9
(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

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

By (4.3.14) and Lemma 4.3.19, $\mathcal{K}(\mu)$ is the μ-completion of $\mathcal{I}$. Therefore Corollary 4.3.9 implies that each μ extreme in $\mathcal{G}_{\mathrm{inv}}(f)$ 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 $\nu,\, \nu' \in \mathcal{G}_{\mathrm{inv}}(f)$. Then there exist $\mathcal{F}$-measurable $h,\, h' \geq 0$ such that $\nu = h\mu$ and $\nu' = h'\mu$. Since $\mu,\, \nu,\, \nu' \in \mathcal{P}_{\mathrm{inv}}(\Omega, \mathcal{F})$, Proposition 4.3.4 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 4.3.18 (b) implies that there exists $h \geq 0$, $\mathcal{I}$-measurable such that $\nu = h\mu$. By Lemma 4.3.19 h is $\mathcal{F}_{-\infty}$-measurable, so Theorem 4.2.6 b) implies that $\nu \in \mathcal{G}(f)$. Therefore $\nu \in \mathcal{G}_{\mathrm{inv}}(f)$.

(c) It is an immediate consequence of (b). $\square$

Chapter 5

Uniqueness, loss of memory and decay of correlations

We present two types of results. First, 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 Georgii (1974) for one-dimensional Gibbs fields. This criterion is known to be optimal for the latter, in the sense that 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 criterion, which corresponds to a well known uniqueness criterion in statistical mechanics (see, for instance, Simon, 1993, Chapter V). This criterion is valid for systems with a compact metric alphabet. We exhibit simple examples where Dobrushin criterion applies but that fall outside the scope of most other known uniqueness criteria (Harris, 1955; Iosifescu and Spataru, 1973; Walters, 1975; Berbee, 1987; Stenflo, 2002; Johansson and Öberg, 2002). We also discuss a slight enhancing of a criterion due to Hulse (1997), which is stronger than Dobrushin's for a finite alphabet and is also valid for countable alphabets.

Our second 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 Simon, 1993), are complementary, both in their precision and in their range of applicability, to similar results available in the literature (Iosifescu, 1992; Bressaud, Fernández and Galves, 1999 and references therein). The results depend on a *sensitivity matrix* that generalizes the Dobrushin ergodic coefficient of Markov chains.

In this chapter, $(E, \mathcal{E})$ is a measurable space and Ω a subset of $E^{\mathbb{Z}}$. 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.

5.1 Uniqueness results

We shall prove two types of uniqueness results. We start with the counterpart of a criterion proven by Georgii (1974) for measures determined by specifications.

Theorem 5.1.1 (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

$$f_{[n,m]}(A \mid \xi) \;\geq\; c\, f_{[n,m]}(A \mid \eta) \quad \text{for all } \xi, \eta \in \Omega\,. \tag{5.1.2}$$

Then there exists at most one chain consistent with f.

The main virtue of this criterion is its generality. Existing uniqueness criteria (Harris, 1955; Iosifescu and Spataru, 1973; Walters, 1955; Berbee, 1987; Hulse, 1997; Stenflo, 2002; Johansson and Öberg, 2002) require that the space E have particular properties (finite, countable, compact), and that the kernels satisfy appropriate non-nullness hypotheses. Many of these criteria are based on summability properties of the sequence of variations:

$$\mathrm{var}_j(f_{\{i\}}) \;\triangleq\; \sup\Big\{\big|f_{\{i\}}(\xi_i \mid \xi_{-\infty}^{i}) - f_{\{i\}}(\eta_i \mid \eta_{-\infty}^{i})\big| : \xi, \eta \in \Omega_{-\infty}^{i},\, \xi_j^i = \eta_j^i\Big\} \tag{5.1.3}$$

for $j < i$.

Proposition 5.1.4
Assume that E is a countable set and $\mathcal{E}$ the discrete σ-algebra. A LIS f satisfies the one-sided boundary-uniformity condition (5.1.2) if it is **uniformly non-null***:*

$$\inf_{i \in \mathbb{Z}} \inf_{\omega \in \Omega_{\leq i}} f_{\{i\}}\big(\omega_i \bigm| \omega_{-\infty}^{i-1}\big) \;>\; 0\,, \tag{5.1.5}$$

and satisfies

$$\sup_{n \in \mathbb{Z}} \sum_{i \geq n} \mathrm{var}_n\left(f_{\{i\}}\right) \;<\; +\infty\,. \tag{5.1.6}$$

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 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 $d_{\rm disc}(a,b) = 1$ if $a \neq b$ and 0 otherwise.

Definition 5.1.7
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 in presence of non-nullness (see Theorem 3.2.20). In the particular case where E finite, continuity is equivalent to $\lim_{j\to-\infty} \mathrm{var}_j \left(f_{\{i\}}\right) = 0$.

Remark 5.1.8
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, the following theorems determine conditions for the existence of *exactly* one compatible measure.

For every $i \in \mathbb{Z}$ and every $\mathcal{F}_{\leq i}$-measurable function h, the d-**oscillation** of h with respect to the site $j \leq i$, is defined by

$$\delta_j^d(h) \;\triangleq\; \sup\left\{ \frac{|h(\xi) - h(\eta)|}{d\left(\xi_j, \eta_j\right)} : \xi, \eta \in \Omega_{-\infty}^i \,,\, \xi \overset{\neq j}{=} \eta \right\} \;, \tag{5.1.9}$$

with the convention $0/0 = 0$ and where we introduced the notation

$$\xi \overset{\neq j}{=} \eta \quad \Longleftrightarrow \quad \xi_i = \eta_i \;,\; \forall\, i \neq j \tag{5.1.10}$$

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

$$\mathcal{B}_d \;\triangleq\; \left\{ \mathcal{F}\text{-measurable } h : \sup_{j\in\mathbb{Z}} \delta_j^d(h) < \infty \right\} \;, \tag{5.1.11}$$

and its restrictions

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

for $\Lambda \subset \mathbb{Z}$. We start with a criterion valid when the alphabet E is countable and discrete. It is a slight enhancing of a result due to Hulse (1997), without shift-invariance hypotheses and with weaker non-nullness requirements. Let us denote

$$\delta_j(h) \;\triangleq\; \delta_j^{d_{\rm disc}}(h) \;=\; \sup\left\{ |h(\xi) - h(\eta)| : \xi, \eta \in \Omega \,,\, \xi \overset{\neq j}{=} \eta \right\} . \tag{5.1.12}$$

Let us say that a LIS f on a $(\Omega, \mathcal{F})$ is **weakly uniformly non-null** if there exists $\varepsilon > 0$ such that

$$\sum_{\sigma_i \in \Omega_{\{i\}}} \min_{\omega \in \Omega^i_{-\infty} : \omega_i = \sigma_i} f_{\{i\}}(\sigma_i \mid \omega^{i-1}_{-\infty}) \;\geq\; \varepsilon \tag{5.1.13}$$

for each $i \in \mathbb{Z}$. We also introduce the singleton functions $f_i \in \mathcal{F}_{\leq i}$, $i \in \mathbb{Z}$, defined by

$$f_i\big(w^i_{-\infty}\big) \;\triangleq\; f_{\{i\}}\big(w_i \,\big|\, w^{i-1}_{-\infty}\big)\;. \tag{5.1.14}$$

Theorem 5.1.15 (Enhanced Hulse)
Assume that E is a countable set with the discrete topology and σ-algebra. Suppose that f is a LIS on $(\Omega, \mathcal{F})$ that is continuous and weakly uniformly non-null, and such that for each $i \in \mathbb{Z}$,

$$\sum_{k \geq 0} \delta_{i-k}\,(f_i) \;<\; +\infty\;. \tag{5.1.16}$$

Then there exists at most one chain consistent with f.

The most general version of Dobrushin's strategy allows the use of a "pavement" 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 5.1.17 (d-sensitivity estimator)
*Let $V \in \mathcal{S}_b$ and f_V a probability kernel on $\mathcal{F}_{\leq m_V} \times \Omega$. A **$d$-sensitivity estimator** for f_V is a nonnegative matrix $\alpha^V = \big(\alpha^V_{ij}\big)_{i,j\in\mathbb{Z}}$ such that $\alpha^V_{ij} = 0$ if $i \notin V$ or $j \notin V_-$ and*

$$\delta^d_j\,(f_V h) \;\leq\; \sum_{i \in V} \delta^d_i(h)\,\alpha^V_{ij} \tag{5.1.18}$$

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

Theorem 5.1.19 (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

$$\sum_{j \in V_-} \alpha^V_{ij} \;<\; 1 \tag{5.1.20}$$

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

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

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 $\overset{\circ}{f}_V$ the projection of each kernel f_V over Ω_V:

$$\overset{\circ}{f}_V\left(A \mid \omega_{-\infty}^{l_V-1}\right) \;\triangleq\; f_V\big(\{\sigma_V\in A\}\mid \omega_{-\infty}^{l_V-1}\big),\quad \forall\, A\in\mathcal{F}_V,\ \forall\,\omega_{-\infty}^{l_V-1}\in\Omega_{-\infty}^{l_V-1}\ ,$$

then the VKR distances between these projections are

$$\begin{aligned}\Big\|\overset{\circ}{f}_V\left(\,\cdot\mid \xi_{-\infty}^{l_V-1}\right)-\overset{\circ}{f}_V\left(\,\cdot\mid \eta_{-\infty}^{l_V-1}\right)\Big\|_{d_V} &=\\ \sup\Big\{\Big|\overset{\circ}{f}_V\left(h\mid \xi_{-\infty}^{l_V-1}\right)-\overset{\circ}{f}_V\left(h\mid \eta_{-\infty}^{l_V-1}\right)\Big| &: h\in\mathcal{B}_d(V)\,,\ \mathrm{osc}_V(h)\leq 1\Big\}\end{aligned} \tag{5.1.21}$$

where $\mathrm{osc}_V(h)=\sup\big\{|h(\sigma_V)-h(\omega_V)|/d_V(\sigma_V,\omega_V)\big\}$. Equivalently (see, for instance, Dudley, 2002, Section 11.8),

$$\begin{aligned}\Big\|\overset{\circ}{f}_V\left(\,\cdot\mid \xi_{-\infty}^{l_V-1}\right)-\overset{\circ}{f}_V\left(\,\cdot\mid \eta_{-\infty}^{l_V-1}\right)\Big\|_{d_V} &=\\ \inf\Big\{\int d(\sigma_V,\omega_V)\,\rho(d\sigma_V,d\omega_V) &: \rho\in\mathcal{P}(\Omega\times\Omega)\\ \text{with marginals }\ \overset{\circ}{f}_V\left(\,\cdot\mid \xi_{-\infty}^{l_V-1}\right) &\text{ and }\ \overset{\circ}{f}_V\big(\,\cdot\mid \eta_{-\infty}^{l_V-1}\big)\Big\}.\end{aligned} \tag{5.1.22}$$

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

$$C_{ij}^V(f)\;\triangleq\;\sup_{\substack{\xi,\eta\in\Omega_{-\infty}^{l_V-1}\\ \xi\overset{\neq j}{=}\eta}}\frac{\Big\|\overset{\circ}{f}_V\left(\,\cdot\mid\xi\right)-\overset{\circ}{f}_V\left(\,\cdot\mid\eta\right)\Big\|_{d_V}}{d\left(\xi_j,\eta_j\right)}\ ,\quad i\in V, j\in V_- \tag{5.1.23}$$

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

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

$$C_{ij}^{\{i\}}(f)\;=\;\delta_j(f_i) \tag{5.1.24}$$

and a sufficient condition for Dobrushin criterion (5.1.2) is, therefore,

$$\sum_{j<i}\delta_j(f_i)\;<\;1\ ,\quad i\in\mathbb{Z}\ . \tag{5.1.25}$$

This explicitly shows the difference between Hulse's and Dobrushin's criteria for these particular systems. Besides the absence of non-nullness hypotheses, an advantage of Dobrushin 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 three uniqueness criteria given in Theorems 5.1.1, 5.1.15 and 5.1.19 give a very strong form of uniqueness.

Definition 5.1.26 (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*

$$f_{\Lambda}^{(\Gamma,\omega)}(\,\cdot \mid \xi) \;=\; f_{\Lambda}(\,\cdot \mid \omega_{\Gamma_-}\,\xi_{\Gamma})\,, \quad \Lambda \in \mathcal{S}_b\,,\ \Lambda \subset \Gamma \tag{5.1.27}$$

admits at most one consistent unique chain.

The three criteria given above involve bounds valid for all past conditions. They remain, therefore, valid if only particular pasts are considered as in (5.1.27). This observation proofs the following corollary.

Corollary 5.1.28
If a LIS satisfies the hypotheses of either Theorem 5.1.1 or Theorem 5.1.15 or Theorem 5.1.19, then it also satisfies a HUC.

We remark that, for similar reasons, the criteria of Harris (1955), Stenflo (2002) and Johansson and Öberg (2002) also imply the validity of a HUC.

5.2 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 5.2.1
*A d-***sensitivity matrix** *for a LIS f is a matrix of the form*

$$\alpha_{ij} \;\triangleq\; \begin{cases} \alpha_{ij}^{\{i\}} & \text{if } i > j \\ 0 & \text{otherwise} \end{cases} \tag{5.2.2}$$

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

Theorem 5.2.3 (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 $h \in \mathcal{B}_d(\Lambda)$,*

$$\delta_j^d\,(f_\Lambda h) \;\leq\; \sum_{k\in\Lambda} \delta_k^d(h) \left[\sum_{l=1}^{|\Lambda|} (P_\Lambda \alpha)^l\right]_{kj} \tag{5.2.4}$$

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

$$\gamma_i \;\triangleq\; \sum_{j<i} \alpha_{ij}\, e^{F(i,j)} \;<\; 1\,, \tag{5.2.5}$$

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

$$\delta_j^d\,(f_\Lambda h) \;\leq\; \frac{\gamma_\Lambda}{1-\gamma_\Lambda} \sum_{k\in\Lambda} \delta_k^d(h)\, e^{-F(k,j)}\,, \tag{5.2.6}$$

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

Remarks

5.2.7 In the Markovian case $\alpha_{ij} = 0$ if $|i-j| > 1$. Then expression (5.2.4) implies that for $h \in \mathcal{F}_{\{n\}}$

$$\delta_{-1}^d\Big(f_{[0,n]}(h)\Big) \;\leq\; \gamma^n\, \delta_n^d(h) \tag{5.2.8}$$

with $\gamma = \sup_i \sum_j \alpha_{ij}$. For d discrete and estimators (5.1.23), γ is known as the **Dobrushin ergodic coefficient**. If, in addition, E is countable, $\Omega = E^{\mathbb{Z}}$ and f shift-invariant, then

$$\gamma \;=\; 1 - \min_{\sigma_{-1},\omega_{-1}\in E} \sum_{\omega_0\in E} f_{\{0\}}\big(\omega_0 \bigm| \sigma_{-1}\big) \wedge f_{\{0\}}\big(\omega_0 \bigm| \omega_{-1}\big)\,. \tag{5.2.9}$$

5.2.10 If the alphabet E is countable and the metric discrete we can use the estimators (5.1.24). With this choice, (5.2.6) implies

$$\begin{aligned} &\delta_j(f_i) \;\leq\; \text{const}\, e^{-F(i,j)} \\ &\qquad\Longrightarrow\quad \delta_{-n}[f_{[0,m]}(A)] \;\leq\; \text{const}\, e^{-F(m,-n)}\,,\ A \in \mathcal{F}_{\{m\}}\,. \end{aligned} \tag{5.2.11}$$

Published loss-of-memory results (Iosifescu, 1992; Bressaud, Fernández and Galves, 1999) resort instead to the variations (5.1.3). Comparisons can only be made through the obvious inequalities

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

For LIS with an exponentially decaying dependence on the past, (5.2.11) 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 Bressaud, Fernández and Galves (1999) 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 (5.2.11) is also a power law but with a power decreased by one unit. Bressaud, Fernández and Galves (1999) 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 5.2.3 applies. Consider, for instance, the 2-letter alphabet $E = \{0,1\}$ and a shift-invariant LIS defined by singletons

$$f\big(\omega_0 = 1 \bigm| \omega_{-\infty}^{-1}\big) \;=\; \sum_{i\leq 0} a_i\,\omega_i\;, \tag{5.2.12}$$

for a sequence $\{a_i\}_{i\leq 0}$ of non-negative numbers. The estimators (5.1.24) yield a sensitivity matrix

$$\alpha_{ij} \;=\; \delta_j\left(f_{\{i\}}\right) \;=\; a_{i-j} \tag{5.2.13}$$

for $i > j$, and zero otherwise. Theorem (5.2.3) is therefore applicable as long as $\sum_{i\leq 0} a_i < 1$. On the other hand, for each $0 < \varepsilon < 1$, the choice

$$a_{-k} \;=\; \frac{1-\varepsilon}{M_\varepsilon}\,\frac{1}{k^{1+\varepsilon}} \tag{5.2.14}$$

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

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

for $i - j \geq 2$. Thus, this LIS is not covered by the results of Iosifescu (1992) or of Bressaud, Fernández and Galves (1999). It also does not satisfy any uniqueness criteria except Hulse's.

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 Simon, 1993). Their proofs, presented in Section 5.4, follow the guidelines of the statistical mechanical proofs. They require a compact Ω. We observe that example (5.2.12)–(5.2.14) shows that our results are complementary to those existing in the literature, which are based on variations rather than oscillations (Bressaud, Fernández and Galves, 1999, and references therein).

Theorem 5.2.15
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

$$\left\| \overset{\circ}{f}_{\{i\}} \left(\cdot \mid \omega \right) - \overset{\circ}{\widetilde{f}}_{i} \left(\cdot \mid \omega \right) \right\|_d \;\leq\; b_i(\omega) \tag{5.2.16}$$

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 $\Lambda \in \mathcal{S}_b$

$$\left| \mu(h) - \widetilde{\mu}(h) \right| \;\leq\; \sum_{k \in \Lambda \cup \Lambda_-} \widetilde{\mu}\left(b_k\right)\, \delta_k^d \left(f_{[k+1, m_\Lambda]} h \right) \tag{5.2.17}$$

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

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

Theorem 5.2.18
Assume Ω compact and let f be a LIS on $(\Omega, \mathcal{F})$ that is continuous and 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 $m_\Delta < l_\Lambda$,

$$Cor_\mu \left(h_1, h_2 \right) \;\leq\; \frac{D^2}{4} \sum_{k \leq m_\Delta} \delta_k^d \left(f_{[k+1, m_\Lambda]} h_1 \right) \delta_k^d \left(f_{[k+1, m_\Delta]} h_2 \right) \tag{5.2.19}$$

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

Next corollary offers a more quantitative consequence of this theorem. For all $\Lambda \in \mathcal{S}_b$ we define the Λ-projection

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

For a matrix $\left(A_{kj} \right)_{k,j \in \mathbb{Z}}$ with nonnegative entries, we denote

$$\left[\frac{A}{1 - A} \right]_{kj} \;\triangleq\; \sum_{n \geq 1} \left[A^n \right]_{kj} \;. \tag{5.2.20}$$

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

Corollary 5.2.21

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 $h_2 \in \mathcal{B}_d(]-\infty, m_\Delta])$,*

$$Cor_\mu\,(h_1,h_2) \;\le\; \frac{D^2}{4} \sum_{k\le m_\Delta} \sum_{l\in\Lambda} \delta_l^d(h_1) \left[\frac{P_\Lambda\alpha}{1-P_\Lambda\alpha}\right]_{lk} \delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right)\;. \tag{5.2.22}$$

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

$$Cor_\mu\,(h_1,h_2) \;\le\; \frac{D^2}{4} \sum_{l\in\Delta} \sum_{m\in\Lambda} \delta_m^d(h_1)\, \delta_l^d(h_2)\, A_{ml}\;, \tag{5.2.23}$$

where

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

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

Proposition 5.2.24

If (α_{ij}) is a matrix satisfying (5.2.5), then for each $\Lambda \in \mathcal{S}_b$

$$\left[\frac{(P_\Lambda\alpha)}{1-(P_\Lambda\alpha)}\right]_{kj} \;\le\; \frac{\gamma_\Lambda}{1-\gamma_\Lambda}\; e^{-F(k,j)}\;. \tag{5.2.25}$$

5.3 Proofs on uniqueness

5.3.1 One-sided boundary-uniformity

Lemma 5.3.1

If uniqueness condition (5.1.2) is satisfied, then $\nu \ge c\,\mu$, $\forall\,\mu,\nu \in \mathcal{G}(f)$.

Proof

Let A be a cylinder set and n an integer such that (5.1.2) holds. If μ and ν are consistent with f,

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

Proof of Theorem 5.1.1
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

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

By Theorem 4.2.6 (c), $\nu \in \mathcal{G}(f)$. By the preceding lemma $0 = \nu(B^c) \geq c\,\mu(B^c)$, so $\mu(B) = 1$. □

Proof of Proposition 5.1.4
Call $m(f)$ the infimum (5.1.5) and $V(f)$ the supremum (5.1.6). Through an elementary logarithmic inequality we have that for each $i, j \in \mathbb{Z}$ with $i > j$ and each $\xi, \eta \in \Omega_{\leq i}$ with $\xi_j^i = \eta_j^i$,

$$\frac{f_{\{i\}}\big(\xi_i \mid \xi_{-\infty}^{i-1}\big)}{f_{\{i\}}\big(\eta_i \mid \eta_{-\infty}^{i-1}\big)} \;\geq\; \exp\Big(-\frac{\mathrm{var}_j(f_{\{i\}})}{m(f)}\Big)\,. \tag{5.3.2}$$

Applying the factorization (4.2.4) we conclude that for each $n, m \in \mathbb{Z}$ with $n < m$ and each $\xi, \eta \in \Omega_{\leq m}$ with $\xi_n^m = \eta_n^m$,

$$\frac{f_{[n,m]}\big(\xi_n^m \mid \xi_{-\infty}^{n-1}\big)}{f_{[n,m]}\big(\eta_n^m \mid \eta_{-\infty}^{n-1}\big)} \;\geq\; e^{-V(f)/m(f)}\,. \quad \square \tag{5.3.3}$$

5.3.2 Enhanced Hulse

The criterion relies on the following lemma, whose proof is a straightforward extension of the argument given in Proposition 4.2 in Hulse (1997).

Lemma 5.3.4 (Hulse)
Assume that E is a discrete countable set and let f be a continuous LIS defined on $(\Omega, \mathcal{F})$. Assume that for each $i \in \mathbb{Z}$ and each $\omega, \sigma \in \Omega_{-\infty}^{i-1}$ there exist a coupling $K_{\omega\sigma}^i$ of $f_i(\,\cdot \mid \omega)$ and $f_i(\,\cdot \mid \sigma)$ and a matrix $\{a_{ij} : i, j \in \mathbb{Z}\}$ such that

(a) $K_{\omega\sigma}^i\big(\{\xi_i = \eta_i\}\big) \geq \varepsilon > 0.$

(b) $K_{\omega\sigma}^i\big(\{\xi_i \neq \eta_i\}\big) \leq \sum_{j<i} a_{i\,j}\, d_{\mathrm{disc}}(\omega_j, \sigma_j).$

Then, if $\sum_{j<i} a_{i\,j} < +\infty$ for each $i \in \mathbb{Z}$ there is at most one chain consistent with f.

Proof of Theorem 5.1.15
Let us consider, for each $i \in \mathbb{Z}$ and each $\omega, \sigma \in \Omega_{-\infty}^{i-1}$, the coupling of $f_i(\,\cdot \mid \omega)$ and $f_i(\,\cdot \mid \sigma)$

defined by

$$K^i_{\omega\,\sigma}\,(\xi_i,\eta_i) \;=\; \begin{cases} f_{\{i\}}\,(\xi_i \mid \omega) \wedge f_{\{i\}}\,(\eta_i \mid \sigma) & \text{if } \xi_i = \eta_i \\ \dfrac{\big[f_{\{i\}}\,(\xi_i \mid \omega) - f_{\{i\}}\,(\xi_i \mid \sigma)\big]^+ \big[f_{\{i\}}\,(\eta_i \mid \sigma) - f_{\{i\}}\,(\eta_i \mid \omega)\big]^+}{\sum_{\zeta_i \in E}\big[f_{\{i\}}\,(\zeta_i \mid \omega) - f_{\{i\}}\,(\zeta_i \mid \sigma)\big]^+} & \text{if } \xi_i \neq \eta_i \end{cases} \tag{5.3.5}$$

This is a well known coupling, which is optimal in the sense that it achieves the variational distance:

$$K^i_{\omega\,\sigma}\big(\{\xi_i \neq \eta_i\}\big) \;=\; \frac{1}{2}\sum_{p \in \Omega_i^{\omega\,\sigma}} \Big| f_{\{i\}}(p \mid \omega) - f_{\{i\}}(p \mid \sigma)\Big| \;,$$

where $\Omega_i^{\omega\,\sigma}$ is the set of $\omega_i \in E$ such that both $\omega^i_{-\infty}$ and $\eta_{]-\infty,i-1]}\omega_i$ belong to $\Omega^i_{-\infty}$. The continuity of f therefore implies

$$K^i_{\omega\,\sigma}\big(\{\xi_i \neq \eta_i\}\big) \;\leq\; \sum_{j<i} \frac{|E|}{2}\,\delta_j\,(f_i)\; d_{\rm disc}\,(\omega_j,\sigma_j)\;\;. \tag{5.3.6}$$

On the other hand, the weak uniform non-nullness (5.1.13) of f implies that there exists $\varepsilon > 0$ such that for each $i \in \mathbb{Z}$,

$$K^i_{\omega\,\sigma}\big(\{\xi_i = \eta_i\}\big) \;=\; \sum_{p \in \Omega_i^{\omega\,\sigma}} f_{\{i\}}(p \mid \omega) \wedge f_{\{i\}}(p \mid \sigma) \;\geq\; \varepsilon\;. \tag{5.3.7}$$

The theorem follows from (5.3.7), (5.3.6) and Hulse lemma. $\square$

5.3.3 Dobrushin uniqueness

The following bound is the basic tool of the theory.

Lemma 5.3.8 (Multisite dusting lemma)
Let $V \in \mathcal{S}_b$, f_V a probability kernel on $\mathcal{F}_{\leq m_V} \times \Omega$ and α^V is a d-sensitivity estimator for f. Then,

$$\delta^d_j\,(f_V h) \quad \begin{cases} = \;0 & \text{if } j \in V \\ \leq \;\delta^d_j(h) + \sum_{k\in V} \delta^d_k(h)\,\alpha^V_{kj} & \text{if } j \in V_-\;, \end{cases} \tag{5.3.9}$$

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

Remark 5.3.10
The name of the lemma comes from a picturesque interpretation due to Michael Aizenman reported in Simon (1993): 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 (5.3.9)], but the dust has been spread over the remaining sites [second line of (5.3.9)]. The estimators give the fraction blown from site to site. In this picture, Dobrushin condition (5.1.20) means that some dust stays in the duster, a fact that allows for an eventual total cleaning.

Proof

The first line in (5.3.9) 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 $h_{V,\omega}(\sigma_V) \triangleq h\big(\omega_{V_-}\,\sigma_V\big)$ for each $\omega \in \Omega$ ("freezing" at ω). For $j \in V_-$ and $\xi, \eta \in \Omega_{V_-}$ such that $\xi \stackrel{\neq j}{=} \eta$, we have

$$\Big|f_V(h \mid \xi) - f_V(h \mid \eta)\Big| \le \Big|\overset{\circ}{f}_V\,(h_{V,\xi} - h_{V,\eta} \mid \xi)\Big| + \Big|\overset{\circ}{f}_V\,(h_{V,\eta} \mid \xi) - \overset{\circ}{f}_V\,(h_{V,\eta} \mid \eta)\Big| . \tag{5.3.11}$$

If we divide throughout by $d(\xi_j, \eta_j)$ and use the estimator bound (5.1.18) we obtain, upon taking the necessary suprema, the second line in (5.3.9). $\square$

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

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

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

Proposition 5.3.12

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 $h \in \mathcal{B}_d(\Lambda^* \cup \Lambda^*_-)$,*

$$\delta^d_j\,(f_{\Lambda^*} h) \le \delta^d_j(h) + \sum_{k \in \Lambda^*} \delta^d_k(h) \left[\sum_{l=1}^{n(\Lambda)} (P_{\Lambda^*}\alpha)^l\right]_{kj} . \tag{5.3.13}$$

(ii) *If Dobrushin condition (5.1.20) is satisfied, then for every $j \in \Lambda^*_-$ and $h \in \mathcal{B}_d(\Lambda^*)$,*

$$\delta^d_j\,(f_{\Lambda^*} h) \le \sum_{k \in \Lambda^*} \delta^d_k(h) \left[\frac{P_{\Lambda^*}\alpha}{1 - P_{\Lambda^*}\alpha}\right]_{kj} . \tag{5.3.14}$$

Proof

We only need to prove (5.3.13). Inequality (5.3.14) is then obtained by bounding the sum in the RHS of (5.3.13) 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 (4.2.5) of the LIS, $\delta_j^d\left(f_{\Delta^*} h\right) = \delta_j^d\left(f_{\Lambda^*} f_{V_{n+1}} h\right)$. Therefore, by the inductive hypothesis,

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

and the multisite dusting lemma 5.3.8 yields

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

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 $m \in V_{n+1}$, $l \in \mathbb{N}$,

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

The last two displays imply that

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

Proof of Theorem 5.1.19

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 $m - i \leq n$. 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

$$\Bigl|\nu(h) - \mu(h)\Bigr| \;\leq\; \iint \Bigl| f_{V_{m-i}^n}\left(h \mid d\xi\right) - f_{V_{m-i}^n}\left(h \mid d\eta\right)\Bigr| \;\nu(d\xi)\,\mu(d\eta) .$$

Therefore, by the continuity of f and (5.3.14),

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

Under condition (5.1.20) the series on the RHS is summable, hence the bound converges to zero as $i \to \infty$. $\square$

5.4 Proofs on loss of memory and mixing

Proof of Theorem 5.2.3 and Proposition 5.2.24

Part (i) of Theorem 5.2.3 is just (5.3.13). The triangular property of F implies that for each $i \in \Lambda^*$,

$$\left[(P_{\Lambda^*}\alpha)^2\right]_{kj} e^{F(k,j)} \;=\; \sum_{i\in\Lambda^*} \alpha_{ki}\,\alpha_{ij}\, e^{F(k,j)} \;\le\; \sum_{i\in\Lambda^*} \alpha_{ki}\, e^{F(k,i)}\, \alpha_{ij}\, e^{F(i,j)} \;.$$

Therefore,

$$\left[(P_{\Lambda^*}\alpha)^2\right]_{kj} e^{F(k,j)} \;\le\; \sum_{j\in\mathbb{Z}} \left[(P_{\Lambda^*}\alpha)^2\right]_{kj}\; e^{F(k,j)} \;\le\; \gamma_{\Lambda^*}^2 \;.$$

Proceeding inductively we obtain

$$\left[(P_{\Lambda^*}\alpha)^n\right]_{kj} \;\le\; \gamma_{\Lambda^*}^n\; e^{-F(k,j)} \tag{5.4.1}$$

for every natural n. This yields (5.2.25) upon summation over n. Combining (5.2.25) with (5.3.14), we obtain (5.2.6). $\square$

Proof of Theorem 5.2.15

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 $n \in \mathbb{N}$,

$$\begin{aligned}\Bigl|\mu(h) - \widetilde{\mu}(h)\Bigr| \;\le\; & \Bigl|\mu\left(f_{[m_\Lambda-n,m_\Lambda]}h\right) - \widetilde{\mu}\left(f_{[m_\Lambda-n,m_\Lambda]}h\right)\Bigr| \\ & + \Bigl|\widetilde{\mu}\left(f_{[m_\Lambda-n,m_\Lambda]}h\right) - \widetilde{\mu}\left(\widetilde{f}_{[m_\Lambda-n,m_\Lambda]}h\right)\Bigr| \;.\end{aligned} \tag{5.4.2}$$

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 $\omega \in \Omega$ as $n \to \infty$ (see Remark 5.1.8). Therefore, by dominated convergence (h is continuous, hence bounded)

$$\Bigl|\mu\left(f_{[m_\Lambda-n,m_\Lambda]}h\right) - \widetilde{\mu}\left(f_{[m_\Lambda-n,m_\Lambda]}h\right)\Bigr| \;\xrightarrow[n\to\infty]{}\; 0 \;. \tag{5.4.3}$$

To bound the last term in (5.4.2) we telescope using the factorization property (4.2.5) for LIS:

$$\begin{aligned}\Bigl|\widetilde{\mu}\bigl(f_{[m_\Lambda-n,m_\Lambda]}h\bigr) - \widetilde{\mu}\bigl(\widetilde{f}_{[m_\Lambda-n,m_\Lambda]}h\bigr)\Bigr| \;\le\; & \Bigl|\mu\bigl(f_{\{m_\Lambda\}}h\bigr) - \widetilde{\mu}\bigl(\widetilde{f}_{\{m_\Lambda\}}h\bigr)\Bigr| \\ & + \sum_{k=m_\Lambda-n}^{m_\Lambda-1} \Bigl|\widetilde{\mu}\bigl(f_{[k,m_\Lambda]}h\bigr) - \widetilde{\mu}\bigl(\widetilde{f}_{\{k\}}f_{[k+1,m_\Lambda]}h\bigr)\Bigr| \;.\end{aligned} \tag{5.4.4}$$

The definition (5.1.21)/(5.1.22) of the VKR distance, implies that

$$\Big|(f_k\, g)(\omega) - (\widetilde{f}_k\, g)(\omega)\Big| \;\le\; \delta_k^d(g)\, \Big\| \overset{\circ}{f}_{\{k\}}\,(\,\cdot\mid\omega) - \overset{\circ}{\widetilde{f}}_{\{k\}}\,(\,\cdot\mid\omega)\Big\|_d\,,$$

for all $k \in \mathbb{Z}$, $\omega \in \Omega_{-\infty}^{k-1}$ and $g \in \mathcal{B}_d(]-\infty,k])$. Hypothesis (5.2.16) implies

$$\Big|\widetilde{\mu}\big(f_{\{k\}}\, g - \widetilde{f}_{\{k\}}\, g\big)\Big| \;\le\; \widetilde{\mu}\,(b_k)\, \delta_k^d(g)\,. \tag{5.4.5}$$

Combining (5.4.4) and (5.4.5) we obtain

$$\Big|\widetilde{\mu}\big(f_{[m_\Lambda-n,m_\Lambda]}h\big) - \widetilde{\mu}\big(\widetilde{f}_{[m_\Lambda-n,m_\Lambda]}h\big)\Big| \;=\; \sum_{k=m_\Lambda-n}^{m_\Lambda-1} \widetilde{\mu}\,(b_k)\;\delta_k^d\big(f_{[k+1,m_\Lambda]}\big) \;+\; \widetilde{\mu}\,(b_i)\;\delta_i^d(h)\,. \tag{5.4.6}$$

To obtain (5.2.17) we insert this bound in (5.4.2), let n tend to infinity and use (5.4.3). □

Proof of Theorem 5.2.18

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, we can suppose that $h_2 \ge 0$, $h_2 \not\equiv 0$ and $\mu(h_2) = 1$ since both sides of (5.2.19) 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

$$\mathrm{Cor}_\mu\,(h_1,h_2) \;=\; \Big|\nu(h_1) - \mu(h_1)\Big| \tag{5.4.7}$$

where ν is the probability measure defined by

$$\nu \;=\; h_2\,\mu\,. \tag{5.4.8}$$

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

$$\widetilde{f}_k \;=\; g_k\, f_{\{k\}} \tag{5.4.9}$$

with

$$g_k \;=\; \begin{cases} 1 & \text{if } k \in [m_\Delta+1, m_\Lambda] \\[2mm] \dfrac{f_{[k+1,m_\Delta]}\big(h_2 \bigm| \cdot\big)}{f_{[k,m_\Delta]}\big(h_2 \bigm| \cdot\big)} & \text{if } k \in]-\infty, m_\Delta]\,. \end{cases} \tag{5.4.10}$$

The function g_k is well defined because $f_{[k,m_\Delta]}h_2 \neq 0$ for every $k \in]-\infty,i]$. Indeed the existence of k such that $f_{[k,q]}h_2 = 0$ would imply, by consistency, that $\mu(h_2) = 0$. This contradicts the fact that $\mu(h_2) = 1$. It is clear that the kernels $\widetilde{f}_k$ satisfy the hypotheses of Theorem (4.2.1), 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 $h \in \mathcal{F}_{\le k}$:

$$\nu\big(\widetilde{f}_k(h)\big) \;=\; \mu(h_2\, f_{\{k\}}(h)\big) \;=\; \mu\big(f_{\{k\}}(h_2\, h)\big) \;=\; \mu(h_2\, h) \;=\; \nu(h)\;. \tag{5.4.11}$$

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

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

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

the last inequality being a consequence of the consistency of μ with f. Upon inserting the definition of g_k [second line in (5.4.10)] we see that there is a term $f_{[k,m_\Delta]}$ in the denominator that can be pulled to the left because of its $\mathcal{F}_{\le k-1}$-measurability. This produces a cancellation with an analogous term in the numerator. We thus obtain

$$\begin{aligned}\nu\big(\widetilde{f}_k(h)\big) \;&=\; \mu\Big(f_{\{k\}}\big[h\, f_{[k,m_\Delta]}(h_2)\big]\Big) \;=\; \mu\Big(f_{\{k\}}\big[f_{[k,m_\Delta]}(h_2\, h)\big]\Big)\\ &=\; \mu(h_2\, h) \;=\; \nu(h)\;.\end{aligned} \tag{5.4.12}$$

The third inequality is due to the $\mathcal{F}_{\le k}$-measurability of h and the fourth one to the consistency of μ with f. Identities (5.4.11) and (5.4.12) 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

$$\Big\|\overset{\circ}{f}_k\,(\,\cdot\mid\omega) - \overset{\circ}{\widetilde{f}}_k\,(\,\cdot\mid\omega)\Big\|_d \;\le\; b_k(\omega)\;. \tag{5.4.13}$$

For starters, we can take

$$b_k = 0 \quad \forall\, k \in [m_\Delta + 1, m_\Lambda]\;, \tag{5.4.14}$$

because $\overset{\circ}{f}_k\,(\,\cdot\mid\omega) = \overset{\circ}{\widetilde{f}}_k\,(\,\cdot\mid\omega)$, 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

$$u(x) \;\triangleq\; g_k\left(\omega_{-\infty}^{k-1}\, x\right) \;=\; \frac{f_{[k+1,m_\Delta]}\big(h_2 \mid \omega_{-\infty}^{k-1}\, x\big)}{f_{[k,m_\Delta]}\big(h_2 \mid \omega\big)} \tag{5.4.15}$$

and the measure

$$\alpha \;\triangleq\; \overset{\circ}{f}_k\,(\,\cdot\mid\omega) \tag{5.4.16}$$

on Ω_k^ω. Notice that

$$\tilde{\mathring{f}}_k\,(\cdot\mid\omega) - \mathring{f}_k\,(\cdot\mid\omega) \;=\; u\,\alpha - \alpha\,. \tag{5.4.17}$$

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

$$m_h \;\triangleq\; \sup_{x\neq y} \frac{h(x)+h(y)}{2d(x,y)}\;,$$

and observe that

$$\big\|h - m_h D\big\|_\infty \;\le\; \frac{D}{2}\,\delta_k^d(h) \tag{5.4.18}$$

We affirm that

$$\Big\|\mathring{f}_k\,(\cdot\mid\omega) - \tilde{\mathring{f}}_k\,(\cdot\mid\omega)\Big\|_d \;\le\; \frac{D}{2}\,\alpha(|u-1|)\;. \tag{5.4.19}$$

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

$$\Big|u\,\alpha(h) - \alpha(h)\Big| \;=\; \Big|\alpha\big[(u-1)(h-m_hD)\big]\Big| \;\le\; \alpha\,(|u-1|)\,\big\|h-m_hD\big\|_\infty\;.$$

From this and (5.4.18), assertion (5.4.19) follows.

We now use Schwarz's inequality to bound

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

and since $\alpha(u)$ minimize $x\longmapsto \alpha\left((u-x)^2\right)$, we obtain

$$\alpha(|u-1|) \;=\; \Big[\alpha\Big((u-m_uD)^2\Big)\Big]^{\frac{1}{2}} \;\le\; \|u-m_uD\|_\infty\;. \tag{5.4.20}$$

The combination of (5.4.19) and (5.4.20) gives (5.4.13) with

$$b_k(\omega) \;\triangleq\; \frac{D^2\,\delta_k^d(u)}{4} \;=\; \frac{D^2\,\delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right)}{4\,f_{[k,m_\Delta]}h_2(\omega)}\;. \tag{5.4.21}$$

3rd stage: We estimate $\nu\,(b_k)$. From (5.4.21):

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

By consistency, $\mu = \mu\,f_{[k,m_\Delta]}$, hence the last factor is just 1. From this and (5.4.14) we conclude that

$$\nu(b_k) \;=\; \begin{cases} 0 & \text{if } k\in[m_\Delta+1, m_\Lambda] \\[2ex] \dfrac{D^2}{4}\,\delta_k^d\left(f_{[k+1,m_\Delta]}h_2\right) & \text{if } k\in\Delta\cup\Delta_-\,. \end{cases} \tag{5.4.22}$$

In view of (5.4.7), (5.4.13) and (5.4.22) imply (5.2.19) by Theorem 5.2.15. □

Proof of Corollary 5.2.21

Part (i) follows from (5.3.14) and (5.2.19), and part (ii) from (5.3.13) and (5.2.22). □

5.5 Comparative summary of results

The following tables summarize the different uniqueness and mixing results available.

Speed for Uniqueness

	Space	Transformation	Non-nullness	Speed for Uniqueness
Harris (1955)	finite	shift	weak	$\sum_{n\geq 1}\prod_{k=1}^{n}\left(1-\frac{\lvert E\rvert}{2}\mathrm{var}_k(g)\right)=+\infty$
Keane (1972)	finite	covering	strong	$\exists\, 0<a<1,\ C<+\infty : \mathrm{var}_k(g)\leq C\,a^k$
Iosifescu and Spataru (1973)	countable	shift	weak	$\sum_{n\geq 0}\prod_{k=0}^{n}(1-\alpha_k)=+\infty$
Walters (1975)	finite	subshift	strong	$\sum_{k\geq 0}\mathrm{var}_k(\log g)<+\infty$
Berbee (1987)	countable	shift	strong	$\sum_{n\geq 1}\exp\left(-\sum_{k=1}^{n}\mathrm{var}_k(\log g)\right)=+\infty$
Hulse (1997)	finite	shift	strong	$\sum_{k\leq 0}\delta_k(g)<+\infty$
Stenflo (2002)	finite	shift	strong	$\sum_{n\geq 1}\prod_{k=1}^{n}\Delta_k(g)=+\infty$
Johansson and Öberg (2002)	finite	covering	strong	$\sum_{k\geq 0}\mathrm{var}_k^2(\log g)<+\infty$
Fernández and Maillard (2003)	compact	shift	no	$\exists\, C>0:\forall\,\omega_l^m,\ \exists\, n$ such that $\forall\,\xi,\eta$ $f_{[n,m]}\left(\omega_l^m \mid \xi\right)\geq C\, f_{[n,m]}\left(\omega_l^m \mid \eta\right)$
Fernández and Maillard (2003)	compact	shift	no	$\sum_{k\geq 0}C_k^d(g)<1$

Links between speeds for uniqueness

Variation

$$
\text{Keane} \;\Rightarrow\; \text{Walters} \;\Rightarrow\; \left\{\begin{array}{c} \text{F. M. 1} \\ \left.\begin{array}{c} \text{I. S.} \\ \text{Harris} \\ \lhd\!\!\!\!\triangledown \\ \text{Berbee} \end{array}\right\} \;\Rightarrow\; \text{Stenflo} \end{array}\right.
$$

$$
\begin{array}{ccc} \text{Berbee} & & \\ \lhd\!\!\!\!\triangledown & & \\ \text{J. Ö.} & \Leftarrow & \text{Harris} \end{array}
$$

Oscillation

$$
\text{Keane} \;\Rightarrow \left\{\begin{array}{c} \text{F. M. 2} \\ \Downarrow \\ \text{Hulse} \end{array}\right\} \lhd\,\rhd \;(\text{Harris, Berbee, J. Ö., Stenflo})
$$

Nomenclature $\Rightarrow$: "faster", $\lhd\rhd$: "complementary"

Speed for decay of correlations

	Space	Transformation	Non-nullness	Speed for Decay of Correlations
Bowen (1975)	countable	shift	strong	$\exists\, 0 < a < 1,\ C < +\infty : \mathrm{var}_k(\log g) \le C\, a^k$
Bressaud, Fernández and Galves (1987)	countable	shift	strong	$\sum_{k \ge 0} \mathrm{var}_k(\log g) < +\infty$
Fernández and Maillard (2003)	compact	shift	no	$\sum_{k \le 0} \delta_k(g) < 1$

Links between speeds for correlations

$$\text{Bowen} \;\Rightarrow \left\{ \begin{array}{c} \text{B. F. G.} \\ \vartriangle \\[-4pt] \triangledown \\ \text{F. M.} \end{array} \right\}$$

Nomenclature $\Rightarrow$: "faster", $\triangleleft\,\triangleright$: "complementary"

Chapter 6

Correspondence between LIS and specifications

We discuss the relationship between discrete-time processes (chains) and one-dimensional Gibbs measures. We consider finite-alphabet (finite-spin) systems, possibly with a grammar (exclusion rule). We establish conditions for a stochastic process to define a Gibbs measure and vice versa. Our conditions generalize well known equivalence results between ergodic Markov chains and fields, as well as the known Gibbsian character of processes with exponential continuity rate. Our arguments are purely probabilistic; they are based on the study of regular systems of conditional probabilities (specifications). Furthermore, we discuss the equivalence of uniqueness criteria for chains and fields and we establish bounds for the continuity rates of the respective systems of finite-volume conditional probabilities.

6.1 Markovian case

The equivalence between Markov processes and fields is a well known result. See, for instance, Chapter 11 of the treatise by Georgii (1988). The equivalence, however, is obtained by eigenvalue-eigenvector considerations which are not readily applicable to non-Markovian processes. In this section we give an alternative to this previous work by using only purely probabilistic arguments.

We consider $E = \mathcal{A}$ a finite alphabet and $\mathcal{A}^{\mathbb{Z}}$ the set of all possible configurations ω. Let us denote by Θ_{Mark} the set of all stochastic matrices with positive entries and by Π_{Mark}

be the set of Markovian positive specifications defined as follows

Definition 6.1.1
A specification γ *is a* **Markovian specification** *if for all* $\Lambda \in \mathcal{S}$ *and all* $\omega \in \mathcal{A}^{\mathbb{Z}}$,

$$\gamma_\Lambda \left(\omega_\Lambda \mid \omega_{\Lambda^c}\right) = \gamma_\Lambda \left(\omega_\Lambda \mid \omega_{\partial\Lambda}\right)$$

where $\partial\Lambda \triangleq \{i \in \mathbb{Z} \setminus \Lambda \ : \ |i-j| = 1 \text{ for some } j \ \in \Lambda\}$ *is the boundary of* Λ.

Spitzer (1971) has proven that there exists a one-to-one correspondence between Θ_{Mark} and Π_{Mark}. His proof is based on the theory of positive matrix and in particular uses the Perron-Frobenius theorem. Here we will give a independent proof which do not depends on the latter theory and we will see in next sections that this new approach will permit us to extend the result to the non-markovian case.

Remark 6.1.2
Let $P = (P(x,y))_{x,y\in A}$ be a stochastic matrix on $\mathcal{A}$ with non-vanishing entries. We note by μ^P the unique stationary distribution of the Markov chain with transition matrix P and $\alpha^P \in]0,\ 1[^{\mathcal{A}}$ the unique probability vector satisfying $\alpha^P\, P = \alpha^P$. Therefore, μ^P is uniquely determined by

$$\mu^P \left(\omega_i^j\right) = \alpha^P \left(\omega_i\right)\, P\left(\omega_i, \omega_{i+1}\right) \cdots\, P\left(\omega_{j-1}, \omega_j\right)$$

for all $i, j \in \mathbb{Z} : i < j$ and $\omega_i^j \in \mathcal{A}_i^j$. Since every finite subset Λ of $\mathbb{Z}$ is of the form $\Lambda = \bigcup_{k=1}^n \{i_k+1, ..., i_k+n_k\}$, where $n \geq 1$ and the sets $\{i_k, i_k{+}1, ..., i_k{+}n_k\}$ are pairwise disjoint, we have

$$\mu^P(\omega_\Lambda \mid \sigma_{\partial\Lambda}) = \prod_{k=1}^n \frac{P\left(\sigma_{i_k}, \omega_{i_k+1}\right)\, P\left(\omega_{i_k+1}, \omega_{i_k+2}\right) \cdots\, P\left(\omega_{i_k+n_k}, \sigma_{i_k+n_k+1}\right)}{P^{n_k+1}\left(\sigma_{i_k}, \sigma_{i_k+n_k+1}\right)}. \tag{6.1.3}$$

Let us define the functions

$$b : \Theta_{\text{Mark}} \longrightarrow \Pi_{Mark},\ P \longmapsto \gamma^P$$

such that

$$\gamma_\Lambda^P \left(\omega \mid \sigma\right) = \mu^P \left(\omega_\Lambda \mid \sigma_{\partial\Lambda}\right) \times \delta_{\sigma_{\partial\Lambda}}\left(\omega_{\partial\Lambda}\right) \tag{6.1.4}$$

for all $\Lambda \in \mathcal{S}$, and ω, $\sigma \in \mathcal{A}^{\mathbb{Z}}$ and

$$c : \Pi_{\text{Mark}} \longrightarrow \Theta_{\text{Mark}},\ \gamma \longmapsto P^\gamma$$

such that

$$P^\gamma\left(\omega_0, \omega_{-1}\right) = \lim_{n\to+\infty} \gamma_{[0,n]}\left(\omega_0 \mid \omega\right) \tag{6.1.5}$$

for all ω_0, $\omega_{-1}\ \in \mathcal{A}$.

Theorem 6.1.6
The maps b and c define a one-to-one correspondence between Θ_{Mark} and Π_{Mark} that preserve the consistent measure. More precisely,

(a) $\mathcal{G}\left(\gamma^P\right) = \{\mu^P\}$ *for all* $P \in \Theta$,
where μ^P *is the only probability measure consistent with* P.

(b) $\mathcal{G}\left(P^\gamma\right) = \{\mu^\gamma\}$ *for all* $\gamma \in \Pi$,
where μ^γ *is the only probability measure consistent with* γ.

Similar results have already been given by Spitzer (1971) and can be found in Georgii (1988) (chapter 3).

Proof We will make the proof in four steps.

1) $\gamma^P \in \Pi$ and $\mathcal{G}\left(\gamma^P\right) = \{\mu_P\}$ for all $P \in \Theta$.

2) $P^\gamma \in \Theta$ and $\mathcal{G}\left(P^\gamma\right) = \{\mu^\gamma\}$ for all $\gamma \in \Pi$.

3) $c \circ b(P) = P$, for all $P \in \Theta$.

4) $b \circ c(\gamma) = \gamma$, for all $\gamma \in \Pi$.

To prove **1)**, we first check that γ^P is a Markovian positive specification. By construction, for all $\Lambda \in \mathcal{S}$ and $B \in \mathcal{F}_{\Lambda^c}$, $\gamma^P_\Lambda\left(B \mid \cdot\right) = \mathbb{1}_B(\cdot)$. Let $\Lambda, \Delta \in \mathcal{S}$ such that $\Lambda \subset \Delta$, $\omega_\Delta \in \mathcal{A}^\Delta$ and $\sigma \in \mathcal{A}^{\mathbb{Z}}$. To show that γ^P is a specification it suffices to prove that

$$\gamma^P_\Delta \gamma^P_\Lambda \left(\omega_\Delta \mid \sigma\right) = \gamma^P_\Delta \left(\omega_\Delta \mid \sigma\right). \tag{6.1.7}$$

By definition (6.1.4) of γ^P we have

$$\begin{aligned}\gamma^P_\Delta \gamma^P_\Lambda \left(\omega_\Delta \mid \sigma\right) &= \int \gamma^P_\Lambda \left(\omega_\Lambda \omega_{\Delta\setminus\Lambda} \mid \eta\right) \gamma^P_\Delta \left(d\eta \mid \sigma\right) = \int_{\{\omega_{\Delta\setminus\Lambda}\}} \gamma^P_\Lambda \left(\omega_\Lambda \mid \eta\right) \gamma^P_\Delta \left(d\eta \mid \sigma\right) \\ &= \mu^P \left(\omega_\Lambda \mid \omega_{\partial\Lambda}\right) \mu^P \left(\omega_{\Delta\setminus\Lambda} \mid \sigma_{\partial\Delta}\right). \end{aligned} \tag{6.1.8}$$

Using one of well-known properties of conditional probabilities we have

$$\mu^P \left(\omega_\Lambda \mid \omega_{\partial\Lambda}\right) \mu^P \left(\omega_{\Delta\setminus\Lambda} \mid \sigma_{\partial\Delta}\right) = \mu^P \left(\omega_\Delta \mid \sigma_{\partial\Delta}\right) \quad \mu^P\text{-a.s.}, \tag{6.1.9}$$

but both sides of this last equality are $\mathcal{F}_{\partial\Delta}$-measurable and μ^P charges each atom of $\mathcal{F}_{\partial\Delta}$. Hence (6.1.9) holds everywhere. Combining this with (6.1.8) we obtain

$$\gamma^P_\Delta \gamma^P_\Lambda \left(\omega_\Delta \mid \sigma\right) = \mu^P \left(\omega_\Delta \mid \sigma_{\partial\Delta}\right) = \gamma^P_\Delta \left(\omega_\Delta \mid \sigma\right).$$

So γ^P is a specification and since γ^P is positive Markovian by construction it follows that

$$\gamma^P \in \Pi.$$

Second we show that $\left|\mathcal{G}\left(\gamma^P\right)\right| = \left\{\mu^P\right\}$. Since γ^P is Markovian, it suffices to prove that

$$\mu^P \in \mathcal{G}\left(\gamma^P\right). \tag{6.1.10}$$

Let $\Lambda \in \mathcal{S}$, $\omega_\Lambda \in \mathcal{A}^\Lambda$ and B any cylinder of $\mathcal{F}_{\Lambda^c}$. Then we have

$$\mu^P\left(\omega_\Lambda \cap B\right) = \mu^P\left(\omega_\Lambda \mid B\right)\mu^P(B) = \int_B \mu^P\left(\omega_\Lambda \mid \sigma\right)\mu^P(d\sigma).$$

But there exists $\Delta \in \mathcal{S}$ such that $\Lambda \cup \partial\Lambda \subset \Delta$ and $B \subset \Delta \setminus \Lambda$ so it follows that

$$\mu^P\left(\omega_\Lambda \cap B\right) = \int_B \mu^P\left(\omega_\Lambda \mid \sigma_{\Delta\setminus\Lambda}\right)\mu^P(d\sigma).$$

Since μ^P is Markovian and $\partial\Lambda \subset \Delta \setminus \Lambda$, we obtain

$$\mu^P\left(\omega_\Lambda \cap B\right) = \int_B \mu^P\left(\omega_\Lambda \mid \sigma_{\partial\Lambda}\right)\mu^P(d\sigma) \triangleq \int_B \gamma^P_\Lambda\left(\omega_\Lambda \mid \sigma\right)\mu^P(d\sigma).$$

Moreover this last equality holds for every $B \in \mathcal{F}_{\Lambda^c}$ and every $\omega_\Lambda \in \mathcal{A}^\Lambda$, so

$$\mathcal{G}\left(\gamma^P\right) = \left\{\mu^P\right\}.$$

To prove **2)**, let $\gamma \in \Pi$. Since γ is a Markovian specification, γ satisfies Dobrushin's uniqueness condition, so there exists a unique measure μ^γ which is consistent with γ and verifies

$$\lim_{n\to+\infty} \gamma_{[0,n]}\left(\omega_0 \mid \sigma\right) = \mu^\gamma\left(\omega_0 \mid \sigma_{-1}\right) \tag{6.1.11}$$

for every $\omega_0,\ \sigma_{-1} \in \mathcal{A}$. So we can define

$$P^\gamma\left(\omega_{-1}, \omega_0\right) = \mu^\gamma\left(\omega_0 \mid \omega_{-1}\right) \tag{6.1.12}$$

for every $\omega_0,\ \omega_{-1} \in \mathcal{A}$. First we show that P^γ has positive entries. Since γ is a positive Markovian specification, let $\omega_0 \in \mathcal{A}$ and $n > 0$, it follows that $\gamma_{[0,n]}\left(\omega_0 \mid \omega_{-1}\ \omega_{n+1}\right) > 0$ for every $\omega_{-1},\ \omega_{n+1} \in \mathcal{A}$, so there exists a positive constant

$$C(\omega_0, n) \triangleq \inf_{\omega_{-1},\,\omega_{n+1}} \gamma_{[0,n]}\left(\omega_0 \mid \omega_{-1}\ \omega_{n+1}\right).$$

By consistency of γ we have for every $m \geq n$,

$$\gamma_{[0,m]}(\omega_0 \mid \omega) = \gamma_{[0,m]}\gamma_{[0,n]}(\omega_0 \mid \omega) \geq C(\omega_0, n),$$

and taking the limit when m tend to infinity in the last inequality, we obtain

$$P^{\gamma}\left(\omega_{-1}, \omega_{0}\right) \geq C(\omega_{0}, n) > 0. \tag{6.1.13}$$

Hence P^{γ} has positive entries and P^{γ} belongs to Θ. Second to prove that $\mathcal{G}\left(P^{\gamma}\right) = \{\mu^{\gamma}\}$. Since P^{γ} has positive entries, it suffices to show that μ^{γ} is consistent with P^{γ} which is trivial in view of (6.1.12).

We proceed with **3)**. Let $P \in \Theta$, then by (6.1.3)

$$\begin{aligned}\gamma^{P}_{[0,n]}\left(\omega_0 \mid \omega_{-1}\,\omega_{n+1}\right) &= \sum_{\omega_1^n} \gamma^{P}_{[0,n]}\left(\omega_0^n \mid \omega_{-1}\,\omega_{n+1}\right) \triangleq \sum_{\omega_1^n} \mu^{P}\left(\omega_0^n \mid \omega_{-1}\,\omega_{n+1}\right)\\ &= \frac{P(\omega_{-1},\omega_0)P^{n+1}(\omega_0,\omega_{n+1})}{P^{n+2}(\omega_{-1},\omega_{n+1})}.\end{aligned}$$

From the ergodic theorem for Markov chains, we have

$$P^{n}(\,\cdot, x) \xrightarrow{n \to +\infty} \nu^{P}(x), \quad \forall\, x \in A.$$

Therefore

$$\gamma^{P}_{[0,n]}\left(\omega_0 \mid \omega_{-1}\,\omega_{n+1}\right) \xrightarrow{n \to +\infty} P(\omega_0, \omega_1),$$

that is

$$c \circ b(P) = P, \quad \text{for all } P \in \Theta.$$

To prove **4)**, let $\gamma \in \Pi$. Then by (6.1.3)

$$\begin{aligned}\gamma^{P^{\gamma}}_{\Lambda}\left(\omega_{\Lambda} \mid \sigma\right) &\triangleq \mu^{P^{\gamma}}\left(\omega_{\Lambda} \mid \sigma_{\partial\Lambda}\right)\\ &= \prod_{k=1}^{n} \frac{P^{\gamma}\left(\sigma_{i_k}, \omega_{i_k+1}\right)\, P^{\gamma}\left(\omega_{i_k+1}, \omega_{i_k+2}\right) \cdots P^{\gamma}\left(\omega_{i_k+n_k}, \sigma_{i_k+n_k+1}\right)}{\displaystyle\sum_{\sigma_{i_k+1}^{i_k+n_k}} \prod_{l=i_k}^{i_k+n_k} P^{\gamma}\left(\sigma_l, \sigma_{l+1}\right)}.\end{aligned}$$

By (6.1.12), this gives

$$\begin{aligned}\gamma^{P^{\gamma}}_{\Lambda}\left(\omega_{\Lambda} \mid \sigma\right) &= \prod_{k=1}^{n} \frac{\mu^{\gamma}\left(\omega_{i_k+1} \mid \sigma_{i_k}\right) \mu^{\gamma}\left(\omega_{i_k+2} \mid \omega_{i_k+1}\right) \cdots \mu^{\gamma}\left(\sigma_{i_k+n_k+1} \mid \omega_{i_k+n_k}\right)}{\displaystyle\sum_{\sigma_{i_k+1}^{i_k+n_k}} \prod_{l=i_k}^{i_k+n_k} \mu^{\gamma}\left(\sigma_{l+1} \mid \sigma_l\right)}\\ &= \mu^{\gamma}\left(\omega_{\Lambda} \mid \sigma_{\partial\Lambda}\right).\end{aligned} \tag{6.1.14}$$

But γ is a Markovian specification and μ^{γ} charges each atom of $\mathcal{F}_{\partial\Lambda}$ therefore

$$\mu^{\gamma}\left(\omega_{\Lambda} \mid \sigma_{\partial\Lambda}\right) = \gamma_{\Lambda}\left(\omega_{\Lambda} \mid \sigma_{\partial\Lambda}\right) = \gamma_{\Lambda}\left(\omega_{\Lambda} \mid \sigma\right). \tag{6.1.15}$$

This combining with (6.1.14) assure that $\gamma^{P^\gamma} = \gamma$, that is

$$b \circ c(\gamma) = \gamma, \quad \text{for all } \gamma \in \Pi.$$

In conclusion of these different steps, b is one-to-one and $b^{-1} = c$. Moreover for all $P \in \Theta$ and all $\gamma \in \Pi$, we have $\mathcal{G}\left(\gamma^P\right) = \{\mu_P\}$ and $\mathcal{G}\left(P_\gamma\right) = \{\mu_\gamma\}$. $\square$

In the sequel, we will generalize this correspondence in a non-Markovian setting. We directly establish consistency-preserving maps between specifications and left-interval specifications. We consider systems with a finite alphabet, possibly with a grammar, that is, with exclusion rules such that the non-excluded configurations form a compact set. We do not assume translation invariance either of the kernels or of the consistent measures. The main limitation of our results is that, in order to insure that the necessary limits are uniquely defined, specifications and processes are required to satisfy a strong uniqueness condition called *hereditary uniqueness condition* (HUC). A second property, called *good future* (GF) is demanded for stochastic processes to guarantee some control of the conditioning with respect to the future. For shift-invariant transition kernels this second condition is stronger than the first one. HUC is verified, for instance, by specifications satisfying Dobrushin and boundary-uniformity criteria (reviewed in Theorem 6.2.9 below). Both GF and HUC are satisfied by a large family of processes, for instance by the chains with summable variations studied by Harris (1955), Ledrappier (1974), Walters (1975),Lalley (1986), Berbee (1987), Bressaud et al (1999), ...

Our results show that under these hypotheses there exist: (i) a map that to each LIS associates a specification such that the process consistent with the former is a Gibbs measure consistent with the latter (Theorem 6.3.12), and (ii) a map that to each specification associates a LIS such that the Gibbs measure consistent with the former is a process consistent with the latter (Theorem 6.3.17). If domain and image match, these maps are inverses of each other. This happens, in particular, in the case of exponentially decreasing continuity rates (Theorem 6.3.20). As part of the proofs, we obtain estimates linking the continuity rates of LIS and specifications related by these maps (Theorem 6.3.19). We also show that the validity of the Dobrushin and boundary-uniformity criteria for the specification implies the validity of analogous criteria for the associated stochastic process (Theorem 6.3.18).

6.2 Preliminary results

Here we consider a finite alphabet $\mathcal{A}$ endowed with the discrete topology and σ-algebra, and Ω a compact subset of $\mathcal{A}^{\mathbb{Z}}$. The space Ω is endowed with the projection $\mathcal{F}$ of the product

σ-algebra associated to $\mathcal{A}^{\mathbb{Z}}$. The space Ω represents admissible "letter configurations", where the admissibility is defined, for instance by some exclusion rule as in Ruelle (1978) or by a "grammar" (subshift of finite type) as in Walters (1975).

Let us recall some notions needed in the sequel.

Definition 6.2.1

1) *A specification γ satisfies a* **hereditary uniqueness condition** *(HUC)* **for a family** *$\mathcal{H}$ of subsets of $\mathbb{Z}$ if for all (possibly infinite) sets $V \in \mathcal{H}$ and all configurations $\omega \in \Omega$, the specification $\gamma^{(V,\omega)}$ defined by*

$$\gamma_\Lambda^{(V,\omega)}(\,\cdot \mid \xi) = \gamma_\Lambda(\,\cdot \mid \omega_{V^c}\,\xi_V)\ , \quad \forall\,\Lambda \in \mathcal{S},\ \Lambda \subset V,\ \ \forall\,\omega_{V^c}\,\xi_V \in \Omega\ , \tag{6.2.2}$$

admits a unique Gibbs measure. The specification satisfies a HUC if it satisfies a HUC for $\mathcal{H} = \mathcal{P}(\mathbb{Z})$.

2) *A LIS f satisfies a* **hereditary uniqueness condition** *(HUC) if for all intervals of the form $V = [i, +\infty[$, $i \in \mathbb{Z}$, or $V = \mathbb{Z}$, and all configurations $\omega \in \Omega$, the LIS $f^{(V,\omega)}$ defined by*

$$f_\Lambda^{(V,\omega)}(\,\cdot \mid \xi) = f_\Lambda(\,\cdot \mid \omega_{V_-}\,\xi_V)\ , \quad \forall\,\Lambda \in \mathcal{S}_b,\ \Lambda \subset V,\ \forall\ \omega_{V^c}\,\xi_V \in \Omega\ , \tag{6.2.3}$$

admits a unique consistent chain.

We list now several, mostly well known, sufficient conditions for hereditary uniqueness. They refer to different ways to bound continuity rates of transition kernels. We start with the relevant definitions.

Definition 6.2.4

(i) *The k-***variation** *of a $\mathcal{F}_{\le i}$-measurable function f_i is defined by*

$$\mathrm{var}_k(f_i)\ \triangleq\ \sup\Big\{\big|f_i(\omega_{-\infty}^i) - f_i(\sigma_{-\infty}^i)\big| : \omega_{-\infty}^i, \sigma_{-\infty}^i \in \Omega_{-\infty}^i\,,\ \omega_{i-k}^i = \sigma_{i-k}^i\Big\}\ .$$

(ii) *The* **oscillation** *of a function $h : \Omega \to \mathbb{R}$, with respect to the site $j \in \mathbb{Z}$, is defined by*

$$\delta_j(h)\ \triangleq\ \sup\Big\{|h(\omega) - h(\sigma)| : \omega, \sigma \in \Omega\,,\ \omega \overset{\neq j}{=} \sigma\Big\}\ . \tag{6.2.5}$$

(iii) *The* **interdependence coefficients** *for a family of probability kernels $\pi = \big(\pi_{\{i\}}\big)_{i\in\mathbb{Z}}$, $\pi_{\{i\}} : \mathcal{F}_i \times \Omega \to [0,1]$ are defined by*

$$C_{ij}(\pi)\ \triangleq\ \sup_{\substack{\xi,\eta\in\Omega \\ \xi \overset{\neq j}{=} \eta}} \big\|\pi_{\{i\}}(\,\cdot \mid \xi) - \pi_{\{i\}}(\,\cdot \mid \eta)\big\| \tag{6.2.6}$$

for all $i, j \in \mathbb{Z}$. Here we use the variation norm.

Theorem 6.2.7
A LIS f on $(\Omega, \mathcal{F})$ satisfies a HUC if it satisfies one of the following conditions:

1) **Harris** *(1955): For each $i \in \mathbb{Z}$, $\sum_{n\geq 1}\prod_{k=1}^{n}\left(1-\frac{|\mathcal{A}|}{2}\mathrm{var}_k(f_i)\right) = +\infty$ and f is weakly non-null on Ω.*

2) **Hulse** *(1997): For each $i \in \mathbb{Z}$, $\sum_{k\geq 0}\delta_{i-k}\left(f_i\right) < +\infty$ and f is continuous and non-null on Ω.*

3) **One-sided Dobrushin**: *For each $i \in \mathbb{Z}$, $\sum_{j<i} C_{ij}(f) < 1$ and f is continuous on Ω.*

4) **One-sided boundary-uniformity**: *There exists a constant $K > 0$ so that for every cylinder set $A = \{x_l^m\} \in \Omega_l^m$ there exists an integer n such that*

$$f_{[n,m]}(A \mid \xi) \geq K\, f_{[n,m]}(A \mid \eta) \quad \textit{for all } \xi, \eta \in \Omega\ . \tag{6.2.8}$$

These results which have been proven in chapter 5 (Corollary 5.1.28) are in fact adaptations of the following well known criteria for specifications.

Theorem 6.2.9
A specification γ on $(\Omega, \mathcal{F})$ satisfies a HUC if it satisfies one of the following conditions:

1) **Dobrushin** *(1968), Lanford (1973): $\sup_{i\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} C_{ij}(\gamma) < 1$ and γ continuous.*

2) **Georgii** *(1974)* **boundary-uniformity**: *There exists a constant $K > 0$ so that for every cylinder set $A \in \mathcal{F}$ there exists $\Lambda \in \mathcal{S}_b$ such that*

$$\gamma_\Lambda(A \mid \xi) \geq K\, \gamma_\Lambda(A \mid \eta) \quad \textit{for all } \xi, \eta \in \Omega\ .$$

We remark that the conditions involve no non-nullness assumption.

6.3 Main results

For a LIS f on Ω let us denote, for each $\Lambda \in \mathcal{S},\ k \geq m_\Lambda$ and $\omega \in \Omega_{-\infty}^k$,

$$c_\Lambda^\omega(f_k) \ \triangleq\ \min_{\sigma_\Lambda \in \Omega_\Lambda}\left\{f_k\left(\sigma_\Lambda\, \omega_{]-\infty,k]\setminus\Lambda}\right) : \sigma_\Lambda\, \omega_{]-\infty,k]\setminus\Lambda} \in \Omega_{-\infty}^k\right\} \tag{6.3.1}$$

and

$$\delta_\Lambda^\omega(f_k) \ \triangleq\ \sum_{j\in\Lambda}\sup\left\{\left|f_k\left(\omega_{-\infty}^k\right) - f_k\left(\sigma_{-\infty}^k\right)\right| : \sigma \in \Omega_{-\infty}^k,\ \sigma \overset{\neq j}{=} \omega\right\}. \tag{6.3.2}$$

Similarly, for a specification γ on Ω, let us denote, for each $\omega \in \Omega$ and each $k, j \in \mathbb{Z}$

$$c_j^\omega(\gamma_k) \triangleq \min_{\sigma_j \in \Omega_{\{j\}}} \left\{ \gamma_k \left(\sigma_j \, \omega_{\{j\}^c} \right) : \sigma_j \, \omega_{\{j\}^c} \in \Omega \right\} \tag{6.3.3}$$

and

$$\delta_j^\omega \left(\gamma_k \right) \triangleq \sup \left\{ |\gamma_k(\omega) - \gamma_k(\sigma)| : \sigma \in \Omega, \ \sigma \overset{\neq j}{=} \omega \right\} \ . \tag{6.3.4}$$

Definition 6.3.5

(i) *A LIS f on Ω is said to have a* **good future (GF)** *if it is non-null on Ω and for each $\Lambda \in \mathcal{S}$, there exists a sequence $\left\{\varepsilon_k^\Lambda\right\}_{k\in\mathbb{N}}$ of positive numbers such that $\sum_k \varepsilon_k^\Lambda < +\infty$ for which*

$$\sup_{\omega \in \Omega_{-\infty}^k} c_\Lambda^\omega(f_k)^{-1} \, \delta_\Lambda^\omega(f_k) \ \leq \ \varepsilon_k^\Lambda \tag{6.3.6}$$

for each $k \geq m_\Lambda$.

(ii) *A LIS f on Ω is said to have an* **exponentially-good future (EGF)** *if it is non-null on Ω and there exists a real $a > 1$ such that*

$$\limsup_{k\to\infty} a^{|k-j|} \sup_{\omega \in \Omega_{-\infty}^k} c_j^\omega(f_k)^{-1} \, \delta_j^\omega(f_k) \ < \ \infty \tag{6.3.7}$$

for all $j \in \mathbb{Z}$.

(iii) *A specification γ on Ω is said to have an* **exponentially-good future (EGF)** *if it is non-null on Ω and there exists a real $a > 1$ such that*

$$\limsup_{k\to\infty} a^{|k-j|} \sup_{\omega \in \Omega} c_j^\omega(\gamma_k)^{-1} \, \delta_j^\omega \left(\gamma_k \right) \ < \ \infty \tag{6.3.8}$$

for all $j \in \mathbb{Z}$.

Definition 6.3.9

Let us introduce the following sets.

$\Theta \triangleq \{$*LIS f continuous and non-null on* $\Omega\}$

$\Pi \triangleq \{$*specifications γ continuous and non-null on* $\Omega\}$,

$\Theta_1 \triangleq \{f \in \Theta : f$ *has a GF*$\}$,

$\Pi_1 \triangleq \{\gamma \in \Pi : |\mathcal{G}(\gamma)| = 1\}$,

$\Pi_2 \triangleq \{\gamma \in \Pi : \gamma$ *satisfies a HUC over all* $[i, +\infty[, \ i \in \mathbb{Z}\}$,

$\Theta_2 \triangleq \{f \in \Theta : f$ *satisfies a HUC*$\}$,

$\Theta_3 \triangleq \{f \in \Theta : f$ *has an EGF*$\}$,

$\Pi_3 \triangleq \{\gamma \in \Pi : \gamma$ *has an EGF*$\}$.

We remark that each of the LIS or specifications of any of the preceding sets has at least one consistent measure. This is because the (interesting part of) the configuration space is compact and the LIS or specifications are assumed to be continuous. Indeed, as the space of probability measures on a compact space is weakly compact, every sequence of measures $\gamma_{\Lambda_n}(\,\cdot \mid \omega^{\{n\}})$ or $f_{\Lambda_n}(\,\cdot \mid \omega^{\{n\}})$, for (Λ_n) an exhausting sequence of regions and $(\omega^{\{n\}})$ a sequence of configurations, has a weakly convergent subsequence. By continuity of the transitions the limit is respectively a Gibbs measure or a consistent chain.

Consider the function

$$F_{\Lambda,n}\left(\omega_\Lambda \mid \omega\right) \;\triangleq\; \frac{f_{[l_\Lambda,n]}\left(\omega_{l_\Lambda}^{n} \mid \omega_{\Lambda_-}\right)}{f_{[l_\Lambda,n]}\left(\omega_{\Lambda^c\cap[l_\Lambda,n]} \mid \omega_{\Lambda_-}\right)} \tag{6.3.10}$$

for all $\Lambda \in \mathcal{S}$, $n > m_\Lambda$ and $\omega \in \Omega$. The continuity of f implies that the functions $F_{\Lambda,n}(\omega_\Lambda \mid \cdot\,)$ are continuous on Ω_{Λ^c} for each $\omega_\Lambda \in \Omega_\Lambda$. We use these functions to introduce the map

$$b : \Theta_1 \longrightarrow \Pi\ ,\ f \mapsto \gamma^f$$

defined by

$$\gamma_\Lambda^f\left(\omega_\Lambda \mid \omega\right) \;\triangleq\; \lim_{n\to+\infty} F_{\Lambda,n}(\omega_\Lambda \mid \omega) \tag{6.3.11}$$

for all $\Lambda \in \mathcal{S}$ and $\omega \in \Omega$.

Theorem 6.3.12 (LIS $\rightsquigarrow$ specification)

1) *The map b is well defined. That is, for $f \in \Theta_1$*

- **(a)** *the limit (6.3.11) exists for all $\Lambda \in \mathcal{S}$ and $\omega \in \Omega$.*
- **(b)** *γ^f is a specification on $(\Omega, \mathcal{F})$.*
- **(c)** *$\gamma^f \in \Pi$.*

2)
- **(a)** *For each (finite or infinite) interval V and each $\omega \in \Omega$, $\mathcal{G}\left(f^{(V,\omega)}\right) \subset \mathcal{G}\left(\left(\gamma^f\right)^{(V,\omega)}\right)$.*
- **(b)** *For $f \in b^{-1}(\Pi_1)$, $\mathcal{G}(f) = \mathcal{G}(\gamma^f) = \left\{\mu^f\right\}$, where μ^f is the only chain consistent with f.*
- **(c)** *The map b restricted to $b^{-1}(\Pi_1)$ is one-to-one.*

Remarks

6.3.13 The set of stationary non-null LIS satisfying Hulse condition is strictly included in Θ_1.

6.3.14 Since for all $k \geq 1$, $\mathrm{var}_k(g) \geq \delta_k(g)$, Θ_1 includes the set of stationary non-null LIS with summable variation.

Consider now the map

$$c : \Pi_2 \to \Theta_2 \ , \ \gamma \mapsto f^\gamma \tag{6.3.15}$$

defined by

$$f^\gamma_\Lambda \left(A \mid \omega_{\Lambda_-} \right) \ \triangleq \ \lim_{k \to +\infty} \gamma_{\Lambda \cup \Lambda_+^{(k)}} \left(A \mid \omega \right) \tag{6.3.16}$$

for all $\Lambda \in \mathcal{S}_b,\ A \in \mathcal{F}_\Lambda$, and $\omega \in \Omega$ for which the limit exists.

Theorem 6.3.17 (specification $\rightsquigarrow$ LIS)

1) *The map c is well defined. That is, for $\gamma \in \Pi_2$*

- **(a)** *the limit (6.3.16) exists for all $\Lambda \in \mathcal{S}_b,\ A \in \mathcal{F}_\Lambda,\ \omega_{\Lambda_-} \in \Omega_{\Lambda_-}$ and is independent of ω_{Λ_+}.*
- **(b)** *f^γ is a LIS on $(\Omega, \mathcal{F})$.*
- **(c)** *$f^\gamma \in \Theta_2$.*

2)
- **(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.*

In addition a LIS of the form f^γ satisfies the following properties.

Theorem 6.3.18
Let $\gamma \in \Pi_2$.

(a) *If γ satisfies Dobrushin uniqueness condition, then so does f^γ.*

(b) *If γ satisfies the boundary-uniformity uniqueness condition, then so does f^γ.*

Theorem 6.3.19 (Continuity rates)
Let $\omega \in \Omega$ and $j \in \mathbb{Z}$.

1) *For $f \in \Theta_1$ and $\Lambda \in \mathcal{S}$*

- **(a)** *if $j > m_\Lambda$ then $\delta^\omega_j \left(\gamma^f_\Lambda \right) \leq 2 \sum_{i \geq j} c^\omega_j(f_i)^{-1}\, \delta^\omega_j(f_i)$.*
- **(b)** *if $j < l_\Lambda$ then $\delta^\omega_j(\gamma^f_\Lambda) \leq 1 - \prod_{i=l_\Lambda}^{+\infty} \frac{1 - c^\omega_j(f_i)^{-1}\, \delta^\omega_j(f_i)}{1 + c^\omega_j(f_i)^{-1}\, \delta^\omega_j(f_i)}$.*

2) *For $\gamma \in \Pi_2$, $\Lambda \in \mathcal{S}_b$ and $j < l_\Lambda$, $\delta^\omega_j(f^\gamma_\Lambda) \leq 1 - \prod_{i=l_\Lambda}^{+\infty} \frac{1 - c^\omega_j(\gamma_k)^{-1}\, \delta^\omega_j(\gamma_i)}{1 + c^\omega_j(\gamma_k)^{-1}\, \delta^\omega_j(\gamma_i)}$.*

Under suitable conditions the maps b and c are reciprocal.

Theorem 6.3.20 (LIS $\leftrightsquigarrow$ specification)

(a) $b \circ c = \mathrm{Id}$ *over* $c^{-1}(\Theta_1)$ *and* $\mathcal{G}(f^\gamma) = \mathcal{G}(\gamma) = \{\mu^\gamma\}$.

(b) $c \circ b = \mathrm{Id}$ *over* $b^{-1}(\Pi_2)$ *and* $\mathcal{G}(\gamma^f) = \mathcal{G}(f) = \{\mu^f\}$.

(c) *b and c establish a one-to-one correspondence between* Θ_3 *and* Π_3 *that preserves the consistent measure.*

We remark that Θ_3 includes the well studied processes with Holdërian transition rates (see, for instance, Lalley, 1986, or Keller, 1998). Part (c) of the theorem shows, in particular, the equivalence between such processes and Bowen's Gibbs measures.

6.4 Proofs

We start with a collection of results used for several proofs.

Lemma 6.4.1
Consider $\Lambda \in \mathcal{S}$, $\Delta \subset \Lambda^c$ *and* π *a probability kernel over* $(\mathcal{F}_\Lambda \otimes \mathcal{F}_\Delta\,,\, \Omega_\Lambda \times \Omega_\Delta)$ *such that* $\pi\,(A_\Delta \mid \cdot\,) = \mathbb{1}_{A_\Delta}(\,\cdot\,)$, $\forall\, A_\Delta \in \mathcal{F}_\Delta$. *Then, for all* $\omega \in \Omega_\Lambda \times \Omega_\Delta$,

$$\pi(\,\cdot \mid \omega) \;=\; [\mathring{\pi}_\Lambda\,(\,\cdot \mid \omega) \otimes \delta_{\omega_\Delta}](\,\cdot\,)$$

where $\mathring{\pi}_\Lambda\,(\,\cdot \mid \omega)$ *is the restriction of* $\pi(\,\cdot \mid \omega)$ *to* $\mathcal{F}_\Lambda$ *and* δ_{ω_Δ} *is the Dirac mass at* ω_Δ.

Proof If $A = A_\Lambda \times A_\Delta \in \mathcal{F}_\Lambda \otimes \mathcal{F}_\Delta$ and $\omega \in \Omega_\Lambda \times \Omega_\Delta$

$$\pi\left(A_\Delta \,\middle|\, \omega\right) \;=\; \pi\left(A_\Lambda \times A_\Delta \,\middle|\, \omega\right) + \pi\left(A_\Lambda^c \times A_\Delta \,\middle|\, \omega\right)$$

and

$$\pi\left(A_\Lambda \,\middle|\, \omega\right) \;=\; \pi\left(A_\Lambda \times A_\Delta \,\middle|\, \omega\right) + \pi\left(A_\Lambda \times A_\Delta^c \,\middle|\, \omega\right)\;.$$

Hence,

$$\pi\,(A_\Lambda \times A_\Delta) \;\le\; \pi\,(A_\Lambda) \,\wedge\, \pi\,(A_\Delta) \;\le\; \pi\,(A_\Lambda)\; \mathbb{1}_{A_\Delta}.$$

Analogously,

$$\pi\,(A_\Lambda \times A_\Delta^c) \;\le\; \pi\,(A_\Lambda)\; \mathbb{1}_{A_\Delta^c}\;.$$

On the other hand,

$$\pi\,(A_\Lambda \times A_\Delta) + \pi\,(A_\Lambda \times A_\Delta^c) \;=\; \pi\,(A_\Lambda)\; \mathbb{1}_{A_\Delta} + \pi\,(A_\Lambda)\; \mathbb{1}_{A_\Delta^c}\;.$$

The last three displays imply

$$\pi(A) = \pi\,(A_\Lambda)\;\; \mathbb{1}_{A_\Delta} \quad\Longleftrightarrow\quad \pi\,(A_\Delta) = \;\mathbb{1}_{A_\Delta}\;.\;\square$$

In particular, LIS and specifications are completely defined by the families of their restrictions.

Proposition 6.4.2
Let $\omega \in \Omega$, $\Lambda \in \mathcal{S}$ and $n \in \mathbb{Z}$, $n \geq m_\Lambda$.

(a) *For any $\beta \in \Omega$,*

$$f_{n+1}\left(\omega_{n+1} \mid \omega_{]-\infty,n]}\right) \ \substack{\leq \\ \geq}\ f_{n+1}\left(\omega_{n+1} \mid \beta_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right) \left[1 \pm \frac{\delta_\Lambda^\omega(f_{n+1})}{c_\Lambda^\omega(f_{n+1})}\right] . \tag{6.4.3}$$

(b) *For $j < l_\Lambda$ and $\sigma \in \Omega$ with $\sigma \overset{\neq j}{=} \omega$*

$$F_{\Lambda,n}(\omega_\Lambda \mid \sigma) - F_{\Lambda,n}(\omega_\Lambda \mid \omega) \ \substack{\leq \\ \geq} \ \pm \left[1 - \prod_{i=l_\Lambda}^{n} \frac{1 - c_j^\omega(f_i)^{-1}\, \delta_j^\omega(f_i)}{1 + c_j^\omega(f_i)^{-1}\, \delta_j^\omega(f_i)}\right] \times \begin{cases} F_{\Lambda,n}(\omega_\Lambda \mid \sigma) \\ F_{\Lambda,n}(\omega_\Lambda \mid \omega) \end{cases} . \tag{6.4.4}$$

(c)

$$F_{\Lambda,n+1}(\omega_\Lambda \mid \omega) \ \substack{\leq \\ \geq}\ F_{\Lambda,n}(\omega_\Lambda \mid \omega) \left[1 \pm \frac{\delta_\Lambda^\omega(f_{n+1})}{c_\Lambda^\omega(f_{n+1})}\right] \tag{6.4.5}$$

Proof If we telescope $f_{n+1}\left(\omega_{n+1} \mid \xi_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right) - f_{n+1}\left(\omega_{n+1} \mid \eta_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right)$, transforming ξ_Λ into η_Λ letter by letter, we have

$$f_{n+1}\left(\omega_{n+1} \mid \xi_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right) - f_{n+1}\left(\omega_{n+1} \mid \eta_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right) \ \leq\ \delta_\Lambda^\omega(f_{n+1})$$

for any $\xi_\Lambda, \eta_\Lambda \in \Omega_\Lambda$. To obtain **(6.4.3)** we simply use this inequality twice, assigning ω_Λ to ξ_Λ, β_Λ to η_Λ, and vice versa.

To prove **(6.4.4)** we use definition (6.3.10)

$$F_{\Lambda,n}(\omega_\Lambda \mid \omega) = \frac{f_{[l_\Lambda,n]}\left(\omega_{l_\Lambda}^n \mid \omega_{\Lambda_-}\right)}{\sum_{\beta_\Lambda} f_{[l_\Lambda,n]}\left(\beta_\Lambda\, \omega_{\Lambda^c \cap [l_\Lambda,n]} \mid \omega_{\Lambda_-}\right)}$$

and the factorization

$$f_{[k,n]}\left(\omega_k^n \mid \omega_{-\infty}^{k-1}\right) = \prod_{i=k}^{n} f_i\left(\omega_i \mid \omega_{-\infty}^{i-1}\right) .$$

We then apply inequalities (6.4.3) to bound each of the factors by similar factors with conditioning configuration σ.

To obtain **(6.4.5)** we apply the LIS-reconstruction formula (4.2.5) with $m = n+1$ which yields

$$F_{\Lambda,n+1}(\omega_\Lambda \mid \omega) = \frac{f_{[l_\Lambda,n]}\left(\omega_{l_\Lambda}^n \mid \omega_{\Lambda_-}\right)\, f_{n+1}\left(\omega_{n+1} \mid \omega_{]-\infty,n]}\right)}{\sum_{\beta_\Lambda}\left[f_{[l_\Lambda,n]}\left(\beta_\Lambda\, \omega_{\Lambda^c \cap [l_\Lambda,n]} \mid \omega_{\Lambda_-}\right)\, f_{n+1}\left(\omega_{n+1} \mid \beta_\Lambda\, \omega_{]-\infty,n]\setminus\Lambda}\right)\right]} .$$

In the denominator, only β_Λ with $f_{n+1}\left(\omega_{n+1} \mid \omega_{]-\infty,n]\setminus\Lambda}\,\beta_\Lambda\right) \neq 0$ contribute. We use inequalities (6.4.3) for these. □

Lemma 6.4.6

Let $f \in \Theta_1$, $\omega \in \Omega$ and $\Lambda \in \mathcal{S}$. Consider the sequence $F_{\Lambda,n}\left(\omega_\Lambda \mid \omega\right)$, $n \geq m_\Lambda$ defined by (6.3.10). Then the limit (6.3.11) exits and satisfies

$$\left|\gamma_\Lambda^f\left(\omega_\Lambda \mid \omega_{\Lambda^c}\right) - F_{\Lambda,n}\left(\omega_\Lambda \mid \omega\right)\right| \leq \sum_{k\geq n+1} c_\Lambda^\omega(f_k)^{-1}\,\delta_\Lambda^\omega(f_k). \tag{6.4.7}$$

Proof From (6.4.5), plus the fact that $F_{\Lambda,n} \in [0,1]\;\forall\, n \geq m$, we obtain

$$\left|F_{\Lambda,n+1}\left(\omega_\Lambda \mid \omega\right) - F_{\Lambda,n}\left(\omega_\Lambda \mid \omega\right)\right| \leq c_\Lambda^\omega(f_{n+1})^{-1}\,\delta_\Lambda^\omega(f_{n+1}).$$

Therefore the summability of $c_\Lambda^\omega(f_k)^{-1}\,\delta_\Lambda^\omega(f_k)$ for $k \geq m$ implies the summability of the sequence $\left|F_{\Lambda,k+1}\left(\omega_\Lambda \mid \omega\right) - F_{\Lambda,k}\left(\omega_\Lambda \mid \omega\right)\right|$. In particular $\left(F_{\Lambda,n}(\omega_\Lambda \mid \omega)\right)_{n\geq m}$ is a Cauchy sequence so the limit $\lim_{n\to+\infty} F_{\Lambda,n}\left(\omega_\Lambda \mid \omega\right) \triangleq \gamma_\Lambda^f\left(\omega_\Lambda \mid \omega\right)$ exits and satisfies (6.4.7) for each $\omega \in \Omega$. □

6.4.1 LIS $\rightsquigarrow$ specification

Proof of Theorem 6.3.12

Lemma (6.4.6) proves Item **1) (a)**.

To prove item **1) (b)**, we observe that $\gamma_\Lambda^f(A \mid \cdot)$ is clearly $\mathcal{F}_{\Lambda^c}$-measurable for every $\Lambda \in \mathcal{S}$ and every $A \in \mathcal{F}$. Moreover condition (b) of Definition 4.1.1 together with the presence of the indicator function $\mathbb{1}_{\omega_{\Lambda^c\cap[l_\Lambda,m_\Lambda]}}$ in the denominator of (6.3.10) imply that $\gamma_\Lambda^f(B \mid \cdot) = \mathbb{1}_B(\cdot)$ for every $\Lambda \in \mathcal{S}$ and every $B \in \mathcal{F}_{\Lambda^c}$. Therefore it suffices to show that

$$\sum_{\omega_{\Delta\setminus\Lambda}} \gamma_\Lambda^f\left(\omega_\Lambda \mid \omega_{\Lambda^c}\right)\gamma_\Delta^f\left(\omega_{\Delta\setminus\Lambda} \mid \omega_{\Delta^c}\right) \;=\; \gamma_\Delta^f\left(\omega_\Lambda \mid \omega_{\Delta^c}\right) \tag{6.4.8}$$

for each $\Lambda, \Delta \in \mathcal{S}$ such that $\Lambda \subset \Delta$ and each $\omega \in \Omega$. Let us denote, for each $\Gamma \subseteq \Delta$, each integer $n \geq l_\Gamma$ and $\omega_{-\infty}^n \in \Omega_{-\infty}^n$,

$$G_{\Gamma,n}\left(\,\cdot \mid \omega_{\Gamma^c}\right) \;\triangleq\; f_{[l_\Gamma,n]}\left(\,\cdot\;\mathbb{1}_{\omega_{\Gamma^c\cap[l_\Gamma,n]}}\;\middle|\;\omega_{\Gamma_-}\right).$$

Definition (6.3.10)–(6.3.11) becomes

$$\gamma_\Delta\left(h \mid \omega\right) \;=\; \lim_{n\to+\infty} \frac{G_{\Delta,n}\left(\,\mathbb{1}_{\omega_\Delta} \mid \omega_{\Delta^c}\right)}{G_{\Delta,n}\left(1 \mid \omega_{\Delta^c}\right)}. \tag{6.4.9}$$

Using the reconstruction property (4.2.5) of LIS with $l = l_\Delta$, $n = l_\Lambda - 1$ and $m = n$, we obtain

$$G_{\Delta,n}(\,\mathbb{1}_{\omega_{\Delta\setminus\Lambda}} \mid \omega_{\Delta^c}) \;=\; G_{\Delta,l_\Lambda-1}(\,\mathbb{1}_{\omega_{\Delta\setminus\Lambda}} \mid \omega_{\Delta^c}) \times G_{\Lambda,n}(1 \mid \omega_{\Lambda^c})$$

and

$$G_{\Lambda,n}(\,\mathbb{1}_{\omega_\Lambda} \mid \omega_{\Lambda^c}) \times G_{\Delta,l_\Lambda-1}(\,\mathbb{1}_{\omega_{\Delta\setminus\Lambda}} \mid \omega_{\Delta^c}) \;=\; G_{\Delta,n}(\,\mathbb{1}_{\omega_\Delta} \mid \omega_{\Delta^c})\;.$$

Therefore

$$\frac{G_{\Lambda,n}(\,\mathbb{1}_{\omega_\Lambda} \mid \omega_{\Lambda^c})}{G_{\Lambda,n}(1 \mid \omega_{\Lambda^c})} \times \frac{G_{\Delta,n}(\,\mathbb{1}_{\omega_{\Delta\setminus\Lambda}} \mid \omega_{\Delta^c})}{G_{\Delta,n}(1 \mid \omega_{\Delta^c})} \;=\; \frac{G_{\Delta,n}(\,\mathbb{1}_{\omega_\Delta} \mid \omega_{\Delta^c})}{G_{\Delta,n}(1 \mid \omega_{\Delta^c})}\;. \tag{6.4.10}$$

Identity (6.4.8) follows from (6.4.9) and (6.4.10).

We proceed with item **1) (c)**. By (6.4.7) and the summability of the bound ϵ_k [defined in (6.3.6)], $F_{\Lambda,n}(\omega_\Lambda \mid \cdot\,)$ converges uniformly to $\gamma_\Lambda(\omega_\Lambda \mid \cdot\,)$. As each $F_{\Lambda,n}$ is continuous on Ω, so is γ^f_Λ. Let us fix k_0 such that $\epsilon_k < 1$ for $k \geq k_0$. By (6.3.6) and the lower bound in (6.4.5)

$$\gamma^f_\Lambda \;\geq\; F_{\Lambda,k_0} \prod_{k=k_0}^{\infty} (1-\epsilon_k)\;.$$

The right-hand side is strictly positive on Ω due to the non-nullness of F_{Λ,k_0} and the summability of the ϵ_k. Hence γ^f_Λ is non-null on Ω.

To prove assertion **2)(a)** we consider $\mu \in \mathcal{G}\left(f^{(V,\omega)}\right)$ and denote

$$G^{(V,\omega)}_{\Lambda,n}\,(\,\cdot \mid \sigma_{\Lambda^c}) \;\triangleq\; f_{[l_\Lambda,n]}\Big(\,\cdot\;\mathbb{1}_{\sigma_{\Lambda^c\cap[l_\Lambda,n]}}\;\Big|\; \omega_{V_-}\,\sigma_{\Lambda_-\setminus V_-}\Big)$$

for all $\Lambda \in \mathcal{S}_b : \Lambda \subset V$ and $\omega, \sigma : \omega_{V_-}\,\sigma_{\Lambda_-\setminus V_-} \in \Omega_{\Lambda_-}$. By a straightforward extension of (6.4.9), the dominated convergence theorem and the consistency of μ with respect to $G^{(V,\omega)}_{\Lambda,n}$

$$\begin{aligned}\mu\,\gamma^f_\Lambda\,(\,\mathbb{1}_{\omega_\Lambda}) &= \lim_{n\to+\infty} \mu\, G^{(V,\omega)}_{\Lambda,n}\left(\frac{G^{(V,\omega)}_{\Lambda,n}\,(\,\mathbb{1}_{\omega_\Lambda} \mid \cdot\,)}{G^{(V,\omega)}_{\Lambda,n}\,(1 \mid \cdot\,)}\right)\\ &= \lim_{n\to+\infty} \mu\, G^{(V,\omega)}_{\Lambda,n}\,(\,\mathbb{1}_{\omega_\Lambda})\,.\end{aligned}$$

Applying the consistency hypothesis a second time we obtain $\mu\,\gamma^f_\Lambda = \mu$.

Assertion **2)(b)** is an immediate consequence of 2) (a) and of the fact that $|\mathcal{G}(\gamma)| = 1$ for all $\gamma \in \Pi_1$.

Finally we prove **2) (c)**. Let f^1 and f^2 be two LIS on $(\Omega, \mathcal{F})$, both in $b^{-1}(\Pi_1)$, and such that $\gamma^{f^1} = \gamma^{f^2}$. By 2) (c), $\mu^{f^1} = \mu^{f^2} \triangleq \mu$. The non-nullness of f^1 and f^2 on Ω implies that μ charges all open sets in Ω. Therefore, f^1_Λ and f^2_Λ coincide, on Ω, with the unique continuous realization of $E_\mu\left(\,\cdot \mid \mathcal{F}_{\Lambda_-}\right)$. $\square$

6.4.2 Specification $\rightsquigarrow$ LIS

Let us introduce the *spread* of a (bounded) function h on Ω:

$$\mathrm{Spr}(h) \;=\; \sup(h) - \inf(h)\;.$$

Lemma 6.4.11

1) *Let γ be a specification on Ω.*

(a) *If there exists an exhausting sequence of regions $\Lambda_n \subset \mathbb{Z}$ such that*

$$\lim_{n\to+\infty} \mathrm{Spr}\,(\gamma_{\Lambda_n} h) = 0 \tag{6.4.12}$$

for each continuous $\mathcal{F}$-measurable function h, then $|\mathcal{G}(\gamma)| \leq 1$.

(b) *If γ is continuous and $|\mathcal{G}(\gamma)| \leq 1$, then (6.4.12) holds for all exhausting sequences of regions $\Lambda_n \subset \mathbb{Z}$ and all continuous $\mathcal{F}$-measurable function h.*

2) *Let f be a LIS on Ω*

(a) *If for each $i \in \mathbb{Z}$ and each continuous $\mathcal{F}_{\leq i}$-measurable continuous function h*

$$\lim_{n\to+\infty} \mathrm{Spr}\left(f_{[i-n,i]} h\right) = 0 \;, \tag{6.4.13}$$

then $|\mathcal{G}(f)| \leq 1$.

(b) *If f is continuous and $|\mathcal{G}(f)| \leq 1$, then (6.4.13) is verified for all $i \in \mathbb{Z}$ and all continuous $\mathcal{F}_{\leq i}$-measurable continuous function h.*

Proof We proof part 1), the proof of 2) is similar. The obvious spread-reducing relation

$$\inf_{\widetilde{\omega}\in\Omega} h(\widetilde{\omega}) \;\leq\; (\gamma_\Lambda)h(\omega) \;\leq\; \sup_{\widetilde{\omega}\in\Omega} h(\widetilde{\omega}) \;,$$

valid for every bounded measurable function h on Ω and every configuration $\omega \in \Omega$, plus the consistency condition (3.1.2) imply that the sequence $\{\sup(\gamma_{\Lambda_n} h)\}$ is decreasing (and bounded below by $\inf h$), while the sequence $\{\inf(\gamma_{\Lambda_n} h)\}$ is increasing (and bounded above by $\sup h$). Therefore, if $\mu, \nu \in \mathcal{G}(\gamma)$,

$$\mu(h) - \nu(h) \;\leq\; \sup(\gamma_{\Lambda_n} h) - \inf(\gamma_{\Lambda_n} h) \tag{6.4.14}$$

for every n, which yields

$$|\mu(h) - \nu(h)| \;\leq\; \mathrm{Spr}(\gamma_{\Lambda_n} h) \tag{6.4.15}$$

for every n. This proves item **1)(a)**.

Regarding **1)(b)**, we observe that, as Ω is compact, there exist optimizing boundary conditions $\{\sigma^{(n)}\}$ and $\{\eta^{(n)}\}$ such that $(\gamma_{\Lambda_n} h)\left(\sigma^{(n)}\right) = \sup(\gamma_{\Lambda_n} h)$. and $(\gamma_{\Lambda_n} h)\left(\eta^{(n)}\right) = \inf(\gamma_{\Lambda_n} h)$. (Of course, both sequences of boundary conditions depend on h). Let $\overline{\rho}$ and $\underline{\rho}$ be respective accumulation point of the sequences of measures $\left\{\gamma_{\Lambda_n}(\,\cdot \mid \sigma^{(n)})\right\}$ and $\left\{\gamma_{\eta_n}(\,\cdot \mid \sigma^{(n)})\right\}$ (they exist by compactness). Then, $\overline{\rho}, \underline{\rho} \in \mathcal{G}(\gamma)$ (due to the continuity of γ) and

$$\lim_n \mathrm{Spr}(\gamma_{\Lambda_n} h) \;\leq\; \overline{\rho}(h) - \underline{\rho}(h) \;. \tag{6.4.16}$$

Hence the uniqueness of the consistent measure implies (6.4.12). We learnt this argument from Michael Aizenman (private communication). □

Our last auxiliary result refers to the following notion.

Definition 6.4.17
A **global specification** γ *over* $(\Omega, \mathcal{F})$ *is a family of probability kernels* $\{\gamma_V\}_{V\subset\mathbb{Z}}$, $\gamma_V : \mathcal{F}\times\Omega \to [0,1]$ *such that for all* $V \subset \mathbb{Z}$

(a) *For each* $A \in \mathcal{F}$, $\gamma_V(A \mid \cdot\,) \in \mathcal{F}_{V^c}$.

(b) *For each* $B \in \mathcal{F}_{V^c}$ *and* $\omega \in \Omega$, $\gamma_V(B \mid \omega) = \mathbb{1}_B(\omega)$.

(c) *For each* $W \subset \mathbb{Z} : W \supset V$, $\gamma_W\,\gamma_V = \gamma_W$.

Proposition 6.4.18
Let γ *be a continuous specification over* $(\Omega, \mathcal{F})$ *which satisfies a HUC. Then* γ *can be extended into a continuous global specification such that for every subset* $V \subset \mathbb{Z}$,

$$\gamma_V\left(h \mid \omega_{V^c}\right) \;\triangleq\; \lim_{\Lambda\uparrow V} \gamma_\Lambda\left(h \mid \omega\right) \tag{6.4.19}$$

for all continuous functions $h \in \mathcal{F}$ *and all* $\omega \in \Omega$. *Moreover for all* $V \subset \mathbb{Z}$ *and all* $\omega \in \Omega$,

$$\mathcal{G}\left(\gamma^{(V,\omega)}\right) \;=\; \left\{\gamma_V\left(\,\cdot \mid \omega\right)\right\}. \tag{6.4.20}$$

Georgii (1988) gives a proof of this proposition in the Dobrushin regime (Theorem 8.23). The same proof extends, with minor adaptations, under a HUC (see Fernández and Pfister, 1997).

Proof of Theorem 6.3.17

Items **1) (a)–(b)** are proven in Proposition (6.4.18).

There are three things to prove regarding **1) (c)**:

(i) Continuity of f^γ. This is, in fact, an application of Proposition (6.4.18).

(ii) Non-nullness of f^γ. Consider $\Lambda \in \mathcal{S}$, $\omega \in \Omega$, $n \geq m_\Lambda$ and $k \geq 0$. By the non-nullness and the continuity of γ and the compactness of Ω_{Λ^c}, there exists $\widetilde{\omega} \in \Omega_{\Lambda^c}$ such that

$$0 \;<\; \gamma_\Lambda\left(\omega_\Lambda \mid \widetilde{\omega}\right) \;=\; \inf_{\omega\in\Omega_\Lambda^c} \gamma_\Lambda\left(\omega_\Lambda \mid \omega\right) \;\triangleq\; c\left(\Lambda, \omega_\Lambda\right).$$

Therefore by the consistency of γ

$$\begin{aligned} f^\gamma_\Lambda\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right) \;&=\; \lim_{k\to\infty} \gamma_{[l_\Lambda, n+k]}\left(\omega_\Lambda \mid \omega\right) \;=\; \lim_{k\to\infty}\Big(\gamma_{[l_\Lambda,n+k]}\gamma_\Lambda\Big)\left(\omega_\Lambda \mid \omega\right) \\ &\geq\; c(\Lambda, \omega_\Lambda) \;>\; 0\,. \end{aligned}$$

(iii) Hereditary uniqueness. Let us fix $\omega \in \Omega$ and $V \in \{[j,+\infty[,\ j \in \mathbb{Z}\} \cup \mathbb{Z}$. For each $i \in \mathbb{Z}$ and $h \in \mathcal{F}_{\leq i}$. We have

$$\mathrm{Spr}\left(f^{\gamma\,(V,\omega)}_{[i-k,i]} h\right) \;\leq\; \lim_{n\to+\infty} \mathrm{Spr}\left(\gamma^{(V,\omega)}_{[i-k,i+n]} h\right) \;.$$

As, by hypothesis, each specification $\gamma^{(V,\omega)}$ admits an unique Gibbs measure, it follows from lemma 6.4.11 1) (b) that

$$\lim_{k\to+\infty} \mathrm{Spr}\left(f^{\gamma\,(V,\omega)}_{[i-k,i]} h\right) \;=\; 0\;.$$

This proves that $\left|\mathcal{G}\left(f^{\gamma\,(V,\omega)}\right)\right| = 1$ by lemma 6.4.11 2) (a).

The uniqueness part of assertion **2) (a)** is contained in the just proven hereditary uniqueness. To show that $\mu^\gamma \in \mathcal{G}(f^\gamma)$, consider $\Lambda \in \mathcal{S}_b$ and h a continuous $\mathcal{F}_{\leq m_\Lambda}$-measurable function. By the dominated convergence theorem

$$\mu^\gamma f^\gamma_\Lambda(h) = \lim_{n\to+\infty} \int \gamma_{\Lambda\cup\Lambda^{(n)}_+}\left(h \mid \xi\right) \mu^\gamma(d\xi).$$

The consistency of μ^γ with respect to γ implies, hence, that $\mu^\gamma \in \mathcal{G}(f^\gamma)$.

To prove assertion **2) (b)**, let γ^1 and γ^2 such that $f^{\gamma^1} = f^{\gamma^2}$. By 2) (a) $\mu^{\gamma^1} = \mu^{\gamma^2} \triangleq \mu$. The non-nullness of γ^1 and γ^2 implies that μ charges all open sets on Ω. Therefore for each $\Lambda \in \mathcal{S}$, γ^1_Λ and γ^2_Λ coincide with the unique continuous realization of $E_\mu\left(\,\cdot \mid \mathcal{F}_{\Lambda^c}\right)$. □

Proof of Theorem 6.3.18

To prove item **(a)**, let us recall one of the equivalent definitions of the variational distance between probability measures over $(\Omega_i, \mathcal{F}_i)$

$$\|\mu - \nu\| \;=\; \sup_{h\in\mathcal{F}_i} \frac{|\mu(h) - \nu(h)|}{\mathrm{Spr}(h)}\;.$$

For a proof of this result see for example Georgii (1988) (section 8.1). By the consistency of $\mathring{\gamma}_i$ with respect to $\gamma_{[i,i+k]},\ k \geq 0$

$$\mathring{f}^\gamma_i\left(\,\cdot \mid \omega^{i-1}_\infty\right) \;\triangleq\; \lim_{k\to+\infty} \mathring{\gamma}_{[i,i+k]}\left(\,\cdot \mid \omega\right) \;=\; \lim_{k\to+\infty} \gamma_{[i,i+k]}\;\mathring{\gamma}_i\left(\,\cdot \mid \omega\right)\;.$$

Therefore, by dominated convergence,

$$C_{ij}(f^\gamma) \;\leq\; \sup_{\substack{\xi,\eta\in\Omega \\ \xi^{i-1}_{-\infty} \overset{\neq j}{=} \eta^{i-1}_{-\infty}}} \left\|\mathring{\gamma}_i\left(\,\cdot \mid \xi\right) - \mathring{\gamma}_i\left(\,\cdot \mid \eta\right)\right\|\,. \tag{6.4.21}$$

Since γ is continuous, we can do an infinite telescoping of (6.4.21) to obtain

$$C_{ij}(f^\gamma) \;\leq\; \sum_{\substack{k=j \\ \text{or } k>i}} C_{ik}(\gamma)\;.$$

Thus

$$\sum_{j:j<i} C_{ij}(f^\gamma) \;\leq\; \sum_{j:j\neq i} C_{ij}(\gamma) < 1\;.$$

To show assertion **(b)**, consider $\gamma \in \Pi_2$ for which there exists a constant $K > 0$ such that for every cylinder set $A = \{x_l^m\} \in \Omega_l^m$ there exist integers n, p satisfying

$$\gamma_{[n,p]}(A \mid \xi) \;\geq\; K\,\gamma_{[n,p]}(A \mid \eta) \quad \text{for all } \xi, \eta \in \Omega\;.$$

Hence, by consistency of γ, we have that for some fixed $\sigma \in \Omega$ and for each $k \geq 0$

$$\gamma_{[n,p+k]}(A \mid \xi) \;=\; \int \gamma_{[n,p]}(A \mid \omega)\,\gamma_{[n,p+k]}(d\omega \mid \xi) \;\geq\; K\,\gamma_{[n,p]}(A \mid \sigma)\;.$$

In a similar way we obtain

$$\gamma_{[n,p+k]}(A \mid \eta) \;\leq\; \frac{1}{K}\,\gamma_{[n,p]}(A \mid \sigma)\;.$$

We conclude that for each $k \geq 0$

$$\gamma_{[n,p+k]}(A \mid \xi) \;\geq\; K^2\,\gamma_{[n,p+k]}(A \mid \eta)\;.$$

Letting $k \to \infty$ we obtain, due to definition (6.3.16), that $f^\gamma_{[n,m]}(A \mid \xi) \geq K^2 f^\gamma_{[n,m]}(A \mid \eta)$.
□

6.4.3 LIS $\leftrightsquigarrow$ specification

Proof of Theorem 6.3.19

Assertion **1) (a)** is a direct consequence of inequality (6.4.7) of Lemma 6.4.6 with $\Lambda = \{k\}$ and $n = j - 1$.

Assertion **1) (b)** follows from the $n \to \infty$ limit of inequalities (6.4.4) and the fact that $0 \leq F_{\Lambda,n} \leq 1$.

To prove assertion **2)**, let $k, j \in \mathbb{Z}$ such that $j < k$ and consider $\omega, \sigma \in \Omega$ such that $\omega \overset{\neq j}{=} \sigma$. As a direct consequence of definitions 6.3.3–6.3.4 we have that, for all $i \geq k$,

$$\big(1 - c_j^\omega(\gamma_i)^{-1}\,\delta_j^\omega(\gamma_i)\big) \times \gamma_i(\omega_i \mid \omega_{-\infty}^{i-1}\,\omega_{i+1}^{+\infty}) \;\leq\; \gamma_i(\sigma_i \mid \sigma_{-\infty}^{i-1}\,\sigma_{i+1}^{+\infty}) \tag{6.4.22}$$

and

$$\gamma_i(\sigma_i \mid \sigma_{-\infty}^{i-1}\,\sigma_{i+1}^{+\infty}) \;\leq\; \big(1 + c_j^\omega(\gamma_i)^{-1}\,\delta_j^\omega(\gamma_i)\big) \times \gamma_i(\omega_i \mid \omega_{-\infty}^{i-1}\,\omega_{i+1}^{+\infty})\,. \tag{6.4.23}$$

By the specification reconstruction formula (Theorem 3.1.8 (ii)) with $\Lambda = \{n+1\}$ and $\Gamma = [l_\Lambda, n]$ we have

$$\gamma_{[l_\Lambda, n+1]}\left(\sigma_\Lambda \mid \sigma_{\Lambda_-}\,\sigma_{m_\Lambda+1}^{+\infty}\right) \;=\; \sum_{\sigma_{m_\Lambda+1}^{n}} \frac{\gamma_{n+1}\left(\sigma_{n+1} \mid \sigma_{-\infty}^{n}\,\sigma_{n+2}^{+\infty}\right)}{\displaystyle\sum_{\xi_{n+1}} \frac{\gamma_{n+1}\left(\xi_{n+1} \mid \sigma_{-\infty}^{n}\,\sigma_{n+2}^{+\infty}\right)}{\gamma_{[l_\Lambda,n]}\left(\sigma_{l_\Lambda}^{n} \mid \sigma_{\Lambda_-}\,\xi_{n+1}\,\sigma_{n+2}^{+\infty}\right)}}.$$

Using (6.4.22) and (6.4.23) it is easy to show, by induction over $n \geq m_\Lambda + 1$, that

$$\gamma_{[l_\Lambda,n]}\left(\omega_\Lambda \mid \xi_{\Lambda_-}\,\omega_{n+1}^{+\infty}\right) \;\leq\; \gamma_{[l_\Lambda,n]}\left(\omega_\Lambda \mid \eta_{\Lambda_-}\,\omega_{n+1}^{+\infty}\right) \times \prod_{i=k}^{n} \frac{1 - c_j^\omega(\gamma_k)^{-1}\,\delta_j^\omega(\gamma_i)}{1 + c_j^\omega(\gamma_k)^{-1}\,\delta_j^\omega(\gamma_i)}$$

for all $\xi, \eta \in \Omega : \xi \stackrel{\neq j}{=} \eta \stackrel{\neq j}{=} \omega$. Taking the limit when n tends to infinity, we obtain 2). □

Proof of Theorem 6.3.20

For the proof of item **(a)** we consider $\gamma \in \Pi_2$ such that $f^\gamma \in \Theta_1$ and fix $\Lambda \in \mathcal{S}$ and $\omega \in \Omega$. By definition of the maps b and c [see (6.3.10)–(6.3.11) and (6.3.16)], we have that

$$\gamma_\Lambda^{f^\gamma}\left(\omega_\Lambda \mid \omega\right) \;=\; \lim_{n\to+\infty}\lim_{k\to+\infty} \frac{\gamma_{[l_\Lambda,n+k]}\left(\omega_\Lambda\,\omega_{\Lambda^c\cap[l_\Lambda,n]} \mid \omega_{\Lambda_-}\,\omega_{n+k+1}^{+\infty}\right)}{\gamma_{[l_\Lambda,n+k]}\left(\omega_{\Lambda^c\cap[l_\Lambda,n]} \mid \omega_{\Lambda_-}\,\omega_{n+k+1}^{+\infty}\right)}\,. \tag{6.4.24}$$

The consistency of γ_Λ and $\gamma_{[l_\Lambda,n+k]}$ implies

$$\begin{aligned}&\gamma_{[l_\Lambda,n+k]}\left(\omega_\Lambda\,\omega_{\Lambda^c\cap[l_\Lambda,n]} \mid \omega_{\Lambda_-}\,\omega_{n+k+1}^{+\infty}\right) \;=\\ &\sum_{\xi_{n+1}^{n+k}} \gamma_\Lambda\left(\omega_\Lambda \mid \omega_{\Lambda^c\cap]-\infty,n]}\,\xi_{n+1}^{n+k}\,\omega_{n+k+1}^{+\infty}\right)\,\gamma_{[l_\Lambda,n+k]}\left(\omega_{\Lambda^c\cap[l_\Lambda,n]}\,\xi_{n+1}^{n+k} \mid \omega_{\Lambda_-}\,\omega_{n+k+1}^{+\infty}\right)\,.\end{aligned} \tag{6.4.25}$$

By continuity of $\gamma_\Lambda\left(\omega_\Lambda \mid \cdot\,\right)$ we have that, for each $\varepsilon > 0$,

$$\left|\gamma_\Lambda\left(\omega_\Lambda \mid \omega_{\Lambda^c\cap]-\infty,n]}\,\xi_{n+1}^{n+k}\,\omega_{n+k+1}^{+\infty}\right) - \gamma_\Lambda\left(\omega_\Lambda \mid \omega\right)\right| \;<\; \varepsilon$$

for n large enough uniformly in k. Combining this with (6.4.24)–(6.4.25) we conclude that

$$\left|\gamma_\Lambda^{f^\gamma}\left(\omega_\Lambda \mid \omega\right) - \gamma_\Lambda\left(\omega_\Lambda \mid \omega\right)\right| \;<\; \varepsilon$$

for every $\varepsilon > 0$. Therefore $\gamma^{f^\gamma} = \gamma$.

To prove item **(b)**, consider $f \in \Theta_1$ such that $\gamma^f \in \Pi_2$ and fix $\Lambda \in \mathcal{S}_b$ and $\omega \in \Omega$. Let us denote $V = [l_\Lambda, +\infty[$. Since γ^f satisfies a HUC equation, (6.4.19) and definition (6.3.16) yield

$$f_\Lambda^{\gamma^f}\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right) \triangleq \lim_{n\to+\infty} \gamma^f_{[l_\Lambda, m_\Lambda+n]}\left(\omega_\Lambda \mid \omega\right) = \gamma_V^f\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right) . \tag{6.4.26}$$

Combining (6.4.20) with assertion 2) (a) of Theorem 6.3.12 we obtain that

$$\mathcal{G}\left(f^{(V,\omega)}\right) = \left\{\gamma_V^f\left(\cdot \mid \omega_{\Lambda_-}\right)\right\}.$$

Therefore

$$\gamma_V^f\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right) = \gamma_\Lambda^f\left(f_\Lambda^{(V,\omega)}\left(\omega_\Lambda \mid \cdot\right) \mid \omega_{\Lambda_-}\right) = f_\Lambda\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right).$$

The last equality is a consequence of the definition (6.2.3). By (6.4.26) this implies that

$$f_V^{\gamma^f}\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right) = f_\Lambda\left(\omega_\Lambda \mid \omega_{\Lambda_-}\right).$$

Item **(c)** is a direct consequence of Theorem 6.3.19 and the following result. □

Lemma 6.4.27
Let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function and $(u_i)_{i\in\mathbb{N}}$ be a sequence taking values in $]0,1[$ for which there exists $m \geq 0$ such that $u_i \leq m\, h(i)$. Then there exists $M \geq 0$ such that

$$1 - \prod_{i=k}^{+\infty} \frac{1-u_i}{1+u_i} \leq M\, H(k-1) ,$$

where $H(x) = \int_x^{+\infty} h(t)\, dt$.

The proof is left to the reader. □

Bibliography

[1] H. Berbee. Chains with infinte connections: Uniqueness and Markov representation. *Prob. Th. Rel. Fields*, 76:243–53, 1987.

[2] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Springer-Verlag, Berlin, 1975. Lecture Notes in Mathematics, Vol. 470.

[3] M. Bramson and S. A. Kalikow. Nonuniqueness in g-functions. *Israeli J. Math.*, 84:153–60, 1993.

[4] 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/`).

[5] P. Bühlmann and A. J. Wyner. Variable length Markov chains. *Ann. Statist.*, 27:480–513, 1999.

[6] Z. Coelho and A. N. Quas. Criteria for $\overline{d}$-continuity. *Trans. Amer. Math. Soc.*, 350(8):3257–3268, 1998.

[7] R. L. Dobrushin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of probability and its applications*, 13:197–224, 1968.

[8] W. Doeblin and R. Fortet. Sur les chaînes à liaisons complétes. *Bull. Soc. Math. France*, 65:132–148, 1937.

[9] R. Fernández and C.-E. Pfister. Global specifications and nonquasilocality of projections of Gibbs measures. *The Annals of Probability*, 25(3):1284–1315, 1997.

[10] H. Fllmer. A covariance estimate for Gibbs measures. *J. Funct. Anal.*, 3:387–395, 1982.

[11] J. Fröhlich and T. Spencer. The Phase Transition in the One-Dimensional Ising Model with $\frac{1}{r^2}$ Interaction Energy. *Commn. Math. Phys.*, 84:87–101, 1982.

[12] H.-O. Georgii. Stochastische Felder und ihre Anwendung auf Interaktionssysteme. *Lecture Notes, Institut für Angewandte Mathematik, Universität Heidelberg*, 1974.

[13] H.-O. Georgii. *Gibbs Measures and Phase Transitions.* Walter de Gruyter (de Gruyter Studies in Mathematics, Vol. 9), Berlin–New York, 1988.

[14] T. E. Harris. On chains of infinite order. *Pacific J. Math.*, 5:707–24, 1955.

[15] P. Hulse. A class of unique g-measure. *Ergod. Th. and Dynam. Sys.*, 17:1383–1392, 1997.

[16] 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.

[17] M. Iosifescu and S. Grigorescu. *Dependence with Complete Connections and its Applications.* Cambridge University Press, Cambridge, UK, 1990.

[18] A. Johansson and A. Öberg. Square summability of variations of g-functions and uniqueness of g-measures. Preprint, 2002.

[19] S. Kalikow. Random Markov processes and uniform martingales. *Isr. J. Math.*, 71(1):33–54, 1990.

[20] M. Keane. Strongly mixing g-measures. *Inventiones Math.*, 16:309–24, 1972.

[21] G. Keller. *Equilibrium states in ergodic theory*, volume 42 of *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge, 1998.

[22] O. K. Kozlov. Gibbs description of a system of random variables. *Probl. Inform. Transmission*, 10:258–265, 1974.

[23] S. P. Lalley. Regeneration representation for one-dimensional Gibbs states. *Ann. Prob.*, 14:1262–71, 1986.

[24] S. P. Lalley. Regenerative in one-dimensional Gibbs states and chains with complete connections. *Resenhas IME-USP*, 4:249–80, 2000.

[25] O. E. Lanford. *Entropy and equilibrium states in classical statistical mechanics.* In: A. Lenard (ed.), Statiscal mechanics and mathematical problems, Battelle Seattle Rencontres 1971, LPNh 20, pp. 1-113., 1973.

[26] F. Ledrappier. Principe variationnel et systèmes dynamiques symboliques. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 30:185–202, 1974.

[27] O. Onicescu and G. Mihoc. Sur les chaînes statistiques. *C. R. Acad. Sci. Paris*, 200:511—12, 1935.

[28] A. N. Quas. Non-ergodicity for $\mathcal{C}^1$ expanding maps and g-measures. *Ergod. Th. and Dynam. Sys.*, 16:531–543, 1996.

[29] A. Raftery. A model for high-order Markov chains. *J. R. Statist. Soc. B*, 47:528–539, 1985.

[30] D. Ruelle. *Thermodynamic formalism.* Encyclopedia of Mathematics 5, New-York: Addison Wesley, 1978.

[31] B. Simon. *The Statistical Mechanics of Lattice Gases.* Series in physics. Princeton, 1993.

[32] F. Spitzer. Random fields and interacting particle systems. Notes on lectures at the 1971 MAA Summer Seminar Williamstown, Mass., Math. Assoc. America., 1971.

[33] W. G. Sullivan. Potentials for almost Markovian random fields. *Comm. Math. Phys.*, 33:61–74, 1973.

[34] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.*, 214:375–87, 1975.

Résumé

Nous introduisons un formalisme de mécanique statistique pour l'étude des processus stochastiques discrets (chaînes) pour lesquels on prouve : (i) des propriétés générales de chaînes extrémales, incluant la trivialité de la tribu queue, les corrélations à courtes portées, la réalisation via des limites à volumes infinis et l'ergodicité, (ii) deux nouvelles conditions pour l'unicité de la chaîne cohérente, (iii) des résultats de perte de mémoire et des propriétés de mélange pour des chaînes sous le régime de Dobrushin. Ces résultats sont complémentaires de ceux existant dans la litérature, et généralisent les résultats markoviens basés sur le cœfficient d'ergodicité de Dobrushin. D'autre part, on considère des systèmes à alphabet fini, pouvant avoir une grammaire. On établit des conditions pour qu'une chaîne définisse une mesure de Gibbs et vice-versa. Nos conditions généralisent les résultats d'équivalence bien connus entre les chaînes et les champs markoviens, aussi bien que le caractère gibbsien des processus ayant un taux de continuité exponentiel. Nos arguments sont purement probabilistes; ils sont basés sur l'étude de systèmes réguliers de probabilités conditionnelles (spécifications). De plus, on discute de l'équivalence des critères d'unicité pour les chaînes et les champs et on établit des bornes pour les taux de continuité des systèmes respectifs de probabilités conditionnelles. On prouve également un théorème auxiliaire de (re)construction pour les spécifications en partant de conditionnement sur un site, qui s'applique dans un cadre plus général.