Une Approche vers la Description et l'Identification d'une Classe de Champs Aléatoires

Serguei Dachian

Thèse de Doctorat
de l'Université Paris 6

Spécialité : **Mathématiques**

Présentée par

Sergueï Dachian

pour obtenir le grade de Docteur de l'Université Paris 6

Sujet de la thèse : **Une approche vers la description et l'identification d'une classe de champs aléatoires.**

Soutenue le 21 janvier 1999 devant le jury composé de :

Directeur de thèse	:	**D. Bosq** (Université Paris 6)
Directeur de thèse	:	**Yu. A. Kutoyants** (Université du Maine)
Rapporteur	:	**F. Comets** (Université Paris 7)
Rapporteur	:	**X. Guyon** (Université Paris 1)
Examinateur	:	**J.-P. Lepeltier** (Université du Maine)

Remerciements

Mes remerciements s'adressent tout d'abord à mes directeurs de recherche D. Bosq et Yu. A. Kutoyants qui m'ont fait découvrir le domaine de la statistique mathématique des processus.

Special thanks should be addressed here to B. S. Nahapetian who has introduced me to the theory of random fields, and without whom this work would surely not be possible. His fruitful advices and ideas have always directed my studies and reflections.

Je tiens à exprimer ma reconnaissance à tout le personnel du "Laboratoire de Statistique et Processus" de l'Université du Maine pour l'accueil chaleureux et pour la place qui m'a été offerte au sein du laboratoire.

Je remercie également F. Comets et X. Guyon de l'intérêt qu'ils ont manifesté en acceptant d'être rapporteurs. Leur remarques m'ont beaucoup aidé à améliorer ce travail.

Je voudrais exprimer ma reconnaissance envers la jeune équipe du "Laboratoire de Statistique et Processus" de l'Université du Maine, à savoir C. Aubry, A. Dabye, S. Iacus, I. Negri, A. Matoussi et B. Saussereau, pour les discussions que l'on a pu échanger.

Last but not least, I wish to thank my parents and friends who have look at my studies with great interest from far away.

Résumé

Une nouvelle approche de la description des champs aléatoires sur le réseau entier ν-dimensionnel $\mathbb{Z}^\nu$ est présentée. Les champs aléatoires sont décrits en terme de certaines fonctions de sous-ensembles de $\mathbb{Z}^\nu$, à savoir les P-fonctions, les Q-fonctions, les H-fonctions, les Q-systèmes, les H-systèmes et les systèmes ponctuels. La corrélation avec la description Gibbsienne classique est montrée. Une attention particulière est portée au cas quasilocal. Les champs aléatoires non-Gibbsiens sont aussi considérés. Un procédé général pour construire des champs aléatoires non-Gibbsiens est donné. La solution du problème de Dobrushin concernant la description d'un champ aléatoire par ses distributions conditionnelles ponctuelles est déduite de notre approche.

Ensuite, le problème de l'estimation paramétrique pour les champs aléatoires de Gibbs est considéré. Le champ est supposé spécifié en terme d'un système ponctuel local invariant par translation. Un estimateur du système ponctuel est construit comme un rapport de certaines fréquences conditionnelles empiriques. Ses consistances exponentielle et $\mathbf{L}^p$ uniformes sont démontrées. Finalement, le problème nonparamétrique de l'estimation d'un système ponctuel quasilocal est considéré. Un estimateur du système ponctuel est construit par la méthode de "sieves". Ses consistances exponentielle et $\mathbf{L}^p$ sont prouvées dans des cadres différents. Les résultats sont valides indépendamment de la non-unicité et de la perte de l'invariance par translation.

Mots clés : *champs aléatoires, champs aléatoires de Gibbs, champs aléatoires non-Gibbsiens, localité, quasilocalité, P-fonctions, Q-fonctions, H-fonctions, Q-systèmes, H-systèmes, systèmes ponctuels, estimation paramétrique, estimation nonparamétrique, méthode de "sieves", consistance.*

Abstract

A new approach towards description of random fields on the ν-dimensional integer lattice $\mathbb{Z}^\nu$ is presented. The random fields are described by means of some functions of subsets of $\mathbb{Z}^\nu$, namely P-functions, Q-functions, H-functions, Q-systems, H-systems and one-point systems. Interconnection with classical Gibbs description is shown. Special attention is paid to quasilocal case. Non-Gibbsian random fields are also considered. A general scheme for constructing non-Gibbsian random fields is given. The solution to Dobrushin's problem concerning the description of random field by means of its one-point conditional distributions is deduced from our approach.

Further the problems of parametric estimation for Gibbs random fields is considered. The field is supposed to be specified through a translation invariant local one-point system. An estimator of one-point system is constructed as a ratio of some empirical conditional frequencies, and its uniform exponential and $\mathbf{L}^p$ consistencies are proved. Finally the nonparametric problem of estimation of quasilocal one-point systems is considered. An estimator of one-point system is constructed by the method of sieves, and its exponential and $\mathbf{L}^p$ consistencies are proved in different setups. The results hold regardless of non-uniqueness and translation invariance breaking.

Key words: *random fields, Gibbs random fields, non-Gibbsian random fields, locality, quasilocality, P-functions, Q-functions, H-functions, Q-systems, H-systems, one-point systems, parametric estimation, nonparametric estimation, method of sieves, consistency.*

Table des matières

Introduction

Cette thèse est constituée de deux parties. La Partie I traite de la description des champs aléatoires et la Partie II de l'identification des champs aléatoires.

Description des champs aléatoires

La théorie des champs aléatoires de Gibbs sur le réseau entier ν-dimensionnel $\mathbb{Z}^\nu$, $\nu \geqslant 1$, trouve ses origines dans la physique statistique. Elle est devenue une théorie mathématique rigoureuse principalement grâce aux travaux de R. L. Dobrushin dans les années soixante. On pourra se référer à ses travaux précurseurs [8] – [10]. Une présentation exhaustive de la théorie des champs aléatoires de Gibbs peut être trouvée dans le livre de H.-O. Georgii [12] où l'auteur, tout en restant dans la plus grande généralité, donne un grand nombre d'exemples et de détails.

Dans la première partie de ce travail (Chapitres I–VI) on présente une nouvelle approche de la description des champs aléatoires sur le réseau $\mathbb{Z}^\nu$ à valeurs dans un espace d'états fini $\mathscr{X}$. Une attention plus particulière est portée au cas où l'espace d'états est $\mathscr{X} = \{0,1\}$.

L'idée sous-jacente utilisée en physique statistique est de décrire les champs aléatoires par des spécifications de Gibbs exprimées par des potentiels d'interaction. L'idée principale de notre approche est d'exprimer les spécifications directement en terme des Hamiltoniens sans utiliser la notion de potentiel d'interaction. C'est une approche très générale qui nous permet aussi de décrire des champs aléatoires non-Gibbsiens.

On donne la représentation, en nos termes, de certains champs aléatoires non-Gibbsiens. De plus, on présente un procédé général de construction de champs aléatoires non-Gibbsiens. Notons que le rôle des champs aléatoires non-Gibbsiens dans la physique statistique est de plus en plus important. Le sujet est actuellement devenu le centre d'intérêt de plusieurs travaux (voir par exemple R. B. Israel [16], J. L. Lebowitz et C. Maes [18], R. H. Schonmann [23], A. van Enter, R. Fernandez et A. Sokal [25]).

Remarquons aussi que l'approche proposée permet de donner la solution d'un vieux problème posé par Dobrushin concernant la description d'un champ aléatoire par ses distributions conditionnelles ponctuelles. On présente une condition nécessaire et suffisante pour qu'un système de distributions conditionnelles ponctuelles soit un sous-système d'une spécification.

Dans le Chapitre I, on rappelle des notions et des résultats bien connus de la théorie des champs aléatoires, plus particulièrement de la théorie des champs aléatoires de Gibbs.

Dans le Chapitre II, on donne une alternative équivalente à la description de Kolmogorov des champs aléatoires. Cette description alternative, qui est basée sur une généralisation de la notion de fonction de corrélation à volume infini, fait apparaître la nature combinatoire de notre approche. La notion de P-fonction est introduite dans le but d'effectuer cette généralisation.

Dans le Chapitre III, on montre que l'on peut construire des P-fonctions comme limites de fonctions de corrélation à volume fini (ou plutôt leurs généralisations). Ces dernières sont exprimées en terme de fonctions de partition généralisées (Q-fonctions) ou, de manière équivalente, en terme de facteurs de Boltzmann généralisés (H-fonctions). Dans notre cas les H-fonctions sont des fonctions positives arbitraires. Ensuite on introduit les systèmes de distribution de probabilités consistants dans le sens de Dobrushin. Ces systèmes correspondent aux distributions conditionnelles dans les volumes finis avec condition extérieure vide (vacuum). On décrit ces systèmes en terme des Q-fonctions et/ou H-fonctions correspondantes. Finalement, on donne en terme de développement "cluster" d'une Q-fonction, une condition suffisante générale pour l'existence d'une P-fonction limite.

Même si les Q-fonctions nous permettent de construire des P-fonctions (et donc des champs aléatoires), elles sont insuffisantes pour décrire des champs aléatoires car elles déterminent uniquement les distributions conditionnelles dans les volumes finis avec condition extérieure vide, mais pas toute la spécification. Pour remédier à cela, on introduit au Chapitre IV des systèmes consistants de Q-fonctions (ou, de manière équivalente, de H-fonctions) que l'on appelle Q-systèmes (respectivement H-systèmes). On prouve que les spécifications "vacuum" (ou, autrement dit, les spécifications faiblement positives) peuvent être décrites par ces Q-systèmes et/ou H-systèmes. On montre que les spécifications

que nous décrivons peuvent être non-Gibbsiennes et on donne un procédé général pour construire des spécifications non-Gibbsiennes.

En regardant attentivement la définition d'un H-système (Q-système) consistant on remarque que l'information contenue dans un H-système (Q-système) est redondante. Ainsi, on peut envisager la description des spécifications par des systèmes plus simples que les H-systèmes et/ou Q-systèmes. Effectivement, on montre dans le Chapitre V que l'on peut décrire des spécifications "vacuum" par des sous-systèmes ponctuels de H-systèmes consistants que l'on appelle "systèmes ponctuels" (one-point systems). Notons ici qu'en introduisant ces systèmes ponctuels on donne la solution d'un vieux problème posé par Dobrushin concernant la description des champs aléatoires par ses distributions conditionnelles ponctuelles. La condition figurant dans la définition de système ponctuel n'est rien d'autre que la condition nécessaire et suffisante pour qu'un système de distributions conditionnelles ponctuelles soit un sous-système d'une spécification. Finalement on donne dans ce chapitre une condition nécessaire et suffisante pour qu'un système ponctuel soit Gibbsien.

Dans le Chapitre VI on se concentre sur la description des spécifications quasi-locales car elles sont très importantes dans la théorie des champs aléatoires. D'abord on considère les spécifications "vacuum" et on applique les résultats des Chapitres IV et V en donnant une condition nécessaire et suffisante pour qu'un H-système (respectivement Q-système, système ponctuel) corresponde à une spécification quasi-locale. Ensuite on remplace la condition "vacuum" (condition de positivité faible) par une condition légèrement différente, et on montre que dans ce cas on peut décrire les spécifications par des H-fonctions et/ou Q-fonctions qui satisfont certaines conditions supplémentaires.

Toutes nos considérations sont menées dans le cas de l'espace d'états $\mathscr{X} = \{0,1\}$. Dans tous les chapitres, on montre les généralisations possibles dans le cas d'un espace d'états fini arbitraire. La plupart des résultats pourraient aussi être généralisés dans le cas d'un espace d'états infini, mais cela nécessiterait plus de notations et d'hypothèses topologiques.

Cette première partie de la thèse a été effectuée en collaboration avec B. S. Nahapetian de l'Institut de Mathématiques, Érévan, Arménie. Certains résultats de cette partie ont été présentés dans [4], [6] et [7]. Notons finalement qu'une approche similaire pour des processus ponctuels a été considérée dans le travail de

R. V. Ambartzumian et H. S. Sukiasian [1].

Identification des champs aléatoires

L'inférence statistique pour les champs aléatoires de Gibbs est très intéressante et très importante car les résultats peuvent être appliqués dans ce qui est communément appelé le "traitement d'image". L'inférence statistique paramétrique pour les champs aléatoires de Gibbs est actuellement bien développée dans le cadre Gibbsien classique. L'état actuel de cette théorie est bien présenté dans le livre de X. Guyon [14]. On pourra aussi se rapporter à des références citées dans ce livre sur les travaux de F. Comets, B. Gidas, M. Janžura, D.K. Pickard, L. Younes, *et al.* Pour plus d'informations sur le traitement d'image et sur l'inférence statistique paramétrique pour les champs aléatoires de Gibbs, un lecteur intéressé peut aussi voir [3], [11], [15], [21], [22] and [26] – [112].

Contrairement à l'inférence statistique paramétrique pour les champs aléatoires de Gibbs, l'inférence nonparamétrique parait être moins étudiée. On peut mentionner ici une prépublication de C. Ji [15] où l'auteur considère le cadre Gibbsien classique quand le champ aléatoire est décrit par un potentiel d'interaction de paire à décroissance exponentielle. Pour ce modèle il étudie un estimateur "sieve" de ce qu'il appelle les "caractéristiques locales". La démonstration qu'il présente nécessite quelques rectifications.

Dans la deuxième partie de ce travail (Chapitres VII–VIII), on considère le problème de l'inférence statistique pour les champs aléatoires. Plus précisément on se concentre sur les champs aléatoires spécifiés en terme de systèmes ponctuels invariants par translation (stationnaires), ces derniers constituants une paramétrisation des champs aléatoires appropriée à l'inférence statistique.

On considère d'abord le problème d'estimation des systèmes ponctuels locaux. Évidemment, le problème est paramétrique dans ce cas. On suppose que $\boldsymbol{h} \in \mathscr{H}^V_{A,B}$ est un système ponctuel inconnu qui induit un ensemble $\mathscr{G}(\boldsymbol{h})$ de champs aléatoires de Gibbs ($\mathscr{H}^V_{A,B}$ est ici une certaine classe de systèmes ponctuels locaux). On observe une réalisation d'un champ aléatoire $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ dans une fenêtre d'observation Λ_n (le cube symétrique de coté n centré à l'origine de $\mathbb{Z}^\nu$) et, se basant sur les données $\boldsymbol{x}_n = \boldsymbol{x}_{\Lambda_n} \subset \Lambda_n$ générées par ce champ aléatoire $\mathbf{P}$, on veut estimer $\boldsymbol{h}$.

On construit un estimateur $\widehat{\boldsymbol{h}}_n$ comme un rapport de certaines fréquences conditionnelles empiriques et on démontre sa consistance exponentielle uniforme, c'est-à-dire

$$\sup_{\boldsymbol{h}\in\mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\| > \varepsilon\Big) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^{\nu}},$$

et sa consistance $\mathbf{L}^p$ uniforme pour tout $p \in (0,\infty)$, c'est-à-dire

$$\sup_{\boldsymbol{h}\in\mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \Big(\mathbf{E}\,\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p\Big)^{1/p} \leqslant n^{-(\nu/2-\sigma)},$$

où σ est une constante strictement positive arbitrairement petite, la norme considérée est la norme de la convergence uniforme, n est supposé être suffisamment grand, et les constantes $C, \alpha > 0$ sont déterminées par A, B et V.

Notons ici que dans [3], F. Comets obtient aussi la consistance exponentielle de l'estimateur du maximum de vraisemblance en utilisant la théorie des grandes déviations.

En général, le problème d'estimation pour les champs aléatoires de Gibbs est rendu difficile par des phénomènes classiques de la théorie des champs aléatoires de Gibbs tels que la non-unicité ($|\mathscr{G}| > 1$) et la perte de l'invariance par translation. Dans notre travail les résultats sont établis sans se soucier de ces aspects car ils sont valides uniformément sur $\mathscr{G}$, indépendamment du fait que $|\mathscr{G}| = 1$ ou non.

Ensuite on considère le problème nonparamétrique d'estimation des systèmes ponctuels dans le cas où ils sont quasi-locaux. On construit un estimateur en combinant les idées utilisées dans le cas paramétrique et l'idée principale de la méthode de "sieves" (introduit par U. Grenander dans [13]) qui consiste à approximer un paramètre infini-dimensionnel par des paramètres fini-dimensionnels. On démontre la consistance exponentielle et la consistance $\mathbf{L}^p$ de notre estimateur "sieve" dans des cadres différents.

Certains aspects sont similaires au travail de C. Ji [15]. En effet, nos systèmes ponctuels ressemblent effectivement aux "caractéristiques locales" et on étudie le même estimateur "sieve". Mais, contrairement à [15], on se situe dans un cadre beaucoup plus général et on estime l'objet même (système ponctuel) qui décrit le champ aléatoire.

Finalement notons ici que tous les résultats de cette deuxième partie sont valides dans le cas d'un espace d'états fini arbitraire. Notons aussi que certains résultats de cette partie ont été présentés dans [5].

Part I

Description of random fields

I. Auxiliary results from the theory of random fields

In this chapter we recall some well known notions and results from the theory of random fields, and particularly from Gibbs random fields theory. The exposition is based on the book of H.-O. Georgii [12]. We also set up in this chapter the notations that will be used in the sequel throughout this work.

I.1. Random fields, conditional probabilities

We consider random fields on the ν-dimensional integer lattice $\mathbb{Z}^\nu$, *i.e.*, probability measures on $(\Omega, \mathscr{F}) = \big(\mathscr{X}^{\mathbb{Z}^\nu}, \mathscr{F}_0^{\mathbb{Z}^\nu}\big)$ where $(\mathscr{X}, \mathscr{F}_0)$ is some *state space*, *i.e.*, space of values of a single variable. Usually the space $\mathscr{X}$ is assumed to be endowed with some topology $\mathscr{T}_0$, and $\mathscr{F}_0$ is assumed to be the Borel σ-algebra for this topology.

In this work we concentrate on the case when $\mathscr{X}$ is finite, $\mathscr{T}_0$ is the discrete topology (the topology consisting of all subsets of $\mathscr{X}$) and $\mathscr{F}_0$ is the total σ-algebra (the σ-algebra consisting of all subsets of $\mathscr{X}$), that is $\mathscr{F}_0 = \mathscr{T}_0 = \exp(\mathscr{X})$. Note that in this case $\mathscr{X}$ can also be considered as a metric space with $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise. Note also that in this case the state space is complete and compact, and hence $(\Omega, \mathscr{T}) = \big(\mathscr{X}^{\mathbb{Z}^\nu}, \mathscr{T}_0^{\mathbb{Z}^\nu}\big)$ is also complete, compact and metrizable. It seems that most of the results can be generalized to the case of infinite state space $\mathscr{X}$ under some additional topological assumptions like completeness, compactness, separability, *etc.*

A very important and the most interesting one is the $\{0,1\}$ case, that is, $\mathscr{X} = \{0,1\}$ and $\mathscr{F}_0 = \mathscr{T}_0 = \exp\big(\{0,1\}\big)$. In this case, each element $\boldsymbol{x} \in \mathscr{X}^\Lambda$ is uniquely determined by the subset X of Λ where the configuration $\boldsymbol{x}$ assumes the value 1 (in physical terminology this subset is occupied by particles). Therefore we can identify any configuration $\boldsymbol{x}$ on Λ with the corresponding subset X of Λ. In the sequel, when considering the $\{0,1\}$ case, we will not make difference between this two notions and will write, for example, $\boldsymbol{x} \subset \Lambda$ for a configuration $\boldsymbol{x}$ on Λ.

Denote by $\mathscr{E}$ the set of all finite subsets of $\mathbb{Z}^\nu$, *i.e.*, let $\mathscr{E} = \big\{\Lambda \subset \mathbb{Z}^\nu \,:\, |\Lambda| < \infty\big\}$ where $|\Lambda|$ is the number of points of the set Λ. Let us note that $\mathscr{E}$ is countable. Note also that by definition $\mathscr{F}$ is the smallest σ-algebra on Ω containing all the cylinder events

$$\big\{\boldsymbol{x} \in \Omega \,:\, \boldsymbol{x}_{\Lambda} \in A\big\}, \quad \Lambda \in \mathscr{E},\ A \in \mathscr{F}_0^\Lambda.$$

Here and in the sequel $\boldsymbol{x}_{\Lambda} = \{x_t,\ t \in \Lambda\}$ is the *subconfiguration* (*restriction*) on Λ of the configuration $\boldsymbol{x} = \{x_t,\ t \in \mathbb{Z}^\nu\}$. Note that in the $\{0,1\}$ case we can write this as $\boldsymbol{x}_{\Lambda} = \boldsymbol{x} \cap \Lambda$. In general, if $\boldsymbol{x} \in \mathscr{X}^K$ and $\Lambda \subset \mathbb{Z}^\nu$, then $\boldsymbol{x}_{\Lambda}$ is understood as a configuration $\{x_t,\ t \in K \cap \Lambda\}$ on $K \cap \Lambda$.

For any $\Lambda \in \mathscr{E} \setminus \emptyset$ let us consider the space $\mathscr{X}^\Lambda$ of all configurations on Λ. A probability distribution on $\mathscr{X}^\Lambda$ is denoted by $\mathbf{P}_\Lambda = \big\{\mathbf{P}_\Lambda(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{X}^\Lambda\big\}$. For convenience of notations we agree that for $\Lambda = \emptyset$ there exists only one probability distribution $\mathbf{P}_\emptyset(\boldsymbol{\emptyset}) = 1$ on the space $\mathscr{X}^\emptyset = \{\boldsymbol{\emptyset}\}$ where $\boldsymbol{\emptyset}$ is understood as a configuration consisting of absolutely nothing (the only possible configuration on the empty set).

For any $\Lambda \in \mathscr{E}$ and $I \subset \Lambda$ we denote

$$\big(\mathbf{P}_\Lambda\big)_I(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \mathscr{X}^{\Lambda \setminus I}} \mathbf{P}_\Lambda(\boldsymbol{x} \oplus \boldsymbol{y}), \quad \boldsymbol{x} \in \mathscr{X}^I. \tag{I.1}$$

The probability distribution $\big(\mathbf{P}_\Lambda\big)_I$ on $\mathscr{X}^I$ is the *restriction* of $\mathbf{P}_\Lambda$ on I. Here $\boldsymbol{x} \oplus \boldsymbol{y}$ is understood as a configuration on Λ equal to $\boldsymbol{x}$ on I and to $\boldsymbol{y}$ on $\Lambda \setminus I$. Note that for the $\{0,1\}$ case this corresponds to a usual set union, and so the formula (I.1) can be rewritten as

$$\big(\mathbf{P}_\Lambda\big)_I(\boldsymbol{x}) = \sum_{\boldsymbol{y} \subset \Lambda \setminus I} \mathbf{P}_\Lambda(\boldsymbol{x} \cup \boldsymbol{y}), \quad \boldsymbol{x} \subset I.$$

Definition I.1. — *A system of probability distributions* $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ *is called consistent in Kolmogorov's sense if for any* $\Lambda \in \mathscr{E}$ *and* $I \subset \Lambda$ *we have* $\big(\mathbf{P}_\Lambda\big)_I = \mathbf{P}_I$, *i.e.*, $\big(\mathbf{P}_\Lambda\big)_I(\boldsymbol{x}) = \mathbf{P}_I(\boldsymbol{x})$ *for all* $\boldsymbol{x} \in \mathscr{X}^I$.

It is well known that any system of probability distributions consistent in Kolmogorov's sense determines some probability measure on $(\Omega, \mathscr{F})$ (or, equivalently, some random field on $\mathbb{Z}^\nu$) for which it is the system of *finite-dimensional distributions*.

Before introducing the concept of conditional distribution of a random field, let us recall some combinatorial facts about nets (sequences) of real numbers indexed by elements of $\mathscr{E}$, as well as the notion of their convergence.

Let $\boldsymbol{b} = \{b_{_R},\ R \in \mathscr{E}\}$ be a net, *i.e.*, a real-valued function on $\mathscr{E}$, and let us define

$$a_{_\Lambda} = \sum_{R \subset \Lambda} b_{_R}\ , \qquad \Lambda \in \mathscr{E}. \tag{I.2}$$

Then one can express the function $\boldsymbol{b}$ in terms of the function $\boldsymbol{a} = \{a_{_\Lambda},\ \Lambda \in \mathscr{E}\}$, by "inversing" the formula (I.2) in the following way:

$$b_{_R} = \sum_{J \subset R} (-1)^{|R \setminus J|} a_{_J}\ , \qquad R \in \mathscr{E}. \tag{I.3}$$

The formula (I.3) is sometimes called *inclusion-exclusion formula* and sometimes *Möbius formula*.

In our opinion this formula is very important in description of random fields. Even if not used explicitly, it is implicitly present behind any approach. One can encounter this formula in many works devoted to description of random fields (see, for example, [2], [12], [17], [20], [24] and [25]). Our approach, presented in the following chapters, is heavily based on this formula.

Let us also remark, that an arbitrary real-valued function $\boldsymbol{a} = \{a_{_\Lambda},\ \Lambda \in \mathscr{E}\}$ on $\mathscr{E}$ can be represented in the form (I.2). For that, it is sufficient to define the function $\boldsymbol{b}$ by the formula (I.3). Note that the representation is unique. Note also, that this representation is noting but a generalisation to the case of nets of the formula

$$a_n = a_0 + (a_1 - a_0) + \cdots + (a_n - a_{n-1}),$$

permitting to represent an arbitrary sequence as a series.

Let us now introduce the notion of convergence of nets.

Definition I.2. — *Let $\{a_{_\Lambda},\ \Lambda \in \mathscr{E}\}$ be an arbitrary real-valued function on $\mathscr{E}$ and let $T \subset \mathbb{Z}^\nu$ be an infinite subset of $\mathbb{Z}^\nu$.*

1) *We say that $\lim\limits_{\Lambda \uparrow T} a_{_\Lambda} = a_{_T}$ if for any sequence $\Lambda_n \in \mathscr{E}$ such that $\Lambda_n \uparrow T$ we have the convergence $\lim\limits_{n \to \infty} a_{_{\Lambda_n}} = a_{_T}$.*

2) *As we have already mentioned, there exists some unique function $\{b_{_R},\ R \in \mathscr{E}\}$ such that $a_{_\Lambda} = \sum\limits_{R \subset \Lambda} b_{_R}$ for all $\Lambda \in \mathscr{E}$. We say that the convergence $\lim\limits_{\Lambda \uparrow T} a_{_\Lambda} = a_{_T}$ is "absolute" if the series $\sum\limits_{R \in \mathscr{E}\,:\, R \subset T} b_{_R}$ not only converges to $a_{_T}$ but is also absolutely convergent.*

Now we can finally introduce the concept of conditional distribution of a random field.

Let $\mathbf{P}$ be a random field. It is well known that for any $\Lambda \in \mathscr{E}$ there exist for $\mathbf{P}_{\Lambda^c}$-almost all $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}$ the following limits

$$\boldsymbol{q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \lim_{\widetilde{\Lambda} \uparrow \Lambda^c} \frac{\mathbf{P}_{\Lambda \cup \widetilde{\Lambda}}(\boldsymbol{x} \oplus \overline{\boldsymbol{x}}_{\widetilde{\Lambda}})}{\mathbf{P}_{\widetilde{\Lambda}}(\overline{\boldsymbol{x}}_{\widetilde{\Lambda}})}\ , \quad \boldsymbol{x} \in \mathscr{X}^\Lambda.$$

Any system

$$\boldsymbol{\mathcal{Q}} = \left\{ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c} \right\}$$

of probability distributions in various finite volumes Λ with various boundary conditions $\overline{\boldsymbol{x}}$ on Λ^c such that for all $\Lambda \in \mathscr{E}$ we have $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}} = \boldsymbol{q}_\Lambda^{\overline{\boldsymbol{x}}}$ for $\mathbf{P}_{\Lambda^c}$-almost all $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}$ is called *conditional distribution* of the random field $\mathbf{P}$. Note that if $\boldsymbol{\mathcal{Q}}$ is a conditional distribution of a random field $\mathbf{P}$ then in general, for a particular $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}$, the conditional distribution $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}$ in the volume Λ with boundary condition $\overline{\boldsymbol{x}}$ is not necessarily equal to $\boldsymbol{q}_\Lambda^{\overline{\boldsymbol{x}}}$ even if the last one is well-defined (*i.e.*, the corresponding limits exist).

It is also well known that any conditional distribution $\boldsymbol{\mathcal{Q}}$ of a random field $\mathbf{P}$ satisfies $\mathbf{P}$-almost surely the condition

$$\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}^{\overline{\boldsymbol{x}}}(\boldsymbol{x} \oplus \boldsymbol{y}) = \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}} \oplus \boldsymbol{y}}(\boldsymbol{x}) \left(\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}^{\overline{\boldsymbol{x}}} \right)_{\widetilde{\Lambda}}(\boldsymbol{y}) \tag{I.4}$$

where $\Lambda, \widetilde{\Lambda} \in \mathscr{E}$, $\Lambda \cap \widetilde{\Lambda} = \emptyset$, $\boldsymbol{x} \in \mathscr{X}^\Lambda$, $\boldsymbol{y} \in \mathscr{X}^{\widetilde{\Lambda}}$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\left(\Lambda \cup \widetilde{\Lambda}\right)^c}$. In fact, this is nothing but the elementary formula

$$\mathbf{P}(A \cap B \mid C) = \mathbf{P}(A \mid B \cap C)\ \mathbf{P}(B \mid C) \tag{I.5}$$

written for our case.

I.2. Specifications, Hamiltonians, potentials

Let us consider an arbitrary system

$$\boldsymbol{\mathcal{Q}} = \left\{ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c} \right\}$$

of probability distributions in finite volumes with boundary conditions. If we want this system to be a conditional distribution of some random field $\mathbf{P}$, then we need to suppose that it satisfies $\mathbf{P}$-almost surely the condition (I.4). However, we do not know *a priori* the random field $\mathbf{P}$. Therefore we need to require that the condition (I.4) holds always, rather than almost surely. This leads us to introduce the following

Definition I.3. — *A system*

$$\boldsymbol{\mathcal{Q}} = \Big\{ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}} \Big\}$$

of probability distributions in finite volumes with boundary conditions is called specification if for any $\Lambda, \widetilde{\Lambda} \in \mathscr{E}$ such that $\Lambda \cap \widetilde{\Lambda} = \emptyset$ and for any $\boldsymbol{x} \in \mathscr{X}^{\Lambda}$, $\boldsymbol{y} \in \mathscr{X}^{\widetilde{\Lambda}}$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{(\Lambda \cup \widetilde{\Lambda})^{\mathrm{c}}}$ we have

$$\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}^{\overline{\boldsymbol{x}}}(\boldsymbol{x} \oplus \boldsymbol{y}) = \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}} \oplus \boldsymbol{y}}(\boldsymbol{x}) \left(\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}^{\overline{\boldsymbol{x}}} \right)_{\widetilde{\Lambda}}(\boldsymbol{y}). \tag{I.6}$$

Sometimes such systems are also called systems of distributions in finite volumes with boundary conditions consistent in Dobrushin's sense.

In Gibbs random fields theory a random field is described through a specification $\boldsymbol{\mathcal{Q}} = \big\{ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}} \big\}$ wich is assumed to have the following Gibbsian form:

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \frac{\exp\big(-U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\big)}{\sum\limits_{\boldsymbol{y} \in \mathscr{X}^{\Lambda}} \exp\big(-U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{y})\big)}\ , \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \in \mathscr{X}^{\Lambda},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}},$$

where the system $\boldsymbol{\mathcal{U}} = \big\{ U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}), \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \in \mathscr{X}^{\Lambda},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}} \big\}$ is called *Hamiltonian*, $U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x})$ is called *(total) conditional energy of $\boldsymbol{x}$ in Λ under boundary condition $\overline{\boldsymbol{x}}$*, $\exp\big(U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\big)$ is called *Boltzmann factor*, the denominator is called *partition function*, and the Hamiltonian is assumed to be given by the formula

$$U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \ = \sum_{J\,:\,\emptyset \neq J \subset \Lambda} \ \sum_{\widetilde{J} \in \mathscr{E}\,:\,\widetilde{J} \subset \Lambda^{\mathrm{c}}} \Phi\big(\boldsymbol{x}_{J} \oplus \boldsymbol{x}_{\widetilde{J}}\big), \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \in \mathscr{X}^{\Lambda},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}},$$

where $\boldsymbol{\Phi} = \big\{ \Phi(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{X}^{J} \text{ for some } J \in \mathscr{E} \setminus \{\emptyset\} \big\}$ is some function taking values in $\mathbb{R} \cup \{+\infty\}$ (sometimes only real-valued functions are considered) called *interaction potential*. Here and in the sequel we admit that $\exp(-\infty) = 0$, $(+\infty) + (+\infty) = a + (+\infty) = (+\infty) + a = +\infty$ for all $a \in \mathbb{R}$ and that any sum over an empty space of indexes is equal to 0, *i.e.*, $U_\emptyset^{\overline{\boldsymbol{x}}}(\boldsymbol{\phi}) = 0$ for all $\overline{\boldsymbol{x}} \in \Omega$. Let us note that in general, if one lets the potential to take the value $+\infty$, the Gibbsian form is not well-defined, since the denominator in the definition of $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x})$ can be equal to 0 (say $U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{y}) = +\infty$ for all $\boldsymbol{y} \in \mathscr{X}^{\Lambda}$). So one needs to suppose the potential to be reasonable enough to avoid such situations. Clearly this situation does not occur if one considers a real-valued potential. Neither it occurs in the case of the so-called "*vacuum potentials*" which will be considered below. Note also that in general the system $\boldsymbol{\mathcal{U}}$ is not well-defined, since in the second sum

the summation is taken over an infinite space of indexes. For this reason the interaction potentials are always supposed to be such that the limits

$$U_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \lim_{\Delta \uparrow \mathbb{Z}^\nu} U_{\Lambda,\Delta}^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \tag{I.7}$$

exist and are in $\mathbb{R} \cup \{+\infty\}$ for all $\Lambda \in \mathscr{E}, \boldsymbol{x} \in \mathscr{X}^\Lambda$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}$. Here

$$U_{\Lambda,\Delta}^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \;=\; \sum_{J\,:\,\emptyset \neq J \subset \Lambda} \;\sum_{\widetilde{J} \subset \Delta \cap \Lambda^{\mathrm{c}}} \Phi\big(\boldsymbol{x}_J \oplus \boldsymbol{x}_{\widetilde{J}}\big), \quad \Lambda, \Delta \in \mathscr{E},\ \boldsymbol{x} \in \mathscr{X}^\Lambda,\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}.$$

Such interaction potentials are called *convergent*. Usually some stronger conditions on the interaction potential are supposed in order to guarantee that it is convergent. For example, often the interaction potential is supposed to be *absolutely summable*, *i.e.*, to satisfy the condition

$$\sum_{J\,:\,t \in J \in \mathscr{E}} \sup_{\boldsymbol{x} \in \mathscr{X}^J} \big|\Phi(\boldsymbol{x})\big| < \infty$$

for each $t \in \mathbb{Z}^\nu$. This condition not only implies that $\boldsymbol{\Phi}$ is convergent but, moreover, that it is *uniformly convergent*, *i.e.*, the limits (I.7) exist, are finite, and the convergence is uniform with respect to $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}$.

Interesting class of potentials is the class of *pair potentials*, *i.e.*, potentials $\boldsymbol{\Phi}$ such that $\Phi(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \in \mathscr{X}^J$ with $|J| \geqslant 2$. Note that the similar condition with $|J| \geqslant 1$ would imply the independence.

Another interesting class of potentials is the class of *finite range potentials*, *i.e.*, potentials $\boldsymbol{\Phi}$ such that $\Phi(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \in \mathscr{X}^J$ with $\operatorname{diam}(J) \geqslant d$ for some fixed $d \in \mathbb{N}$. Here and in the sequel $\operatorname{diam}(J)$ denotes the diameter of the set J in the metric ρ on $\mathbb{Z}^\nu$ defined by the norm

$$\Big\|\big(t^{(1)}, \cdots, t^{(\nu)}\big)\Big\| = \max\Big\{\big|t^{(1)}\big|, \ldots, \big|t^{(\nu)}\big|\Big\}, \quad \big(t^{(1)}, \cdots, t^{(\nu)}\big) \in \mathbb{Z}^\nu.$$

Note that finite range potentials are necessarily convergent, and that real-valued finite range potentials are absolutely summable.

The most simple class of potentials are the *nearest neighbour potentials*, *i.e.*, pair potentials $\boldsymbol{\Phi}$ such that $\Phi(\boldsymbol{x}) \neq 0$ only if $\boldsymbol{x}$ is a singleton, or $\boldsymbol{x} = \{s,t\}$ where s and t are *nearest neighbours*, that is they occupy two neighbour horizontal (or vertical) sites of the lattice.

Now, let us introduce the class of so-called "*vacuum potentials*".

Let us fix some element $\emptyset \in \mathscr{X}$ which will be called *vacuum* and let us denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$ (for the {0,1} case this element is usually 0).

Definition I.4. — *A potential* $\mathbf{\Phi} = \big\{\Phi(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{X}^J \text{ for some } J \in \mathscr{E} \setminus \{\emptyset\}\big\}$ *is called vacuum potential if we have* $\Phi(\boldsymbol{x}) = 0$ *for all* $\boldsymbol{x} \in \mathscr{X}^J$ *such that there exist some* $t \in J$ *satisfying* $x_t = \emptyset$.

The class of vacuum potentials plays very important role in Gibbs random fields theory for two reasons. Firstly, for an arbitrary potential one can find a unique vacuum potential giving the same specification as the initial one. Secondly, vacuum potentials are easier to manipulate. From here on we consider only vacuum potentials. In physical terminology $x_t = \emptyset$ means that this site is not occupied by any particle, while all other values represent different types of particles. In the vacuum case a configuration $\boldsymbol{x}$ on $\mathscr{X}^\Lambda$ is uniquely determined by its subconfiguration $\boldsymbol{y} \in \mathscr{X}^{*I}$ where the set $I \subset \Lambda$ is the set of sites occupied by particles, *i.e.*, $I = \{t \in \Lambda,\ x_t \neq \emptyset\}$. In the sequel we will not make difference between this two notions and will write, for example, $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \subset \Lambda$ for a configuration $\boldsymbol{x}$ on Λ. Note that in $\{0,1\}$ case there exists just one type of particles, and hence we have just a set, as we have already seen earlier. Now we can rewrite all the above formulas in these notations. The Gibbsian form is given by the formula

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \frac{\exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\big)}{\sum\limits_{\boldsymbol{y} \in \mathscr{X}^\Lambda} \exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{y})\big)}\ , \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \in \mathscr{X}^{*I},\ I \subset \Lambda,\ \overline{\boldsymbol{x}} \in \mathscr{X}^{*K},\ K \subset \Lambda^{\mathrm{c}},$$

and the Hamiltonian $\boldsymbol{\mathcal{U}} = \big\{U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \mathscr{E},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{*K},\ K \subset I^{\mathrm{c}}\big\}$ is given by the formula

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \ = \sum_{J\,:\,\emptyset \neq J \subset I} \ \sum_{\widetilde{J} \in \mathscr{E}\,:\,\widetilde{J} \subset K} \Phi\big(\boldsymbol{x}_J \oplus \boldsymbol{x}_{\widetilde{J}}\big)$$

where $\mathbf{\Phi} = \big\{\Phi(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{X}^{*J} \text{ for some } J \in \mathscr{E} \setminus \{\emptyset\}\big\}$ is the potential. Note that the Hamiltonian no longer depends on Λ. In fact, condition of vacuumness implies that for an arbitrary $\Lambda \in \mathscr{E}$ satisfying $I \subset \Lambda \subset K^{\mathrm{c}}$ we get the same value of Hamiltonian. The relation (I.7) can be rewritten as

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \lim_{\Delta \uparrow \mathbb{Z}^\nu} U^{\overline{\boldsymbol{x}}_\Delta}(\boldsymbol{x})$$

and the condition of absolute summability as

$$\sum_{J\,:\,t \in J \in \mathscr{E}} \ \sup_{\boldsymbol{x} \in \mathscr{X}^{*J}} \big|\Phi(\boldsymbol{x})\big| < \infty.$$

In $\{0,1\}$ case the notations are even more simple. The Gibbsian form is given by the formula

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \frac{\exp\bigl(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\bigr)}{\sum\limits_{\boldsymbol{y}\subset\Lambda}\exp\bigl(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{y})\bigr)}\ , \quad \Lambda\in\mathscr{E},\ \boldsymbol{x}\subset\Lambda,\ \overline{\boldsymbol{x}}\subset\Lambda^{\mathrm{c}},$$

and the Hamiltonian $\boldsymbol{\mathcal{U}} = \bigl\{U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}),\quad \boldsymbol{x}\in\mathscr{E} \text{ and } \overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}\bigr\}$ is given by the formula

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \;=\; \sum_{J\,:\,\emptyset\neq J\subset\boldsymbol{x}}\ \sum_{\widetilde{J}\in\mathscr{E}\,:\,\widetilde{J}\subset\overline{\boldsymbol{x}}} \Phi\bigl(J\cup\widetilde{J}\bigr)$$

where $\boldsymbol{\Phi} = \bigl\{\Phi(J),\ J\in\mathscr{E}\setminus\{\emptyset\}\bigr\}$ is the potential. The condition of absolute summability can be rewritten as

$$\sum_{J\,:\,t\in J\in\mathscr{E}}\bigl|\Phi(J)\bigr| < \infty.$$

Let us finally note here that in the vacuum case we clearly have $U^{\overline{\boldsymbol{x}}}(\boldsymbol{\emptyset}) = 0$ for all $\overline{\boldsymbol{x}}\in\Omega$, and hence we have $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{\emptyset}) > 0$ for all $\Lambda\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\in\mathscr{X}^{\Lambda^{\mathrm{c}}}$. Here $\boldsymbol{\emptyset}$ is nothing but the configuration $\emptyset^\Lambda$ identically equal to $\emptyset$ on Λ.

This leads us to introduce the notion of a general "*vacuum specification*".

Definition I.5. — *A system*

$$\boldsymbol{\mathcal{Q}} = \Bigl\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}},\quad \Lambda\in\mathscr{E} \text{ and } \overline{\boldsymbol{x}}\in\mathscr{X}^{\Lambda^{\mathrm{c}}}\Bigr\}$$

of probability distributions in finite volumes with boundary conditions is called vacuum specification if for all $\Lambda\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\in\mathscr{X}^{\Lambda^{\mathrm{c}}}$ we have $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{\emptyset}) > 0$ and if it satisfies the condition (I.6). Sometimes vacuum specifications are also called weakly positive specifications.

Note that for this case the condition (I.6) can be rewritten in an equivalent form

$$\mathbf{Q}_{\Lambda\cup\widetilde{\Lambda}}^{\overline{\boldsymbol{x}}}(\boldsymbol{x}\oplus\boldsymbol{y}) = \frac{\mathbf{Q}_{\Lambda\cup\widetilde{\Lambda}}^{\overline{\boldsymbol{x}}}(\boldsymbol{y})}{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}\oplus\boldsymbol{y}}(\boldsymbol{\emptyset})}\ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}\oplus\boldsymbol{y}}(\boldsymbol{x}). \tag{I.8}$$

Note also that in the $\{0,1\}$ case the condition of vacuumness is just $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\emptyset) > 0$ for all $\Lambda\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\subset\Lambda^{\mathrm{c}}$.

I.3. Description of random fields by their conditional probabilities

The main question of the Gibbs random field theory is the study (under different conditions on the potential) of the set of all random fields having a given Gibbsian specification $\boldsymbol{\mathcal{Q}}$ as a conditional distribution. Is this set empty or not? If it is not empty, is it a singleton or not, *i.e.*, is the field having $\boldsymbol{\mathcal{Q}}$ as a conditional distribution unique or not? In the non-uniqueness case, what can be said about the structure of this set? Another interesting question is the following. Suppose that $\boldsymbol{\Phi}$ (and hence $\boldsymbol{\mathcal{Q}}$) is translation invariant (*i.e.*, invariant with respect to shift operators on $\mathbb{Z}^\nu$ or, in other words, stationary). Are all the random fields having $\boldsymbol{\mathcal{Q}}$ as a conditional distribution translation invariant or not? In the latter case what can be said about the set of translation invariant random fields having $\boldsymbol{\mathcal{Q}}$ as a conditional distribution?

Below, we will state a theorem answering these questions in a more general setup, when the specification $\boldsymbol{\mathcal{Q}}$ is not supposed to have Gibbsian form, but rather is supposed to be "*quasilocal*". To state this theorem we need to introduce some definitions and notations.

We start by giving the following

Definition I.6. — *Let* $\boldsymbol{g} = \left\{g^{\boldsymbol{x}}, \quad \boldsymbol{x} \in \mathscr{X}^{*K} \text{ for some } K \subset \mathbb{Z}^\nu\right\}$ *be an arbitrary real-valued function on* $(\Omega, \mathscr{T})$.

1) *We say that the function* $\boldsymbol{g}$ *is local if it is* $\mathscr{F}_0^\Lambda$ *measurable for some* $\Lambda \in \mathscr{E}$, *i.e., if it depends only on the restriction* $\boldsymbol{x}_\Lambda$ *of* $\boldsymbol{x}$ *on* Λ *or, equivalently, if we have* $g^{\boldsymbol{x}} = g^{\boldsymbol{x}_\Lambda}$ *for all* $\boldsymbol{x} \in \Omega$.

2) *We say that the function* $\boldsymbol{g}$ *is quasilocal if it satisfies one of the following four equivalent conditions:*

(**q.l.1**) *the function* $\boldsymbol{g}$ *is continuous with respect to the topology* $\mathscr{T}$,

(**q.l.2**) *the function* $\boldsymbol{g}$ *is a uniform limit of local functions,*

(**q.l.3**) *we have* $\lim\limits_{I \uparrow \mathbb{Z}^\nu} g^{\boldsymbol{x}_I} = g^{\boldsymbol{x}}$ *uniformly on* $\boldsymbol{x} \in \Omega$, *i.e.,*

$$\sup_{\boldsymbol{x} \in \Omega} \left| g^{\boldsymbol{x}_I} - g^{\boldsymbol{x}} \right| \underset{I \uparrow \mathbb{Z}^\nu}{\longrightarrow} 0,$$

(**q.l.4**) *we have*

$$\sup_{\boldsymbol{x}, \boldsymbol{y} \in \Omega \,:\, \boldsymbol{x}_I = \boldsymbol{y}_I} \left| g^{\boldsymbol{x}} - g^{\boldsymbol{y}} \right| \underset{I \uparrow \mathbb{Z}^\nu}{\longrightarrow} 0.$$

The equivalence of these four conditions is well known and easily follows from the compactness of the space $(\Omega, \mathscr{T})$. Note that quasilocal functions are bounded functions, since they are continuous functions on a compact. Note also that local functions are clearly quasilocal.

Definition I.7. — *A specification* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}\right\}$ *is called (quasi)local if for all* $\Lambda \in \mathscr{E}$ *and* $\boldsymbol{x} \in \mathscr{X}^\Lambda$ *the function* $\left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{z}}_{\Lambda^c}}(\boldsymbol{x}), \quad \overline{\boldsymbol{z}} \in \Omega\right\}$ *is (quasi)local, i.e., if for all* $\Lambda \in \mathscr{E}$ *and* $\boldsymbol{x} \in \mathscr{X}^\Lambda$ *the quantity*

$$\varphi_{\boldsymbol{x},\Lambda}(I) = \sup_{\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}} \left|\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}_I}(\boldsymbol{x}) - \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\right|$$

tends to 0 *as* $I \uparrow \mathbb{Z}^\nu$ *(for the quasilocal case) or equals to* 0 *for* I *sufficiently large (for local case). A random field* $\mathbf{P}$ *is called (quasi)local if it has a (quasi)local conditional distribution.*

Note that the quasilocality is obviously true, for example, for Gibbsian specifications with uniformly convergent interaction potentials, and the locality, for Gibbsian specifications with finite range interaction potentials.

Now let us introduce the following convergence in the space $\mathscr{P}$ of all random fields defined on $\mathbb{Z}^\nu$ and taking values in the state space $\mathscr{X}$. We will say that a sequence $\mathbf{P}^{(n)}$ of random fields converges to some random field $\mathbf{P}$ if for all $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \in \mathscr{X}^\Lambda$ we have $\lim\limits_{n\to\infty} \mathbf{P}_\Lambda^{(n)}(\boldsymbol{x}) = \mathbf{P}_\Lambda(\boldsymbol{x})$. Note that we obtain this convergence if we consider the space $\mathscr{P}$ as a subset of the Banach space of all bounded functions $\boldsymbol{r} = \left\{r_\Lambda(\boldsymbol{x}), \quad \Lambda \in \mathscr{E} \text{ and } \boldsymbol{x} \in \mathscr{X}^\Lambda\right\}$ with the norm

$$\|\boldsymbol{r}\| = \sup_{\Lambda \in \mathscr{E}} \frac{1}{2n(\Lambda)} \sum_{\boldsymbol{x} \in \mathscr{X}^\Lambda} \left|r_\Lambda(\boldsymbol{x})\right|$$

where $n(\Lambda)$ is some enumeration of elements of $\mathscr{E}$ (*i.e.*, n is an arbitrary bijection from $\mathscr{E}$ on $\mathbb{N}$). Note also that the space $\mathscr{P}$ is a closed convex subset of this Banach space and, moreover, can be shown to be a compact set by usual "diagonal method".

A random field $\mathbf{P} \in \mathscr{P}$ is called *tail-trivial* if it is trivial on the *tail* σ*-algebra* $\mathscr{F}_\infty = \bigcap\limits_{\Lambda \in \mathscr{E}} \mathscr{F}_0^{\Lambda^c}$, *i.e.*, for all $A \in \mathscr{T}$ we have $\mathbf{P}(A) = 1$ or $\mathbf{P}(A) = 0$.

A random field $\mathbf{P} \in \mathscr{P}$ is called *translation invariant* if for all $\Lambda \in \mathscr{E}$, $\boldsymbol{x} \in \mathscr{X}^\Lambda$ and $t \in \mathbb{Z}^\nu$ we have $\mathbf{P}_\Lambda(\boldsymbol{x}) = \mathbf{P}_{\Lambda+t}(\boldsymbol{x} + t)$. Here and in the sequel $\Lambda + t$ denotes

the set $\{s+t\,:\,s\in\Lambda\}$ and $\boldsymbol{x}+t$ denotes the configuration $\boldsymbol{y}\in\mathscr{X}^{\Lambda+t}$ defined by $y_{s+t}=x_s$ for all $s\in\Lambda$. Similarly a specification $\boldsymbol{\mathcal{Q}}$ is called *translation invariant* if for all $\Lambda\in\mathscr{E}$, $\boldsymbol{x}\in\mathscr{X}^{\Lambda}$ and $\overline{\boldsymbol{x}}\in\mathscr{X}^{\Lambda^{\mathrm{c}}}$ we have $\mathbf{Q}_{\Lambda}^{\overline{\boldsymbol{x}}}(\boldsymbol{x})=\mathbf{Q}_{\Lambda+t}^{\overline{\boldsymbol{x}}+t}(\boldsymbol{x}+t)$.

A random field $\mathbf{P}\in\mathscr{P}$ is called *ergodic* if it is translation invariant and is trivial on the σ-algebra $\mathscr{I}=\big\{A\in\mathscr{F}\,:\,A+t=A\text{ for all }t\in\mathbb{Z}^{\nu}\big\}$ of all translation invariant events. Here $A+t=\{\boldsymbol{x}+t\,:\,\boldsymbol{x}\in A\}$. Let us note here that if $\mathbf{P}\in\mathscr{P}$ is translation invariant and tail-trivial, then it is also ergodic.

Let us now recall some notions from convex analysis. Let A be a convex subset of some real vector space. An element $\alpha\in A$ is said to be *extreme* (in A) if $\alpha\neq s\,\beta+(1-s)\,\gamma$ for all $0<s<1$ and all $\beta,\gamma\in A$ with $\beta\neq\gamma$. The set of all extreme elements of A is called *extreme boundary* of A and is denoted by $\operatorname{ex}A$. The convex set A is said to be a *simplex* if any element $\alpha\in A$ can be represented as

$$\alpha=\int\limits_{\operatorname{ex}A}\beta\;\mu_{\alpha}(\mathrm{d}\beta)$$

with the unique weight μ_{α} which is a probability distribution on the space $\operatorname{ex}A$. Recall also that for any set B the minimal convex set A containing B is called *convex hull* of B and that the closure of A is called *closed convex hull* of B and is denoted by c.c.h.(B).

Now, suppose we are given some fixed specification $\boldsymbol{\mathcal{Q}}$.

For each $\Lambda\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\in\mathscr{X}^{\Lambda^{\mathrm{c}}}$ let us consider a random field defined by $\mathbf{Q}_{\Lambda}^{\overline{\boldsymbol{x}}}$ on Λ and equal a.s. to $\overline{\boldsymbol{x}}$ outside Λ. This random field is called *random field in finite volume* Λ *with boundary condition* $\overline{\boldsymbol{x}}$.

Further, if for some sequence $\Lambda_n\in\mathscr{E}$ of finite volumes such that $\Lambda_n\uparrow\mathbb{Z}^{\nu}$ and some sequence $\overline{\boldsymbol{x}}_n\in\mathscr{X}^{\Lambda_n^{\mathrm{c}}}$ of boundary conditions these random fields converge to some random field $\mathbf{P}$, then this random field $\mathbf{P}$ is called *limiting Gibbs random field for random fields in finite volumes* (or shortly *limiting Gibbs random field*) for $\boldsymbol{\mathcal{Q}}$. We denote the set of all limiting Gibbs random fields for $\boldsymbol{\mathcal{Q}}$ by $\mathscr{G}_{\mathrm{lim}}=\mathscr{G}_{\mathrm{lim}}(\boldsymbol{\mathcal{Q}})$.

On the other hand any random field $\mathbf{P}$ having the specification $\boldsymbol{\mathcal{Q}}$ as a conditional distribution is called *Gibbs random field* for $\boldsymbol{\mathcal{Q}}$. We denote the set of all Gibbs random fields for $\boldsymbol{\mathcal{Q}}$ by $\mathscr{G}=\mathscr{G}(\boldsymbol{\mathcal{Q}})$.

In the case when $\boldsymbol{\mathcal{Q}}$ is translation invariant we also denote by $\mathscr{G}_{\mathrm{t.i.}}=\mathscr{G}_{\mathrm{t.i.}}(\boldsymbol{\mathcal{Q}})$ the set of all translation invariant Gibbs random fields for $\boldsymbol{\mathcal{Q}}$.

Note that above we use the traditional term "Gibbs" even though $\boldsymbol{\mathcal{Q}}$ is not necessarily Gibbsian.

Now we can finally state the following

Theorem I.8. — *Let the specification* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}\right\}$ *be quasilocal.*

1) *The set* $\mathscr{G}$ *is a non-empty closed convex set. Moreover,* $\mathscr{G}$ *is a simplex and we have* $\mathrm{ex}\mathscr{G} \subset \mathscr{G}_{\mathrm{lim}}$ *and* $\mathscr{G} = \mathrm{c.c.h.}(\mathscr{G}_{\mathrm{lim}}) = \mathrm{c.c.h.}(\mathrm{ex}\mathscr{G})$*. Finally, a random field* $\mathbf{P} \in \mathscr{G}$ *is extreme* $\left(i.e., \mathbf{P} \in \mathrm{ex}\mathscr{G}\right)$ *if and only if* $\mathbf{P}$ *is tail-trivial.*

2) *If* $\boldsymbol{\mathcal{Q}}$ *is translation invariant then* $\mathscr{G}_{\mathrm{t.i.}} \subset \mathscr{G}$ *is also a non-empty closed convex set. Moreover,* $\mathscr{G}_{\mathrm{t.i.}}$ *is a simplex and we have* $\mathscr{G}_{\mathrm{t.i.}} = \mathrm{c.c.h.}(\mathrm{ex}\mathscr{G}_{\mathrm{t.i.}})$*. Finally, a random field* $\mathbf{P} \in \mathscr{G}_{\mathrm{t.i.}}$ *is extreme* $\left(i.e., \mathbf{P} \in \mathrm{ex}\mathscr{G}_{\mathrm{t.i.}}\right)$ *if and only if* $\mathbf{P}$ *is ergodic.*

3) *The set* $\mathscr{G}$ *is a singleton, i.e.,* $\mathscr{G} = \{\mathbf{P}\}$*, if and only if for any increasing sequence of finite volumes and for any sequence of corresponding boundary conditions the random fields in these finite volumes with these boundary conditions converge to the random field* $\mathbf{P}$*.*

4) *Suppose* $\boldsymbol{\mathcal{Q}}_1$ *and* $\boldsymbol{\mathcal{Q}}_2$ *are Gibbsian specifications corresponding to some uniformly convergent vacuum potentials* $\boldsymbol{\Phi}_1$ *and* $\boldsymbol{\Phi}_2$ *(and hence are quasilocal). Then* $\mathscr{G}(\boldsymbol{\mathcal{Q}}_1) \cap \mathscr{G}(\boldsymbol{\mathcal{Q}}_2) \neq \emptyset \iff \boldsymbol{\Phi}_1 = \boldsymbol{\Phi}_2 \iff \boldsymbol{\mathcal{Q}}_1 = \boldsymbol{\mathcal{Q}}_2 \iff \mathscr{G}(\boldsymbol{\mathcal{Q}}_1) = \mathscr{G}(\boldsymbol{\mathcal{Q}}_2)$.

Remark I.9. — Non-uniqueness and translation invariance breaking are possible. Non-uniqueness means that it is possible to have $|\mathscr{G}| \neq 1$ and even $|\mathscr{G}_{\mathrm{t.i.}}| \neq 1$. Translation invariance breaking means that it is possible to have (in the non-uniqueness case) $\mathscr{G}_{\mathrm{t.i.}} \neq \mathscr{G}$. Moreover, it is possible to have $\mathrm{ex}\mathscr{G}_{\mathrm{t.i.}} \setminus \mathrm{ex}\mathscr{G} \neq \emptyset$ and $\mathrm{ex}\mathscr{G} \setminus \mathscr{G}_{\mathrm{t.i.}} \neq \emptyset$, *i.e.*, the simplex $\mathscr{G}_{\mathrm{t.i.}}$ is not necessarily a face (subsimplex) of the simplex $\mathscr{G}$.

Finally, to conclude this chapter let us give here a sufficient condition for uniqueness of the Gibbs random field for a given quasilocal specification $\boldsymbol{\mathcal{Q}}$. For the convenience of notations in the sequel we will often write t for the set $\{t\}$ consisting of just one point t.

Let us introduce the following

Definition I.10. — *Let* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}},\quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}\right\}$ *be some specification. We say that it satisfies Dobrushin's uniqueness condition if it is quasilocal and we have*

$$\frac{1}{2}\sup_{t\in\mathbb{Z}^\nu}\sum_{s\in\mathbb{Z}^\nu\setminus t}\sup_{\overline{\boldsymbol{x}},\overline{\boldsymbol{y}}}\sum_{x\in\mathscr{X}}\left|\mathbf{Q}_t^{\overline{\boldsymbol{x}}}(x)-\mathbf{Q}_t^{\overline{\boldsymbol{y}}}(x)\right|<1. \tag{I.9}$$

where the second sup *is taken over all pairs* $\overline{\boldsymbol{x}},\overline{\boldsymbol{y}} \in \mathscr{X}^{\mathbb{Z}^\nu\setminus t}$ *such that we have* $\overline{\boldsymbol{x}}_{\mathbb{Z}^\nu\setminus\{s,t\}} = \overline{\boldsymbol{y}}_{\mathbb{Z}^\nu\setminus\{s,t\}}$.

Now we can finally state Dobrushin's uniqueness theorem.

Theorem I.11. — *Let the specification* $\boldsymbol{\mathcal{Q}}$ *satisfy Dobrushin's uniqueness condition. Then* $\mathscr{G}$ *is a singleton, that is we have* $|\mathscr{G}| = 1$. *If we suppose also that* $\boldsymbol{\mathcal{Q}}$ *is translation invariant then* $\mathscr{G}_{\text{t.i.}} = \mathscr{G}$ *is also a singleton.*

These results are synthesis of several theorems from [12]. Note that the main part of the Theorems I.8 and I.11 was first formulated by R.L. Dobrushin in [8] — [10] for Gibbsian case. Note also that the Theorems I.8 and I.11 hold in the case of a finite state space $\mathscr{X}$. The case of infinite state space requires more notations and assumptions. Details for this case can be found in [12].

II. Random fields and P-functions

In this chapter we propose an approach towards description of random fields which is based on a notion of *P-functions*. This notion is a generalization of a notion of infinite-volume correlation functions well known in Gibbs random fields theory. First two sections are devoted to the {0,1} case. The third section shows the way one can generalize these results to the case of arbitrary finite state space $\mathscr{X}$.

II.1. Description of random fields by P-functions

Here we propose an approach towards description of random fields in the {0,1} case. In the proposed approach the classical system of probability distributions consistent in Kolmogorov's sense is replaced by some function on $\mathscr{E}$ (P-function) and the Kolmogorov's consistency condition is replaced by some "non-negativity" condition imposed on certain finite sums with alternating signs of summands.

Definition II.1. — *A real-valued function $\boldsymbol{f} = \{f_J,\ J \in \mathscr{E}\}$ on $\mathscr{E}$ is called P-function if $f_\emptyset = 1$ and for any $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$ we have*

$$\sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} f_{\Lambda \setminus J} \geqslant 0. \tag{II.1}$$

Theorem II.2. — *A system $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ is a system of probability distributions consistent in Kolmogorov's sense if and only if there exists a P-function $\boldsymbol{f}$ such that for any $\Lambda \in \mathscr{E}$ we have*

$$\mathbf{P}_\Lambda(\boldsymbol{x}) = \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} f_{\Lambda \setminus J}, \quad \boldsymbol{x} \subset \Lambda. \tag{II.2}$$

Particularly, for any $\Lambda \in \mathscr{E}$ we have $\mathbf{P}_\Lambda(\emptyset) = f_\Lambda$.

Proof: **1)** Necessity. Let $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ be a system of probability distributions consistent in Kolmogorov's sense. Put $f_\Lambda = \mathbf{P}_\Lambda(\emptyset)$ for all $\Lambda \in \mathscr{E}$. Clearly $f_\emptyset = \mathbf{P}_\emptyset(\emptyset) = 1$. Further we have

$$\begin{aligned}\sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|} f_{\Lambda\setminus J} &= \sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}\,\mathbf{P}_{\Lambda\setminus J}(\emptyset) = \\ &= \sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}\big(\mathbf{P}_\Lambda\big)_{\Lambda\setminus J}(\emptyset) = \\ &= \sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}\sum_{\widetilde{J}\subset J}\mathbf{P}_\Lambda\big(\widetilde{J}\big) = \\ &= \sum_{\widetilde{J}\subset \boldsymbol{x}}\mathbf{P}_\Lambda\big(\widetilde{J}\big)\sum_{J\,:\,\widetilde{J}\subset J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|} = \mathbf{P}_\Lambda(\boldsymbol{x}) \geqslant 0.\end{aligned}$$

The last equality holds due to the following combinatorial relation

$$\sum_{A\,:\,B\subset A\subset C}(-1)^{|C\setminus A|} = \sum_{A\,:\,B\subset A\subset C}(-1)^{|A\setminus B|} = \begin{cases}1 & \text{if } B = C,\\ 0 & \text{if } B \neq C.\end{cases} \tag{II.3}$$

2) Sufficiency. Let $\boldsymbol{f}$ be a P-function. For any $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$ let us put

$$\mathbf{P}_\Lambda(\boldsymbol{x}) = \sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|} f_{\Lambda\setminus J} \geqslant 0$$

and show that $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ is a system of probability distributions consistent in Kolmogorov's sense. For any $\Lambda \in \mathscr{E}$ we have

$$\sum_{\boldsymbol{x}\subset\Lambda}\mathbf{P}_\Lambda(\boldsymbol{x}) = \sum_{\boldsymbol{x}\subset\Lambda}\sum_{J\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|} f_{\Lambda\setminus J} = \sum_{J\subset\Lambda} f_{\Lambda\setminus J}\sum_{\boldsymbol{x}\,:\,J\subset \boldsymbol{x}\subset\Lambda}(-1)^{|\boldsymbol{x}\setminus J|} = f_\emptyset = 1,$$

i.e., $\mathbf{P}$ is a system of probability distributions. Now let us verify its consistency. For any $\Lambda \in \mathscr{E}$, $I \subset \Lambda$ and $\boldsymbol{x} \subset I$ we can write

$$\begin{aligned}\big(\mathbf{P}_\Lambda\big)_I(\boldsymbol{x}) &= \sum_{J\subset\Lambda\setminus I}\ \sum_{\widetilde{J}\subset \boldsymbol{x}\cup J}(-1)^{\left|(\boldsymbol{x}\cup J)\setminus\widetilde{J}\right|} f_{\Lambda\setminus\widetilde{J}} = \\ &= \sum_{J\subset\Lambda\setminus I}\ \sum_{\widetilde{J}_1\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus\widetilde{J}_1|}\sum_{\widetilde{J}_2\subset J}(-1)^{\left|J\setminus\widetilde{J}_2\right|} f_{\Lambda\setminus(\widetilde{J}_1\cup\widetilde{J}_2)} = \\ &= \sum_{\widetilde{J}_1\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus\widetilde{J}_1|}\sum_{\widetilde{J}_2\subset\Lambda\setminus I} f_{\Lambda\setminus(\widetilde{J}_1\cup\widetilde{J}_2)}\sum_{J\,:\,\widetilde{J}_2\subset J\subset\Lambda\setminus I}(-1)^{\left|J\setminus\widetilde{J}_2\right|} = \\ &= \sum_{\widetilde{J}_1\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus\widetilde{J}_1|} f_{I\setminus\widetilde{J}_1} = \mathbf{P}_I(\boldsymbol{x}).\end{aligned}$$

The theorem is proved. □

II.2. Properties and examples of P-functions

Let $\mathscr{B}$ be the Banach space of all bounded functions defined on $\mathscr{E}$ with the norm

$$\|\boldsymbol{b}\| = \sup_{J\in\mathscr{E}} \frac{|b_J|}{n(J)}\ , \quad \boldsymbol{b} = \{b_J,\ J\in\mathscr{E}\} \in \mathscr{B},$$

where $n(J)$ is some enumeration of elements of $\mathscr{E}$ and let $\mathscr{B}\big([0,1]\big)$ be the subset of $\mathscr{B}$ consisting of all functions taking values in $[0,1]$. Note that $\mathscr{B}\big([0,1]\big)$ is a closed convex subset of $\mathscr{B}$ and that the convergence of functions in $\mathscr{B}\big([0,1]\big)$ is equivalent to the "pointwise" convergence, *i.e.*, to the convergence for any $J\in\mathscr{E}$.

Proposition II.3 [Properties of P-functions]**.** — **1)** *The space $\mathscr{B}^P$ of all P-functions is a closed convex subset of $\mathscr{B}\big([0,1]\big)$. Moreover $\mathscr{B}^P$ is compact.*

2) *Let $\boldsymbol{f}$ be a P-function and fix some $T\subset\mathbb{Z}^\nu$. Then the function $\boldsymbol{f}^{|T}$ defined by $f^{|T}_J = f_{T\cap J},\ J\in\mathscr{E}$, is also a P-function. The corresponding random field is the restriction of the original one on T and assumes a.s. the value 0 outside T.*

3) *Let $\boldsymbol{f}$ be a P-function. For any fixed $B\in\mathscr{E}$ such that $f_B>0$ consider the function $\boldsymbol{f}^B$ defined by $f^B_J = \frac{f_{B\cup J}}{f_B}\ ,\ J\in\mathscr{E}$. Then $\boldsymbol{f}^B$ is also a P-function. The corresponding random field is the original one conditioned to be equal 0 on B (and hence assuming a.s. the value 0 on B).*

4) *Consider a family $\mathscr{F} = \big\{\boldsymbol{f}^{(s)}\big\}$ of P-functions depending on the parameter $s\in(0,1)$ and let $p(s)$, $s\in(0,1)$, be a probability density. Then the function $\boldsymbol{g}$ defined by*

$$g_J = \int_0^1 f^{(s)}_J p(s)\,\mathrm{d}s, \quad J\in\mathscr{E},$$

is also a P-function. Corresponding random field is a mixture of the original ones.

5) *Consider a P-function $\boldsymbol{f}$ and let $\varphi:\mathbb{Z}^\nu\longrightarrow T\subset\mathbb{Z}^\nu$ be a bijection. Then the function $\boldsymbol{f}^\varphi$ defined by $f^\varphi_J = f_{\varphi(J)},\ J\in\mathscr{E}$, is also a P-function. The corresponding random field can be viewed as the image of the original one by φ^{-1}, or rather by $\widetilde{\varphi}:\mathscr{X}^{\mathbb{Z}^\nu}\longrightarrow\mathscr{X}^{\mathbb{Z}^\nu}$ corresponding to each $\boldsymbol{x}\in\mathscr{X}^{\mathbb{Z}^\nu}$ a configuration $\widetilde{\varphi}(\boldsymbol{x})\in\mathscr{X}^{\mathbb{Z}^\nu}$ defined by $\widetilde{\varphi}(\boldsymbol{x})_t = \boldsymbol{x}_{\varphi(t)},\quad t\in\mathbb{Z}^\nu$.*

Proof: **1)** The first assertion is evident. The compactness can be easily proved using the usual "diagonal method".

2) and 3) Both this two assertions can be proved by considering the corresponding random field, calculating in it the probabilities of empty configurations and using the Theorem II.2. Note that we can also check directly the conditions of the Definition II.1 using combinatorial formulas. For example, let us check these conditions for the case of 2). We have obviously $f_\emptyset^{|_T} = f_{T\cap\emptyset} = f_\emptyset = 1$. Further we have

$$\begin{aligned}\sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_{\Lambda\setminus J}^{|_T} &= \sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_{T\cap(\Lambda\setminus J)} = \\ &= \sum_{J_1\subset\boldsymbol{x}\cap T}\ \sum_{J_2\subset\boldsymbol{x}\cap T^{\mathrm{c}}}(-1)^{\left|(\boldsymbol{x}\cap T)\setminus J_1\right|}\,(-1)^{\left|(\boldsymbol{x}\cap T^{\mathrm{c}})\setminus J_2\right|}f_{(T\cap\Lambda)\setminus J_1} = \\ &= \sum_{J_1\subset\boldsymbol{x}\cap T}(-1)^{\left|(\boldsymbol{x}\cap T)\setminus J_1\right|}f_{(T\cap\Lambda)\setminus J_1}\ \times\sum_{J_2\subset\boldsymbol{x}\cap T^{\mathrm{c}}}(-1)^{\left|(\boldsymbol{x}\cap T^{\mathrm{c}})\setminus J_2\right|}\geqslant 0\end{aligned}$$

because the first factor is positive by (II.1) and the second one by (II.3). Here and in the sequel $T^{\mathrm{c}} = \mathbb{Z}^\nu \setminus T$ denotes the complement of T.

4) On one hand we have

$$g_\emptyset = \int_0^1 f_\emptyset^{(s)}p(s)\,\mathrm{d}s = \int_0^1 p(s)\,\mathrm{d}s = 1.$$

On the other hand we can write

$$\begin{aligned}\sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}g_J &= \sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}\int_0^1 f_J^{(s)}p(s)\,\mathrm{d}s = \\ &= \int_0^1\left(\sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_J^{(s)}\right)p(s)\,\mathrm{d}s\geqslant 0\end{aligned}$$

because $\sum\limits_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_J^{(s)}\geqslant 0$ and $p(s)\geqslant 0$ for any $s\in(0,1)$.

5) Obviously we have $f_\emptyset^\varphi = f_{\varphi(\emptyset)} = f_\emptyset = 1$. Further, using the fact that φ is a bijection, we can write

$$\begin{aligned}\sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_{\Lambda\setminus J}^{\varphi} &= \sum_{J\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus J|}f_{\varphi(\Lambda\setminus J)} = \sum_{J\subset\boldsymbol{x}}(-1)^{\left|\varphi(\boldsymbol{x})\setminus\varphi(J)\right|}f_{\varphi(\Lambda)\setminus\varphi(J)} = \\ &= \sum_{J_1\subset\varphi(\boldsymbol{x})}(-1)^{\left|\varphi(\boldsymbol{x})\setminus J_1\right|}f_{\varphi(\Lambda)\setminus J_1}\geqslant 0\end{aligned}$$

by (II.1) because $\boldsymbol{f}$ is a P-function. □

Examples II.4. — **1)** Let $\{f_t,\ t \in \mathbb{Z}^\nu\}$ be a family of real numbers such that $0 \leqslant f_t \leqslant 1$ for any $t \in \mathbb{Z}^\nu$. We put

$$f_J = \prod_{t \in J} f_t, \quad J \in \mathscr{E}.$$

Here and in the sequel any product over an empty space of indexes is considered to be equal 1, *i.e.*, $f_\emptyset = 1$. Then $\boldsymbol{f} = \{f_J,\ J \in \mathscr{E}\}$ is a P-function and the corresponding random field is a random field with independent components and with $\mathbf{P}_{\{t\}}(x) = \begin{cases} f_t & \text{if } x = 0 \\ 1 - f_t & \text{if } x = 1 \end{cases}$ for all $t \in \mathbb{Z}^\nu$. The case $f_t \equiv q$ on $\mathbb{Z}^\nu$, $0 \leqslant q \leqslant 1$, corresponds to Bernoulli random field with parameter $p = 1 - q$. In particular, for $q = 0$ we get a random field which assumes a.s. the value 1 on $\mathbb{Z}^\nu$, and for $q = 1$ a random field which assumes a.s. the value 0 on $\mathbb{Z}^\nu$.

2) Fix some $\tau > 0$ and let, for all $q \in [0,1]$, the function $\boldsymbol{f}^{(q)} = \left\{f_J^{(q)},\ J \in \mathscr{E}\right\}$ be defined by $f_J^{(q)} = q^{|J|}$ (this is a Bernoulli random field from the preceding example). Then the function $\boldsymbol{b}$ defined by

$$b_J = \tau \int_0^1 q^{|J|+\tau-1}\, \mathrm{d}q = \frac{\tau}{|J| + \tau}\ , \quad J \in \mathscr{E}, \tag{II.4}$$

is a P-function corresponding to a random field which is a mixture of the Bernoulli random fields. This is an evident consequence of the Proposition II.3–4 where the probability density p is taken to be $p(q) = \tau q^{\tau-1}$, $q \in [0,1]$, and the family $\mathscr{F} = \left\{\boldsymbol{f}^{(q)}\right\}$ is the family of Bernoulli random fields. The system of finite-dimensional distributions of the mixture random field is given by

$$\mathbf{P}_\Lambda(\boldsymbol{x}) = \frac{\tau}{|\Lambda| + \tau} \prod_{i=1}^{|\boldsymbol{x}|} \frac{i}{|\Lambda| + \tau - i}$$

for all $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$. This can be easily proved by induction over a number of points of the set $\boldsymbol{x}$ using the formula (II.2). As we will see later, this random field is non-Gibbsian (for demonstration see the Section VI.2).

3) Let $\boldsymbol{f}$ be a P-function. Using the Proposition II.3–2 with $T = t \times \mathbb{Z}^{\nu-1}$ we get a P-function $\boldsymbol{f}^{\mathrm{proj}}$ defined by $f_J^{\mathrm{proj}} = f_{J \cap (t \times \mathbb{Z}^{\nu-1})}$ where we have fixed some $t \in \mathbb{Z}$. This P-function corresponds to a random field obtained by projection which may be non-Gibbsian even if the original random field is Gibbsian (see, for example, [23] and [25]).

4) Let $\boldsymbol{f}$ be a P-function. Then the function $\boldsymbol{f}^{\mathrm{dec}}$ defined by $f_J^{\mathrm{dec}} = f_{2J}$ where $2J = \{2t,\ t \in J\}$ is a P-function. This is an evident consequence of

the Proposition II.3–5. This P-function corresponds to a random field obtained by "decimation" which is also known to be in general non-Gibbsian even if the original random field is Gibbsian (see, for example, [16] and [25]).

II.3. Generalizations to the case of arbitrary finite state space

As we have seen in the previous sections, in the {0,1} case one can specify completely a random field by specifying just the probabilities of vacuum configurations: $f_\Lambda = \mathbf{P}_\Lambda(\phi)$. Clearly one could have specified a random field by specifying rather the probabilities of configurations not containing vacuums, that is consisting only of 1's. So, one could have defined the P-functions as $f_\Lambda = \mathbf{P}_\Lambda(\Lambda)$. In this case the Definition II.1 and the Theorem II.2 would be rewritten as follows:

Definition II.5. — *A real-valued function $\boldsymbol{f} = \{f_J,\ J \in \mathscr{E}\}$ on $\mathscr{E}$ is called P-function if $f_\phi = 1$ and for any $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$ we have*

$$\sum_{J \subset \Lambda \setminus \boldsymbol{x}} (-1)^{|J|} f_{\boldsymbol{x} \cup J} \geqslant 0.$$

Theorem II.6. — *A system $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ is a system of probability distributions consistent in Kolmogorov's sense if and only if there exists a P-function $\boldsymbol{f}$ such that for any $\Lambda \in \mathscr{E}$ we have*

$$\mathbf{P}_\Lambda(\boldsymbol{x}) = \sum_{J \subset \Lambda \setminus \boldsymbol{x}} (-1)^{|J|} f_{\boldsymbol{x} \cup J}, \quad \boldsymbol{x} \subset \Lambda.$$

Particularly, for any $\Lambda \in \mathscr{E}$ we have $\mathbf{P}_\Lambda(\Lambda) = f_\Lambda$.

The proof is similar to the one of the Theorem II.1.

This version of the theorem is easily generalized to a case of arbitrary finite state space $\mathscr{X}$. That is, in this case one can still specify completely a random field by specifying just the probabilities of configurations not containing vacuums.

Let us consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$.

Definition II.7. — *A real-valued function* $\boldsymbol{f} = \{f_{\boldsymbol{x}},\ \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \mathscr{E}\}$ *is called* P*-function if* $f_{\boldsymbol{\emptyset}} = 1$ *and for any* $\Lambda \in \mathscr{E}$ *and* $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \subset \Lambda$ *we have*

$$\sum_{J \subset \Lambda \setminus I} (-1)^{|J|} \sum_{\boldsymbol{y} \in \mathscr{X}^{*J}} f_{\boldsymbol{x} \oplus \boldsymbol{y}} \geqslant 0.$$

Theorem II.8. — *A system* $\mathbf{P} = \{\mathbf{P}_\Lambda,\ \Lambda \in \mathscr{E}\}$ *is a system of probability distributions consistent in Kolmogorov's sense if and only if there exists a* P*-function* $\boldsymbol{f}$ *such that for any* $\Lambda \in \mathscr{E}$ *we have*

$$\mathbf{P}_\Lambda(\boldsymbol{x}) = \sum_{J \subset \Lambda \setminus I} (-1)^{|J|} \sum_{\boldsymbol{y} \in \mathscr{X}^{*J}} f_{\boldsymbol{x} \oplus \boldsymbol{y}}, \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \subset \Lambda.$$

Particularly, for any $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \in \mathscr{E}$ *we have* $\mathbf{P}_I(\boldsymbol{x}) = f_{\boldsymbol{x}}$.

The proof for this general case is similar to the one corresponding to the {0,1} case. All the properties of P-functions are also easily generalized for this general case.

III. Random fields, Q-functions and H-functions

In the case of Gibbs random fields one can consider infinite-volume correlation functions as limits of finite-volume correlation functions. In the first sections we consider the {0,1} case. We show that in some cases P-functions can also be considered as limits of finite-volume correlation functions (or rather their generalization). The latter ones can be written down via generalized partition functions (Q-functions) or, equivalently, via the generalized Boltzmann factors (H-functions) which are arbitrary non-negative functions in our case. Then we introduce systems of probability distributions (corresponding to conditional distributions in finite volumes with vacuum boundary conditions) consistent in Dobrushin's sense and describe them via corresponding Q-functions and/or H-functions. Further we give, in terms of cluster representation of Q-functions, a general sufficient condition for existence of limiting P-functions. Finally in Section III.4 we show the way one can generalize the notion of H-functions to the case of arbitrary finite state space $\mathscr{X}$.

III.1. Q-functions and H-functions

Let us start by giving the following

Definition III.1. — *A real-valued function $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ on $\mathscr{E}$ is called Q-function if $\theta_J \neq 0$ for all $J \in \mathscr{E}$, $\theta_\emptyset = 1$ and for any $S \in \mathscr{E}$ we have*

$$\sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J \geqslant 0. \tag{III.1}$$

Unlike P-functions, Q-functions are much easier to specify because they have the following simple constructive description.

Theorem III.2. — *A function $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ is a Q-function if and only if there exists a function $\boldsymbol{H} = \{H_S,\ S \in \mathscr{E}\}$, $H_S \geqslant 0$ for all $S \in \mathscr{E}$, $H_\emptyset = 1$, such that for any $\Lambda \in \mathscr{E}$ we have*

$$\theta_\Lambda = \sum_{S \subset \Lambda} H_S. \tag{III.2}$$

This function $\boldsymbol{H}$ is called H-function.

Proof: **1)** Necessity. Let $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ be a Q-function. Put

$$H_S = \sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J, \quad S \in \mathscr{E}. \tag{III.3}$$

Since $\boldsymbol{\theta}$ is a Q-function and according to the definition (III.3) of H_S, we have $H_\emptyset = 1$ and $H_S \geqslant 0$ for all $S \in \mathscr{E}$. Further, for any $\Lambda \in \mathscr{E}$ we can write

$$\sum_{S \subset \Lambda} H_S = \sum_{S \subset \Lambda} \sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J = \sum_{J \subset \Lambda} \theta_J \sum_{S\,:\,J \subset S \subset \Lambda} (-1)^{|S \setminus J|} = \theta_\Lambda.$$

2) Sufficiency. Let $\boldsymbol{H}$ be a H-function and $\theta_\Lambda = \sum\limits_{S \subset \Lambda} H_S$. Clearly $\theta_\emptyset = H_\emptyset = 1$ and $\theta_\Lambda \geqslant H_\emptyset = 1 > 0$ for all $\Lambda \in \mathscr{E}$. Finally, for all $S \in \mathscr{E}$ we have

$$\begin{aligned} \sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J &= \sum_{J \subset S} (-1)^{|S \setminus J|} \sum_{\widetilde{J} \subset J} H_{\widetilde{J}} = \\ &= \sum_{\widetilde{J} \subset S} H_{\widetilde{J}} \sum_{J\,:\,\widetilde{J} \subset J \subset S} (-1)^{|S \setminus J|} = H_S \geqslant 0 \end{aligned}$$

which concludes the proof. □

Since $H_{\boldsymbol{x}} \geqslant 0$ for all $\boldsymbol{x} \in \mathscr{E}$ we can denote $U(\boldsymbol{x}) = -\ln H_{\boldsymbol{x}}$ $\big($we permit the function $\boldsymbol{U} = \{U(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{E}\}$ to take the value $+\infty\big)$. Then (III.2) can be rewritten in the following form

$$\theta_\Lambda = \sum_{\boldsymbol{x} \subset \Lambda} \exp\big(-U(\boldsymbol{x})\big)$$

and we see that $\boldsymbol{H}$ is nothing but Boltzmann factors and $\boldsymbol{\theta}$ is nothing but the partition function defined through a general Hamiltonian $\boldsymbol{U}$ (without boundary conditions) not using an interaction potential.

Proposition III.3. — *Let $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ be a Q-function. Then for any $\Lambda \in \mathscr{E}$ the function*

$$\boldsymbol{f}^{(\Lambda)} = \left\{ f_J^{(\Lambda)} = \frac{\theta_{\Lambda \setminus J}}{\theta_\Lambda},\ J \in \mathscr{E} \right\}$$

is a P-function.

Proof: Let us fix some $\Lambda \in \mathscr{E}$. Obviously $f_\emptyset^{(\Lambda)} = \theta_\Lambda / \theta_\Lambda = 1$. Further, for any

$I \in \mathscr{E}$ and $\boldsymbol{x} \subset I$ we have

$$\sum_{J\subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} f^{(\Lambda)}_{I\setminus J} = \sum_{J\subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|}\ \frac{\theta_{\Lambda\setminus(I\setminus J)}}{\theta_\Lambda} =$$

$$= \frac{1}{\theta_\Lambda} \sum_{J_1\subset \boldsymbol{x}\cap\Lambda}\ \sum_{J_2\subset \boldsymbol{x}\cap\Lambda^{\mathrm{c}}} (-1)^{\left|(\boldsymbol{x}\cap\Lambda)\setminus J_1\right|}\, (-1)^{\left|(\boldsymbol{x}\cap\Lambda^{\mathrm{c}})\setminus J_2\right|} \theta_{(\Lambda\setminus I)\cup J_1} =$$

$$= \frac{1}{\theta_\Lambda} \sum_{J_1\subset \boldsymbol{x}\cap\Lambda} (-1)^{\left|(\boldsymbol{x}\cap\Lambda)\setminus J_1\right|} \theta_{(\Lambda\setminus I)\cup J_1}\ \times \sum_{J_2\subset \boldsymbol{x}\cap\Lambda^{\mathrm{c}}} (-1)^{\left|(\boldsymbol{x}\cap\Lambda^{\mathrm{c}})\setminus J_2\right|}.$$

Let us denote the first sum by F_1 and the second one by F_2. For F_2 we have by (II.3)

$$F_2 = \begin{cases} 1 & \text{if } \boldsymbol{x} \subset \Lambda, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $F_2 \geqslant 0$ and we have to calculate F_1 only for the case $\boldsymbol{x} \subset \Lambda$. Since $\boldsymbol{\theta}$ is a Q-function, for all $S \subset \Lambda \setminus I$ we have $\sum\limits_{J\subset \boldsymbol{x}\cup S} (-1)^{\left|(\boldsymbol{x}\cup S)\setminus J\right|}\theta_J \geqslant 0$ and hence

$$0 \leqslant \sum_{S\subset\Lambda\setminus I}\ \sum_{J\subset \boldsymbol{x}\cup S} (-1)^{\left|(\boldsymbol{x}\cup S)\setminus J\right|}\theta_J =$$

$$= \sum_{J_1\subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J_1|} \sum_{J_2\subset\Lambda\setminus I} \theta_{J_1\cup J_2} \sum_{S\,:\,J_2\subset S\subset\Lambda\setminus I} (-1)^{|S\setminus J_2|} =$$

$$= \sum_{J_1\subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J_1|}\theta_{(\Lambda\setminus I)\cup J_1} = F_1.$$

So, we get (II.1) and hence $\boldsymbol{f}^{(\Lambda)}$ is a P-function. □

Note that using the above mentioned notation $\boldsymbol{U}$ we can write

$$f^{(\Lambda)}_J = \frac{\sum\limits_{\boldsymbol{x}\subset\Lambda\setminus J} \exp\bigl(U(\boldsymbol{x})\bigr)}{\sum\limits_{\boldsymbol{y}\subset\Lambda} \exp\bigl(U(\boldsymbol{y})\bigr)}$$

which is the Gibbsian form for finite-volume correlation functions but for a general Hamiltonian $\boldsymbol{U}$. Note also that since the space $\mathscr{B}^P$ of all P-functions is closed then, if for some sequence $\Lambda_n \in \mathscr{E}$ such that $\Lambda_n \uparrow \mathbb{Z}^\nu$ the P-functions $\boldsymbol{f}^{(\Lambda_n)}$ converge as $n \to \infty$ to some function $\boldsymbol{f}$, this function $\boldsymbol{f}$ is a new P-function which is a generalized limiting (infinite-volume) correlation function. This is a *limiting P-function* and it corresponds to a *limiting random field* $\mathbf{P}$.

III.2. Consistency in Dobrushin's sense

To any Q-function $\boldsymbol{\theta}$ one can associate a system $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ where $\mathbf{Q}_\Lambda = \{\mathbf{Q}_\Lambda(\boldsymbol{x}),\ \boldsymbol{x} \subset \Lambda\}$ and $\mathbf{Q}_\Lambda(\boldsymbol{x})$ is defined by the formula

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta_\Lambda} \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \theta_J, \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \subset \Lambda.$$

This system turns out to be a system of probability distributions. Note that using the notation $\boldsymbol{U}$ and the formulas (III.2) and (III.3) one can rewrite $\mathbf{Q}_\Lambda(\boldsymbol{x})$ in the form

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{\exp\bigl(U(\boldsymbol{x})\bigr)}{\sum\limits_{\boldsymbol{y} \subset \Lambda} \exp\bigl(U(\boldsymbol{y})\bigr)}$$

which is the classical Gibbsian form but for a general Hamiltonian $\boldsymbol{U}$. In general, the system $\mathbf{Q}$ is not consistent in Kolmogorov's sense. It is rather consistent in so-called "Dobrushin's sense".

Definition III.4. — *A system of probability distributions* $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ *is called consistent in Dobrushin's sense if for all* $\Lambda, \widetilde{\Lambda} \in \mathscr{E}$ *such that* $\Lambda \cap \widetilde{\Lambda} = \emptyset$ *and for all* $\boldsymbol{x} \subset \Lambda$ *we have*

$$\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}(\boldsymbol{x}) = \mathbf{Q}_\Lambda(\boldsymbol{x}) \left(\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}\right)_{\widetilde{\Lambda}}(\emptyset). \tag{III.4}$$

Note that in the case when $\mathbf{Q}_\Lambda(\emptyset) > 0$ for all $\Lambda \in \mathscr{E}$ the condition (III.4) can be rewritten in an equivalent form

$$\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}(\boldsymbol{x}) = \frac{\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}(\emptyset)}{\mathbf{Q}_\Lambda(\emptyset)}\ \mathbf{Q}_\Lambda(\boldsymbol{x}).$$

Note also that Dobrushin's consistency condition (III.4) is just a particular case of the condition (I.4) and is satisfied by the system of conditional distributions in finite volumes with vacuum boundary conditions of a random field. Below we will see that under some conditions the system of probability distributions consistent in Dobrushin's sense is indeed the system of conditional distributions in finite volumes with vacuum boundary conditions for the limiting random field. But before let us show how the systems of probability distributions consistent in Dobrushin's sense can be described.

Theorem III.5. — *A system* $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ *is a system of probability distributions consistent in Dobrushin's sense and satisfying* $\mathbf{Q}_\Lambda(\emptyset) > 0$ *for all*

$\Lambda \in \mathscr{E}$ if and only if there exists a Q-function $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ such that for all $\Lambda \in \mathscr{E}$ we have

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta_\Lambda} \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \theta_J, \quad \boldsymbol{x} \subset \Lambda. \tag{III.5}$$

Particularly, for all $\Lambda \in \mathscr{E}$ we have $\mathbf{Q}_\Lambda(\emptyset) = 1/\theta_\Lambda$.

Proof: **1)** Necessity. Let $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ be a system of probability distributions consistent in Dobrushin's sense with $\mathbf{Q}_\Lambda(\emptyset) > 0$ for all $\Lambda \in \mathscr{E}$. Put $\theta_\Lambda = 1/\,\mathbf{Q}_\Lambda(\emptyset)$. We have obviously $\theta_\Lambda \neq 0$ and $\theta_\emptyset = 1$. Further, for any $\Lambda \in \mathscr{E}$ and $J \subset \Lambda$ we can write

$$1 = \sum_{S \subset J} \mathbf{Q}_J(S) = \sum_{S \subset J} \frac{\mathbf{Q}_J(\emptyset)}{\mathbf{Q}_\Lambda(\emptyset)}\ \mathbf{Q}_\Lambda(S) = \frac{\mathbf{Q}_J(\emptyset)}{\mathbf{Q}_\Lambda(\emptyset)} \sum_{S \subset J} \mathbf{Q}_\Lambda(S)$$

or equivalently

$$\theta_J = \theta_\Lambda \sum_{S \subset J} \mathbf{Q}_\Lambda(S).$$

Therefore

$$\sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \theta_J = \theta_\Lambda \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \sum_{S \subset J} \mathbf{Q}_\Lambda(S) = \theta_\Lambda\ \mathbf{Q}_\Lambda(\boldsymbol{x})$$

and we obtain (III.1) and (III.5).

2) Sufficiency. Let $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ be a Q-function. First of all, let us note that for all $\Lambda \in \mathscr{E}$ we have $\theta_\Lambda = \sum\limits_{S \subset \Lambda} H_S \geqslant H_\emptyset = 1 > 0$. Now let us put for any $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta_\Lambda} \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \theta_J = \frac{H_{\boldsymbol{x}}}{\theta_\Lambda} \geqslant 0$$

and prove that $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ is a system of probability distributions consistent in Dobrushin's sense. We have

$$\sum_{\boldsymbol{x} \subset \Lambda} \mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta_\Lambda} \sum_{\boldsymbol{x} \subset \Lambda} H_{\boldsymbol{x}} = \frac{1}{\theta_\Lambda}\ \theta_\Lambda = 1,$$

i.e., the system $\mathbf{Q}$ is a system of probability distributions. Now let us verify its consistency. We have

$$\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}}{\theta_{\Lambda \cup \widetilde{\Lambda}}} = \frac{\theta_\Lambda}{\theta_{\Lambda \cup \widetilde{\Lambda}}}\ \frac{H_{\boldsymbol{x}}}{\theta_\Lambda} = \frac{\mathbf{Q}_{\Lambda \cup \widetilde{\Lambda}}(\emptyset)}{\mathbf{Q}_\Lambda(\emptyset)}\ \mathbf{Q}_\Lambda(\boldsymbol{x}).$$

The theorem is proved. □

Now we can state the theorem showing when the system of probability distributions consistent in Dobrushin's sense is indeed the system of conditional distributions in finite volumes with vacuum boundary conditions for the limiting random field.

Theorem III.6. — *Let $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ be a Q-function and $\mathbf{Q}$ be the corresponding system of probability distributions consistent in Dobrushin's sense. For each $\Lambda \in \mathscr{E}$ we consider the above introduced P-function $\boldsymbol{f}^{(\Lambda)}$.*

1) *Let $\Lambda \in \mathscr{E}$ and let $\mathbf{P}^{(\Lambda)}$ be the random field corresponding to the P-function $\boldsymbol{f}^{(\Lambda)}$. The finite-dimensional distributions of this random field have the following form: for each $I \in \mathscr{E}$ and $\boldsymbol{x} \subset I$ we have*

$$\mathbf{P}_I^{(\Lambda)}(\boldsymbol{x}) = \begin{cases} \big(\mathbf{Q}_\Lambda\big)_{\Lambda\cap I}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \subset \Lambda, \\ 0 & \text{otherwise.} \end{cases} \tag{III.6}$$

2) *One can choose a sequence $\Lambda_n \in \mathscr{E}$ such that $\Lambda_n \uparrow \mathbb{Z}^\nu$ and that the P-functions $\boldsymbol{f}^{(\Lambda_n)}$ converge as $n \to \infty$ to a limiting P-function $\boldsymbol{f}$, i.e., for all $J \in \mathscr{E}$ we have*

$$\lim_{n\to\infty} f_J^{(\Lambda_n)} = f_J.$$

3) *Suppose moreover that for any $J \in \mathscr{E}$ the limit*

$$\lim_{\Lambda\uparrow\mathbb{Z}^\nu} f_J^{(\Lambda)} = f_J \tag{III.7}$$

exists and the convergence is "absolute" in the sense of the definition I.2–2). Then the function $\boldsymbol{f}$ is a limiting P-function and the corresponding limiting random field $\mathbf{P}$ satisfies

$$\boldsymbol{q}_J^{\emptyset}(\emptyset) = 1/\theta_J \tag{III.8}$$

for any $J \in \mathscr{E}$.

Proof: **1)** Using details of the proof of the Proposition III.3 and formulas (II.2) and (III.5) we get

$$\mathbf{P}_I^{(\Lambda)}(\boldsymbol{x}) = \sum_{J\subset\boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} f_{I\setminus J}^{(\Lambda)} = \frac{1}{\theta_\Lambda}\, F_1\, F_2$$

with

$$F_1 = \sum_{S\subset\Lambda\setminus I}\ \sum_{J\subset\boldsymbol{x}\cup S} (-1)^{\left|(\boldsymbol{x}\cup S)\setminus J\right|}\theta_J = \theta_\Lambda \sum_{S\subset\Lambda\setminus I} \mathbf{Q}_\Lambda(\boldsymbol{x}\cup S) = \theta_\Lambda \left(\mathbf{Q}_\Lambda\right)_{\Lambda\cap I}(\boldsymbol{x}).$$

and

$$F_2 = \begin{cases} 1 & \text{if } \boldsymbol{x} \subset \Lambda, \\ 0 & \text{otherwise.} \end{cases}$$

Now the representation (III.6) is evident.

2) This is an obvious consequence of the compactness of the set $\mathscr{B}^P$ of all P-functions.

3) The fact that $\boldsymbol{f}$ is a P-function is also a consequence of the compactness of the set $\mathscr{B}^P$. To verify the relation (III.8) let us fix some sequence $\overline{J}_n \uparrow J^{\mathrm{c}}$. Using the "absoluteness" of the convergence (III.7) we can write

$$\boldsymbol{q}_J^{\emptyset}(\emptyset) = \lim_{n\to\infty} \frac{\mathbf{P}_{J\cup\overline{J}_n}(\emptyset)}{\mathbf{P}_{\overline{J}_n}(\emptyset)} = \lim_{n\to\infty} \frac{f_{J\cup\overline{J}_n}}{f_{\overline{J}_n}} = \lim_{n\to\infty} \lim_{m\to\infty} \frac{f_{J\cup\overline{J}_n}^{(J\cup\overline{J}_m)}}{f_{\overline{J}_n}^{(J\cup\overline{J}_m)}} =$$

$$= \lim_{n\to\infty} \lim_{m\to\infty} \frac{\theta_{\overline{J}_m\setminus\overline{J}_n} \,/\, \theta_{J\cup\overline{J}_m}}{\theta_{J\cup(\overline{J}_m\setminus\overline{J}_n)} \,/\, \theta_{J\cup\overline{J}_m}} = \lim_{n\to\infty} \lim_{m\to\infty} f_J^{\left(J\cup(\overline{J}_m\setminus\overline{J}_n)\right)} =$$

$$= \lim_{n\to\infty} \lim_{m\to\infty} \sum_{R\subset J\cup(\overline{J}_m\setminus\overline{J}_n)} b_R(J) = \lim_{n\to\infty} \sum_{R\in\mathscr{E}\,:\,R\subset\overline{J}_n^{\,\mathrm{c}}} b_R(J) =$$

$$= \sum_{R\subset J} b_R(J) = f_J^{(J)} = \frac{1}{\theta_J}$$

which concludes the proof. □

Remarks III.7. — **1)** The relation (III.8) between the limiting random field and the original system of probability distributions consistent in Dobrushin's sense (Q-function) can be rewritten in the form

$$\boldsymbol{q}_J^{\emptyset}(\emptyset) = \mathbf{Q}_J(\emptyset), \quad J \in \mathscr{E}.$$

Note that the relation

$$\boldsymbol{q}_J^{\emptyset}(\boldsymbol{x}) = \mathbf{Q}_J(\boldsymbol{x}), \quad J \in \mathscr{E},\ \boldsymbol{x} \subset J,$$

also holds. At first sight it seems to be more general than (III.8), but in reality they are equivalent because the systems $\{\boldsymbol{q}_J^{\emptyset},\ J \in \mathscr{E}\}$ and $\{\mathbf{Q}_J,\ J \in \mathscr{E}\}$ of probability distributions are both consistent in Dobrushin's sense and hence they are determined uniquely and in the same manner (more precisely by the formula (III.5)) by the functions $\{\boldsymbol{q}_J^{\emptyset}(\emptyset),\ J \in \mathscr{E}\}$ and $\{\mathbf{Q}_J(\emptyset),\ J \in \mathscr{E}\}$ respectively.

2) In the relation (III.8) one cannot replace $\boldsymbol{q}_J^{\emptyset}(\emptyset)$ by $\mathbf{Q}_J^{\emptyset}(\emptyset)$ coming from an arbitrary conditional distribution $\boldsymbol{\mathcal{Q}}$ of the random field $\mathbf{P}$, because in general

$\mathbf{Q}_J^{\phi}(\phi)$ is not necessarily equal to $\boldsymbol{q}_J^{\phi}(\phi)$ although the last one is well-defined for the random field $\mathbf{P}$.

3) The "absoluteness" of the convergence in (III.7) is essential for the relation (III.8). If the convergence holds but is not "absolute" this relation can fail as shows the following

Example III.8. — Let $\tau > 0$ and consider a function $\theta_J = \frac{|J|+\tau}{\tau}$. It is not difficult to check that this is a Q-function and that the corresponding system of probability distributions consistent in Dobrushin's sense has the following form

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \begin{cases} \frac{\tau}{|\Lambda|+\tau} & \text{if } \boldsymbol{x} = \phi, \\ \frac{1}{|\Lambda|+\tau} & \text{if } |\boldsymbol{x}| = 1, \\ 0 & \text{if } |\boldsymbol{x}| \geqslant 2. \end{cases}$$

For this Q-function the limits in (III.7) exist. In fact, for any $J \in \mathscr{E}$ we have

$$f_J = \lim_{\Lambda \uparrow \mathbb{Z}^\nu} \frac{\theta_{\Lambda \setminus J}}{\theta_\Lambda} = \lim_{\Lambda \uparrow \mathbb{Z}^\nu} \frac{|\Lambda \setminus J| + \tau}{|\Lambda| + \tau} = 1.$$

As we see, the limiting random field is a random field assuming a.s. the value 0 on $\mathbb{Z}^\nu$. Obviously in this random field we have $\boldsymbol{q}_J^{\phi}(\phi) = 1$ for all $J \in \mathscr{E}$ and the relation (III.8) fails.

III.3. Cluster expansions

Now, let us give an example (or rather a whole class of examples) when the convergence in (III.7) is "absolute". This example is a generalization of a wide class of models occurring in the Gibbs random fields theory and called "models allowing cluster expansion". For this we need to introduce some combinatorial notions. For all $\Lambda \in \mathscr{E} \setminus \{\phi\}$ let us fix an arbitrary point $t_\Lambda \in \Lambda$ and denote $\Lambda' = \Lambda \setminus \{t_\Lambda\}$.

Definition III.9. — **1)** *We define a partially ordering in $\mathscr{E}$ in the following way. For $A, B \in \mathscr{E}$ we say that $B \leqslant A$ if there exists an $n \in \mathbb{N}$ and a sequence $B = A_1, A_2, \ldots, A_n = A$ of elements of $\mathscr{E}$ such that we have $A_{i-1} = A_i \setminus t_{A_i}$ for all $i = 2, \ldots, n$.*

2) *A sequence $\gamma = \{B_1, \Gamma_1; \ldots; B_n, \Gamma_n\}$ such that we have $B_1 \leqslant A \in \mathscr{E}$, $B_i, \Gamma_i \in \mathscr{E}$ and $\Gamma_i \cap B_i = t_{B_i}$ for all $i = 1, \ldots, n$, and $B_i \leqslant B_{i-1} \cup \Gamma_{i-1}$ for*

all $i = 2, \dots, n$, is called path beginning at A. The number n is called length of the path γ and the set $\Gamma_1 \cup \dots \cup \Gamma_n \in \mathscr{E}$ is called support of the path γ. The set of all pathes beginning at A and of length n will be denoted by $\mathbf{\Gamma}^{(n)}(A)$ and the set of all pathes beginning at A and with support R by $\mathbf{\Gamma}_R(A)$.

3) *A sequence $\delta = \{\Gamma_1, \dots, \Gamma_n\}$ such that we have $\Gamma_i \neq \emptyset$ and $\Gamma_i \subset \Lambda \in \mathscr{E}$ for all $i = 1, \dots, n$, and $\Gamma_i \cap \Gamma_j = \emptyset$ for any pair $i, j = 1, \dots, n$ with $i \neq j$, is called weak partition of Λ. Note that we allow the partition to be empty, i.e., $n = 0$. The set of all weak partitions of Λ will be denoted by $\mathbf{\Delta}_\Lambda^{\mathrm{w}}$.*

4) *A weak partition $\delta = \{\Gamma_1, \dots, \Gamma_n\}$ of a set $\Lambda \in \mathscr{E}$ is called partition of Λ, if we have $\Gamma_1 \cup \dots \cup \Gamma_n = \Lambda$. The set of all partitions of Λ will be denoted by $\mathbf{\Delta}_\Lambda$.*

Theorem III.10. — **1)** *Let $\boldsymbol{K} = \{K_J,\ J \in \mathscr{E}\}$ be a real-valued function such that*

$$F(\Lambda) = \sum_{\{\Gamma_1, \dots, \Gamma_n\} \in \mathbf{\Delta}_\Lambda} K_{\Gamma_1} \cdots K_{\Gamma_n} \geqslant 0, \quad \Lambda \in \mathscr{E}. \tag{III.9}$$

Then the function $\boldsymbol{\theta} = \{\theta_\Lambda,\ \Lambda \in \mathscr{E}\}$ defined by

$$\theta_\Lambda = \sum_{\{\Gamma_1, \dots, \Gamma_n\} \in \mathbf{\Delta}_\Lambda^{\mathrm{w}}} K_{\Gamma_1} \cdots K_{\Gamma_n}, \quad \Lambda \in \mathscr{E}.$$

is a Q-function.

2) *If, moreover, there exist some $\lambda, \alpha \geqslant 0$ such that $\lambda\,(1 + \sqrt{\alpha}\,)^2 < 1$, and for all $t \in \mathbb{Z}^\nu$ and $n \in \mathbb{N}$ we have*

$$\sum_{\Gamma\,:\,t \in \Gamma \text{ and } |\Gamma| = n} |K_\Gamma| \leqslant \alpha \lambda^n,$$

then for any $J, \Lambda \in \mathscr{E}$ we have the representation

$$f_J^{(\Lambda)} = \sum_{R \subset \Lambda} b_R(J)$$

where

$$b_R(J) = \sum_{\{B_1, \Gamma_1; \dots; B_n, \Gamma_n\} \in \mathbf{\Gamma}_R(J)} (-1)^n\, K_{\Gamma_1} \cdots K_{\Gamma_n}$$

for all $R, J \in \mathscr{E}$, and the series $\sum\limits_{R \in \mathscr{E}} b_R(J)$ converges absolutely for any $J \in \mathscr{E}$. Hence, the conditions of the Theorem III.6–3 are satisfied and there exists a limiting random field $\mathbf{P}$ satisfying (III.8) and corresponding to a P-function $\boldsymbol{f} = \{f_J,\ J \in \mathscr{E}\}$ defined by

$$f_J = \sum_{n=0}^{\infty} \sum_{\{B_1, \Gamma_1; \dots; B_n, \Gamma_n\} \in \mathbf{\Gamma}^{(n)}(J)} (-1)^n\, K_{\Gamma_1} \cdots K_{\Gamma_n}, \quad J \in \mathscr{E}.$$

Proof: **1)** First of all, for any $\Lambda \in \mathscr{E}$, we have

$$\theta_\Lambda = \sum_{R \subset \Lambda} F(R) \geqslant F(\emptyset) = 1 > 0$$

and $\theta_\emptyset = F(\emptyset) = 1$. It remains to verify the condition (III.1). Indeed, for any $S \in \mathscr{E}$, we have

$$\sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J = \sum_{J \subset S} (-1)^{|S \setminus J|} \sum_{R \subset J} F(R) = F(S) \geqslant 0.$$

2) For an arbitrary $t \in \mathbb{Z}^\nu$ and any $V \in \mathscr{E}$ such that $t \in V$ we can write

$$\begin{aligned}
\theta_V = \sum_{R \subset V} F(R) &= \sum_{R \subset V \setminus \{t\}} F(R) + \sum_{R \,:\, t \in R \subset V} F(R) = \\
&= \theta_{V \setminus \{t\}} + \sum_{\Gamma \,:\, t \in \Gamma \subset V} \left(K_\Gamma \sum_{\{\Gamma_1, \dots, \Gamma_n\} \in \mathbf{\Delta}^{\mathrm{w}}_{V \setminus \Gamma}} K_{\Gamma_1} \cdots K_{\Gamma_n} \right) = \\
&= \theta_{V \setminus \{t\}} + \sum_{\Gamma \,:\, t \in \Gamma \subset V} \big(K_\Gamma \, \theta_{V \setminus \Gamma} \big).
\end{aligned}$$

For $J = \emptyset$ the assertion of the theorem is trivial, so let us suppose that $|J| \geqslant 1$, and apply the last equality for $V = \Lambda \setminus J'$ and $t = t_J$. We get

$$\theta_{\Lambda \setminus J'} = \theta_{\Lambda \setminus J} + \sum_{\Gamma \,:\, t_J \in \Gamma \subset \Lambda \setminus J'} \big(K_\Gamma \, \theta_{\Lambda \setminus (J \cup \Gamma)} \big)$$

and hence

$$f^{(\Lambda)}_J = f^{(\Lambda)}_{J'} + \sum_{\Gamma \,:\, t_J \in \Gamma \subset \Lambda \setminus J'} \big(-K_\Gamma \, f^{(\Lambda)}_{J \cup \Gamma} \big).$$

For any $\Lambda \in \mathscr{E}$, we can write the last equation for all $J \in \mathscr{E}_1(\Lambda)$ where $\mathscr{E}_1(\Lambda)$ denotes the set of all non-empty subsets of Λ. So, we will get a system of $2^{|V|} - 1$ linear equations with $2^{|V|} - 1$ unknown variables $f^{(\Lambda)}_J$, $J \in \mathscr{E}_1(\Lambda)$ $\big($note that $f^{(\Lambda)}_\emptyset = 1$, and so we substitute this value in the equations$\big)$. Let us rewrite this system in "operator-matrix" form. For this we introduce the space $\mathscr{B}^{(\Lambda)}$ of vectors $f^{(\Lambda)} = \big(f^{(\Lambda)}_J, \quad J \subset \mathscr{E}_1(\Lambda) \big)$ indexed by non-empty subsets of Λ, endowed with the norm $\big\| f^{(\Lambda)} \big\| = \sup\limits_{J \in \mathscr{E}_1(\Lambda)} \Big(M^{-|J|} \, \big| f^{(\Lambda)}_J \big| \Big)$ where $M > 1$ is some fixed number that we will specify later.

Let us introduce the basis $\big\{ \chi^{(J)}, \quad J \subset \mathscr{E}_1(\Lambda) \big\}$ in the space $\mathscr{B}^{(\Lambda)}$, by putting $\chi^{(J)} = \big(\chi^{(J)}_V, \quad V \subset \mathscr{E}_1(\Lambda) \big)$ with $\chi^{(J)}_V = \mathbb{1}_{\{J = V\}}$.

We define the "generalized shift" operator by the matrix $R = (r_{JV})_{J,V}$ with $r_{JV} = \mathbb{1}_{\{V=J'\}}$. Clearly this operator will associate to each $f^{(\Lambda)} \in \mathscr{B}^{(\Lambda)}$ the vector

$$R\,f^{(\Lambda)} = \big(f^{(\Lambda)}_{J'}, \quad J \subset \mathscr{E}_1(\Lambda)\big) - \sum_{t\in\Lambda} \chi^{(\{t\})}. \tag{III.10}$$

We define also the operator K by the matrix $(k_{JV})_{J,V}$ with $k_{JV} = -K_{V\setminus J'}\,\mathbb{1}_{\{J\subset V\}}$. Clearly this operator will associate to each $f^{(\Lambda)} \in \mathscr{B}^{(\Lambda)}$ the vector $K\,f^{(\Lambda)}$ with coordinates

$$K\,f^{(\Lambda)}_J = \sum_{V\,:\,J\subset V\subset\Lambda} \big(-K_{V\setminus J'}\,f^{(\Lambda)}_V\big) = \sum_{\Gamma\,:\,t_J\in\Gamma\subset\Lambda\setminus J'} \big(-K_\Gamma\,f^{(\Lambda)}_{J\cup\Gamma}\big). \tag{III.11}$$

Combining (III.10) and (III.11) we see that our system of equations is nothing but

$$f^{(\Lambda)} = R\,f^{(\Lambda)} + \sum_{t\in\Lambda} \chi^{(\{t\})} + K\,f^{(\Lambda)}$$

or, equivalently,

$$\big[E - (R+K)\big] f^{(\Lambda)} = \sum_{t\in\Lambda} \chi^{(\{t\})} \tag{III.12}$$

where E is the unit matrix.

Let us now estimate $\|R+K\|$. For this let us note that

$$\big\|R\,f^{(\Lambda)}\big\| \leqslant \sup_{J\in\mathscr{E}_1(\Lambda)} \Big(M^{-|J|}\,\big|f^{(\Lambda)}_{J'}\big|\Big) \leqslant \sup_{J\in\mathscr{E}_1(\Lambda)} \Big(M^{-|J|}\,M^{|J'|}\,\big\|f^{(\Lambda)}\big\|\Big) = \frac{1}{M}\,\big\|f^{(\Lambda)}\big\|$$

and

$$\begin{aligned}
\big\|K\,f^{(\Lambda)}\big\| &\leqslant \sup_{J\in\mathscr{E}_1(\Lambda)} \Bigg(M^{-|J|} \sum_{\Gamma\,:\,t_J\in\Gamma\subset\Lambda\setminus J'} \big|K_\Gamma\,f^{(\Lambda)}_{J\cup\Gamma}\big|\Bigg) \leqslant \\
&\leqslant \big\|f^{(\Lambda)}\big\| \sup_{J\in\mathscr{E}_1(\Lambda)} \sum_{\Gamma\,:\,t_J\in\Gamma\subset\Lambda\setminus J'} \big|K_\Gamma\big|\,\frac{M^{|J\cup\Gamma|}}{M^{|J|}} = \\
&= \frac{\big\|f^{(\Lambda)}\big\|}{M} \sup_{J\in\mathscr{E}_1(\Lambda)} \sum_{\Gamma\,:\,t_J\in\Gamma\subset\Lambda\setminus J'} \big|K_\Gamma\big|\,M^{|\Gamma|} \leqslant \\
&\leqslant \frac{\big\|f^{(\Lambda)}\big\|}{M} \sup_{t\in\mathbb{Z}^\nu} \sum_{n=1}^{\infty} M^n \sum_{\Gamma\,:\,t\in\Gamma \text{ and } |\Gamma|=n} \big|K_\Gamma\big| \leqslant \\
&\leqslant \frac{\alpha\big\|f^{(\Lambda)}\big\|}{M} \sum_{n=1}^{\infty} \big(M^n\,\lambda^n\big) = \frac{\alpha\big\|f^{(\Lambda)}\big\|}{M}\,\frac{M\,\lambda}{1-M\,\lambda} = \frac{\alpha\,\lambda}{1-M\,\lambda}\,\big\|f^{(\Lambda)}\big\|,
\end{aligned}$$

if $M\,\lambda < 1$, *i.e.*, $\lambda < 1/M$. Hence $\big\|R+K\big\| \leqslant \big\|R\big\| + \big\|K\big\| \leqslant \dfrac{1}{M} + \dfrac{\alpha\,\lambda}{1-M\,\lambda}$. The last expression is smaller that 1 if

$$\lambda < \frac{M-1}{M(M+\alpha-1)} \leqslant \frac{1}{M}. \tag{III.13}$$

If we choose $M = 1+\sqrt{\alpha}$, then $\dfrac{M-1}{M(M+\alpha-1)} = \dfrac{1}{(1+\sqrt{\alpha}\,)^2}$, and hence (III.13) is satisfied.

So, we have proved that $\big\|R+K\big\| < 1$, and hence the system (III.12) has a unique solution given by

$$\begin{aligned} f^{(\Lambda)} &= \big[E-(R+K)\big]^{-1} \sum_{t\in\Lambda} \chi^{(\{t\})} = \sum_{n=0}^{\infty} (R+K)^n \sum_{t\in\Lambda} \chi^{(\{t\})} = \\ &= \sum_{p=0}^{\infty} \sum_{(m_1,\ldots,m_{p+1})\,:\,m_i\geqslant 0,\ i=1,\ldots,p+1} \left(R^{m_1}\,K\cdots K\,R^{m_{p+1}} \sum_{t\in\Lambda} \chi^{(\{t\})} \right) \end{aligned} \tag{III.14}$$

and satisfying

$$f_J^{(\Lambda)} \leqslant M^{|J|}\,\frac{\Big\|\sum\limits_{t\in\Lambda} \chi^{(\{t\})}\Big\|}{1-\big\|R+K\big\|} \leqslant M^{|J|-1}\,\frac{1}{1-\frac{1}{M}-\frac{\alpha\,\lambda}{1-M\,\lambda}} = C \tag{III.15}$$

where the constant C does not depend on Λ, but only on J, α and λ.

Let us rewrite (III.14) coordinate by coordinate. For this let us note at first, that the matrix $R^m = \big(r^m_{JV}\big)_{J,V}$ is given by $r^m_{JV} = \mathbb{1}_{\{J=V^{(m)}\}}$ where we denote $V^{(0)} = V$ and $V^{(m)} = \big(V^{(m-1)}\big)'$. Now we can see that

$$\begin{aligned} f_J^{(\Lambda)} &= \sum_{t\in\Lambda} \sum_{p=0}^{\infty} \sum_{(m_1,\ldots,m_{p+1})\,:\,m_i\geqslant 0,\ i=1,\ldots,p+1} \Big(R^{m_1}\,K\,R^{m_2}\cdots K\,R^{m_{p+1}}\,\chi_J^{(\{t\})} \Big) = \\ &= \sum_{t\in\Lambda} \sum_{p=0}^{\infty} \sum_{(m_1,\ldots,m_{p+1})\,:\,m_i\geqslant 0,\ i=1,\ldots,p+1} \\ &\qquad \sum_{(J_1,V_1;\ldots;J_p,V_p)\,:\,J_i,V_i\in\mathscr{P}_1(\Lambda),\ i=1,\ldots,p} \Big(r^{m_1}_{JJ_1}\,k_{J_1V_1}\,r^{m_2}_{V_1J_2}\cdots k_{J_pV_p}\,r^{m_{p+1}}_{V_p\{t\}} \Big) = \\ &= \sum \big((-K_{\Gamma_1})\cdots(-K_{\Gamma_p}) \big) \end{aligned}$$

where the last sum is taken over all sequences $(J_1, V_1; \ldots; J_p, V_p)$ such that all the sets are included in Λ, $\quad J_1 \leqslant J, \quad J_{i+1} \leqslant V_i$ for all $i = 2, \ldots, n$, and $V_i = J_i \cup \Gamma_i$

with some Γ_i such that $J_i \cap \Gamma_i = t_{J_i}$ for all $i = 1, \dots, n$; or, equivalently, over all pathes beginning at J with support included in Λ. So, we have obtained

$$f_J^{(\Lambda)} = \sum_{R \subset \Lambda} \sum_{\{B_1, \Gamma_1; \dots; B_n, \Gamma_n\} \in \mathbf{\Gamma}_R(J)} \big((-K_{\Gamma_1}) \cdots (-K_{\Gamma_n})\big) = \sum_{R \subset \Lambda} b_R(J).$$

The absolute convergence of the series from the last formula follows immediately from the obvious remark, that if we change the signs of K_Γ-s to make them negative, the estimate of the norm of the matrix K remains unchanged, and hence (III.15) is still valid, that is, partial sums of the series with absolute values are bounded by the same constant C. □

This theorem was presented in [4], you can see it for more details. For general ideas about cluster expansion and related techniques see, for example, [19] and [20]. Note that the condition (III.9) is obviously satisfied when, for example, we have $K_\Gamma \geqslant 0$ for all $\Gamma \in \mathscr{E}$.

III.4. Generalizations to the case of arbitrary finite state space

As shows the Theorem III.5, in the {0,1} case one can specify completely a system of probability distributions consistent in Dobrushin's sense by specifying a Q-function. Clearly this theorem can be reformulated in the terms of H-functions in the following way.

Theorem III.11. — *A system $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ is a system of probability distributions consistent in Dobrushin's sense and satisfying $\mathbf{Q}_\Lambda(\emptyset) > 0$ for all $\Lambda \in \mathscr{E}$ if and only if there exists a H-function $\boldsymbol{H}$ such that for all $\Lambda \in \mathscr{E}$ we have*

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}}{\sum\limits_{\boldsymbol{y} \subset \Lambda} H_{\boldsymbol{y}}}, \quad \boldsymbol{x} \subset \Lambda$$

This version of the theorem is easily generalized to a case of arbitrary finite state space $\mathscr{X}$. That is, in this case one can still specify completely a system of probability distributions consistent in Dobrushin's sense by specifying a suitably defined H-function.

Let us consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$.

Definition III.12. — *A real-valued function* $\boldsymbol{H} = \left\{H_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \mathscr{E}\right\}$ *is called H-function if* $H_{\boldsymbol{\phi}} = 1$ *and* $H_{\boldsymbol{x}} \geqslant 0$ *for all* $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \in \mathscr{E}$.

Theorem III.13. — *A system* $\mathbf{Q} = \{\mathbf{Q}_\Lambda,\ \Lambda \in \mathscr{E}\}$ *is a system of probability distributions consistent in Dobrushin's sense and satisfying* $\mathbf{Q}_\Lambda(\boldsymbol{\phi}) > 0$ *for all* $\Lambda \in \mathscr{E}$ *if and only if there exists a H-function* $\boldsymbol{H}$ *such that for all* $\Lambda \in \mathscr{E}$ *we have*

$$\mathbf{Q}_\Lambda(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}}{\sum\limits_{\boldsymbol{y} \in \mathscr{X}^\Lambda} H_{\boldsymbol{y}}}, \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \Lambda.$$

The proof for this general case is similar to the one corresponding to the {0,1} case.

As in the {0,1} case one can put

$$\theta_\Lambda = \sum_{\boldsymbol{x} \in \mathscr{X}^\Lambda} H_{\boldsymbol{x}}$$

for all $\Lambda \in \mathscr{E}$. The system $\boldsymbol{\theta} = \{\theta_\Lambda,\ \Lambda \in \mathscr{E}\}$ so defined plays again the role of partition function. But unfortunately it no longer determines completely the system of probability distributions consistent in Dobrushin's sense.

All the other properties of Q-functions, and all the properties of H-functions are easily generalized for this general case.

III.5. The problem of uniqueness

So, in this chapter we have seen how a random field (P-function) can be constructed via its conditional distributions in finite volumes with boundary conditions (or, equivalently, Q-function or H-function). The natural questions arise. Is this random field uniquely determined by this Q-function (or H-function), *i.e.*, is it the unique one satisfying (III.8) or there are some other random fields satisfying it too? If no, can one describe the set of all such random fields (may be in some class of random fields or under some conditions) as it was done by Dobrushin in [8] – [10].

Example III.14. — Let us consider the function $\theta_\Lambda \equiv 1$ on $\mathscr{E}$. Obviously this is a Q-function and it satisfies all the conditions of this section (even it has a

cluster expansion with arbitrary small λ). The limiting random field is obviously the random field assuming a.s. the value 0 on $\mathbb{Z}^\nu$ and for it we have

$$\boldsymbol{q}_J^{\emptyset}(\emptyset) = 1, \quad J \in \mathscr{E}. \tag{III.16}$$

But for any $\tau > 0$ the random field from the Example II.4–2 also satisfies the condition (III.16) because for any $J \in \mathscr{E}$ using (II.4) we obtain

$$\boldsymbol{q}_J^{\emptyset}(\emptyset) = \lim_{\overline{J} \uparrow J^{\mathrm{c}}} \frac{\mathbf{P}_{J \cup \overline{J}}(\emptyset)}{\mathbf{P}_{\overline{J}}(\emptyset)} = \lim_{\overline{J} \uparrow J^{\mathrm{c}}} \frac{b_{J \cup \overline{J}}}{b_{\overline{J}}} = \lim_{\overline{J} \uparrow J^{\mathrm{c}}} \frac{\left|\overline{J}\right| + \tau}{\left|J \cup \overline{J}\right| + \tau} = 1.$$

This example shows that in order to answer the questions stated above we need to study more carefully not only the conditional distributions in finite volumes with vacuum boundary conditions but the whole conditional distribution of a random field like it was done by Dobrushin in [8] – [10], *i.e.*, to study specifications rather than systems of probability distributions consistent in Dobrushin's sense.

IV. Vacuum specifications, Q-systems and H-systems

In the previous chapter we have seen that systems of probability distributions consistent in Dobrushin's sense are described by Q-functions and H-functions. In this chapter we show that vacuum specifications can be described by some consistent systems of Q-functions (or, equivalently, of H-functions) which we call Q-systems (H-systems) in approximately the same manner. In the first section we introduce this description for the {0,1} case. In the second section we show that the specifications we describe can be non-Gibbsian and give a general tool for constructing such non-Gibbsian specifications. Particularly, this lets us to show that the random fields from the Example II.4–2 are non-Gibbsian. Finally in the last section we show the way one can generalize the notion of H-systems to the case of arbitrary finite state space $\mathscr{X}$.

IV.1. Q-systems and H-systems

Let us start by giving the following

Definition IV.1. — *A system* $\boldsymbol{\Theta} = \left\{\theta_J^{\overline{\boldsymbol{x}}}, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ *is called* Q*-system if* $\theta_J^{\overline{\boldsymbol{x}}} \neq 0$ *for all* $J \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset J^{\mathrm{c}}$*, if* $\theta_{\emptyset}^{\overline{\boldsymbol{x}}} = 1$ *for all* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu}$ *and if for any* $S \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset S^{\mathrm{c}}$ *we have*

$$\sum_{J \subset S} (-1)^{|S \setminus J|} \theta_J^{\overline{\boldsymbol{x}}} \geqslant 0. \tag{IV.1}$$

Just like Q-functions, Q-systems have the following simple constructive description.

Theorem IV.2. — *A system* $\boldsymbol{\Theta} = \left\{\theta_J^{\overline{\boldsymbol{x}}}, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ *is a* Q*-system if and only if there exists a system* $\boldsymbol{\mathcal{H}} = \left\{H_S^{\overline{\boldsymbol{x}}}, \quad S \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset S^{\mathrm{c}}\right\}$*,* $H_S^{\overline{\boldsymbol{x}}} \geqslant 0$ *for all* $S \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset S^{\mathrm{c}}$*,* $H_{\emptyset}^{\overline{\boldsymbol{x}}} = 1$ *for all* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu}$*, such that for any* $\Lambda \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ *we have*

$$\theta_{\Lambda}^{\overline{\boldsymbol{x}}} = \sum_{S \subset \Lambda} H_S^{\overline{\boldsymbol{x}}}. \tag{IV.2}$$

This system $\boldsymbol{\mathcal{H}}$ *is called* H*-system.*

Proof: **1)** Necessity. Let $\mathbf{\Theta} = \left\{\theta^{\overline{\boldsymbol{x}}}_J, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ be a Q-system. Put

$$H^{\overline{\boldsymbol{x}}}_S = \sum_{J\subset S}(-1)^{|S\setminus J|}\theta^{\overline{\boldsymbol{x}}}_J, \quad S \in \mathscr{E},\ \overline{\boldsymbol{x}} \subset S^{\mathrm{c}}. \tag{IV.3}$$

Since θ is a Q-system and according to the definition (IV.3) of $H^{\overline{\boldsymbol{x}}}_S$, we have $H^{\overline{\boldsymbol{x}}}_{\emptyset} = 1$ for all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu$ and $H^{\overline{\boldsymbol{x}}}_S \geqslant 0$ for all $S \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset S^{\mathrm{c}}$. Further, for any $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ we can write

$$\sum_{S\subset\Lambda} H^{\overline{\boldsymbol{x}}}_S = \sum_{S\subset\Lambda}\sum_{J\subset S}(-1)^{|S\setminus J|}\theta^{\overline{\boldsymbol{x}}}_J = \sum_{J\subset\Lambda}\theta^{\overline{\boldsymbol{x}}}_J \sum_{S\,:\,J\subset S\subset\Lambda}(-1)^{|S\setminus J|} = \theta^{\overline{\boldsymbol{x}}}_\Lambda.$$

2) Sufficiency. Let $\boldsymbol{\mathcal{H}}$ be a H-system and $\theta^{\overline{\boldsymbol{x}}}_\Lambda = \sum\limits_{S\subset\Lambda} H^{\overline{\boldsymbol{x}}}_S$. Clearly $\theta^{\overline{\boldsymbol{x}}}_{\emptyset} = H^{\overline{\boldsymbol{x}}}_{\emptyset} = 1$ for any $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu$ and $\theta^{\overline{\boldsymbol{x}}}_\Lambda \geqslant H^{\overline{\boldsymbol{x}}}_{\emptyset} = 1 > 0$ for any $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$. Finally, for all $S \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset S^{\mathrm{c}}$ we have

$$\begin{aligned}\sum_{J\subset S}(-1)^{|S\setminus J|}\theta^{\overline{\boldsymbol{x}}}_J &= \sum_{J\subset S}(-1)^{|S\setminus J|}\sum_{\widetilde{J}\subset J} H^{\overline{\boldsymbol{x}}}_{\widetilde{J}} = \\ &= \sum_{\widetilde{J}\subset S} H^{\overline{\boldsymbol{x}}}_{\widetilde{J}} \sum_{J\,:\,\widetilde{J}\subset J\subset S}(-1)^{|S\setminus J|} = H^{\overline{\boldsymbol{x}}}_S \geqslant 0\end{aligned}$$

which concludes the proof. □

The motivation of introducing Q-systems and H-systems is the fact that they describe vacuum specifications in approximately the same manner in which Q-functions and H-functions describe systems of probability distributions in finite volumes consistent in Dobrushin's sense.

Definition IV.3. — *A H-system $\boldsymbol{\mathcal{H}} = \left\{H^{\overline{\boldsymbol{x}}}_S, \quad S \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset S^{\mathrm{c}}\right\}$ is called consistent if it satisfies the following condition: for any $S_1, S_2 \in \mathscr{E}$ such that $S_1 \cap S_2 = \emptyset$ and any $\overline{\boldsymbol{x}} \subset (S_1 \cup S_2)^{\mathrm{c}}$ we have*

$$H^{\overline{\boldsymbol{x}}}_{S_1\cup S_2} = H^{\overline{\boldsymbol{x}}}_{S_1}\, H^{\overline{\boldsymbol{x}}\cup S_1}_{S_2}. \tag{IV.4}$$

A Q-system $\mathbf{\Theta} = \left\{\theta^{\overline{\boldsymbol{x}}}_J, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ is called consistent, if the corresponding H-system is consistent.

Theorem IV.4. — *A system $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\right\}$ is a vacuum specification if and only if there exists a consistent Q-system*

$\boldsymbol{\Theta} = \left\{\theta^{\overline{\boldsymbol{x}}}_J, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\rm c}\right\}$ *such that for any* $\Lambda \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset \Lambda^{\rm c}$ *we have*

$$\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta^{\overline{\boldsymbol{x}}}_\Lambda} \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} \theta^{\overline{\boldsymbol{x}}}_J, \quad \boldsymbol{x} \subset \Lambda. \tag{IV.5}$$

Particularly, for any $\Lambda \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset \Lambda^{\rm c}$ *we have* $\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset) = 1/\theta^{\overline{\boldsymbol{x}}}_\Lambda$.

Proof: **1)** Necessity. Let $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\rm c}\right\}$ be a specification with $\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset) > 0$ for all $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\rm c}$. Put $\theta^{\overline{\boldsymbol{x}}}_\Lambda = 1/\,\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset)$. We have obviously $\theta^{\overline{\boldsymbol{x}}}_\Lambda \neq 0$ and $\theta^{\overline{\boldsymbol{x}}}_\emptyset = 1$. Further, for any $\Lambda \in \mathscr{E}$, $J \subset \Lambda$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\rm c}$ we can write

$$1 = \sum_{S \subset J} \mathbf{Q}^{\overline{\boldsymbol{x}}}_J(S) = \sum_{S \subset J} \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_J(\emptyset)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset)}\, \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S) = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_J(\emptyset)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset)} \sum_{S \subset J} \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S)$$

or equivalently

$$\theta^{\overline{\boldsymbol{x}}}_J = \theta^{\overline{\boldsymbol{x}}}_\Lambda \sum_{S \subset J} \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S).$$

Therefore

$$\sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} \theta^{\overline{\boldsymbol{x}}}_J = \theta^{\overline{\boldsymbol{x}}}_\Lambda \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} \sum_{S \subset J} \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S) = \theta^{\overline{\boldsymbol{x}}}_\Lambda\, \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\boldsymbol{x})$$

and we obtain (IV.1) and (IV.5). It remains to verify the consistency of the Q-system $\boldsymbol{\Theta}$. Let $\boldsymbol{\mathcal{H}} = \left\{H^{\overline{\boldsymbol{x}}}_S, \quad S \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset S^{\rm c}\right\}$ be the H-system corresponding to this Q-system and let us fix some $S_1, S_2 \in \mathscr{E}$ such that $S_1 \cap S_2 = \emptyset$ and some $\overline{\boldsymbol{x}} \subset (S_1 \cup S_2)^{\rm c}$. On one hand we have

$$H^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2} = \sum_{J \subset S_1 \cup S_2} (-1)^{\left|(S_1 \cup S_2)\setminus J\right|}\, \theta^{\overline{\boldsymbol{x}}}_J = \theta^{\overline{\boldsymbol{x}}}_\Lambda\, \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S_1 \cup S_2) = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(S_1 \cup S_2)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(\emptyset)}$$

where we have chosen $\Lambda = S_1 \cup S_2$. On the other hand, we obtain in the similar manner the equalities

$$H^{\overline{\boldsymbol{x}}}_{S_1} = \sum_{J \subset S_1} (-1)^{|S_1 \setminus J|} \theta^{\overline{\boldsymbol{x}}}_J = \theta^{\overline{\boldsymbol{x}}}_\Lambda\, \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(S_1) = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(S_1)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(\emptyset)}\ ,$$

$$H^{\overline{\boldsymbol{x}} \cup S_1}_{S_2} = \sum_{J \subset S_2} (-1)^{|S_2 \setminus J|} \theta^{\overline{\boldsymbol{x}} \cup S_1}_J = \theta^{\overline{\boldsymbol{x}} \cup S_1}_\Lambda\, \mathbf{Q}^{\overline{\boldsymbol{x}} \cup S_1}_\Lambda(S_2) = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}} \cup S_1}_{S_2}(S_2)}{\mathbf{Q}^{\overline{\boldsymbol{x}} \cup S_1}_{S_2}(\emptyset)}\ ,$$

and hence we have

$$H^{\overline{\boldsymbol{x}}}_{S_1}\, H^{\overline{\boldsymbol{x}} \cup S_1}_{S_2} = \frac{1}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(\emptyset)}\, \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(S_1)}{\mathbf{Q}^{\overline{\boldsymbol{x}} \cup S_1}_{S_2}(\emptyset)}\, \mathbf{Q}^{\overline{\boldsymbol{x}} \cup S_1}_{S_2}(S_2) = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(S_1 \cup S_2)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}(\emptyset)} = H^{\overline{\boldsymbol{x}}}_{S_1 \cup S_2}$$

which concludes the proof of necessity.

2) Sufficiency. Let $\boldsymbol{\Theta} = \big\{\theta^{\overline{\boldsymbol{x}}}_J, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\big\}$ be a consistent Q-system. First of all, let us note that for all $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ we have $\theta^{\overline{\boldsymbol{x}}}_\Lambda = \sum\limits_{S\subset\Lambda} H^{\overline{\boldsymbol{x}}}_S \geqslant H^{\overline{\boldsymbol{x}}}_\emptyset = 1 > 0$ and for all $\Lambda \in \mathscr{E}$, $\boldsymbol{x} \subset \Lambda$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ put

$$\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta^{\overline{\boldsymbol{x}}}_\Lambda} \sum_{J\subset\boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} \theta^{\overline{\boldsymbol{x}}}_J = \frac{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}}{\theta^{\overline{\boldsymbol{x}}}_\Lambda} \geqslant 0.$$

Now let us prove that $\boldsymbol{\mathcal{Q}} = \big\{\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\big\}$ is a specification. We have

$$\sum_{\boldsymbol{x}\subset\Lambda} \mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\boldsymbol{x}) = \frac{1}{\theta^{\overline{\boldsymbol{x}}}_\Lambda} \sum_{\boldsymbol{x}\subset\Lambda} H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = \frac{1}{\theta^{\overline{\boldsymbol{x}}}_\Lambda}\, \theta^{\overline{\boldsymbol{x}}}_\Lambda = 1,$$

i.e., the system $\mathbf{Q}$ is a system of probability distributions in finite volumes with boundary conditions. It remains to verify that the condition (I.8) is satisfied. We have

$$\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}(\boldsymbol{x}\cup\boldsymbol{y}) = \frac{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}\cup\boldsymbol{y}}}{\theta^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}} = \frac{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{y}}\, H^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\boldsymbol{x}}}{\theta^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}} = \frac{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{y}}}{\theta^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}}\, \frac{H^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\boldsymbol{x}}}{\theta^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\Lambda}}\, \theta^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\Lambda} = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}(\boldsymbol{y})\ \mathbf{Q}^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\Lambda}(\boldsymbol{x})}{\mathbf{Q}^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\Lambda}(\emptyset)}\,.$$

The theorem is proved. □

Remark IV.5. — Let us denote $U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = -\ln H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}$ for all $\boldsymbol{x} \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}$ where we permit the system $\boldsymbol{\mathcal{U}}$ to take the value $+\infty$. Then clearly the system $\boldsymbol{\mathcal{U}} = \big\{U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}\big\}$ satisfies the following consistency property: for all $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{E}$ such that $\boldsymbol{x} \cap \boldsymbol{y} = \emptyset$ and all $\overline{\boldsymbol{x}} \subset (\boldsymbol{x}\cup\boldsymbol{y})^{\mathrm{c}}$ we have

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}\cup\boldsymbol{y}) = U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) + U^{\overline{\boldsymbol{x}}\cup\boldsymbol{x}}(\boldsymbol{y}). \tag{IV.6}$$

Now, using the formulas (IV.2) and (IV.3) we can rewrite (IV.5) in the form

$$\mathbf{Q}^{\overline{\boldsymbol{x}}}_\Lambda(\boldsymbol{x}) = \frac{\exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\big)}{\sum\limits_{\boldsymbol{y}\subset\Lambda} \exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{y})\big)}\,, \quad \Lambda \in \mathscr{E},\ \boldsymbol{x} \subset \Lambda,\ \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}.$$

So, we see that our specifications are similar to the usual Gibbsian specifications with only difference that in our case the Hamiltonian $\boldsymbol{\mathcal{U}}$ is an arbitrary system satisfying the condition (IV.6), while in the Gibbsian case it has an explicit form in terms of an interaction potential. Note that in the Gibbsian case the condition (IV.6) is automatically satisfied.

IV.2. Non-Gibbsian random fields

In this section we will show that in our case the specifications may be non-Gibbsian and will describe a simple general scheme for constructing such non-Gibbsian specifications. For this we need the following

Lemma IV.6. — *Let* $\boldsymbol{\Theta} = \left\{\theta_J^{\overline{\boldsymbol{x}}}, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ *be a consistent* Q*-system,* $\boldsymbol{\mathcal{H}} = \left\{H_S^{\overline{\boldsymbol{x}}}, \quad S \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset S^{\mathrm{c}}\right\}$ *be the corresponding consistent* H*-system and* $\boldsymbol{R} = \{R(\overline{\boldsymbol{x}}),\ \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu\}$ *be a real-valued strictly positive function such that* $R(\overline{\boldsymbol{x}}_1) = R(\overline{\boldsymbol{x}}_2)$ *if* $\overline{\boldsymbol{x}}_1 = \overline{\boldsymbol{x}}_2$ *up to a finite number of lattice points. Then the system*

$$\boldsymbol{\mathcal{H}}_{\boldsymbol{R}} = \left\{\left(H_S^{\overline{\boldsymbol{x}}}\right)^{R(\overline{\boldsymbol{x}})}, \quad S \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset S^{\mathrm{c}}\right\}$$

is a consistent H*-system and hence determines some consistent* Q*-system which we denote by* $\boldsymbol{\Theta}_{\boldsymbol{R}}$.

Proof: For any $S_1, S_2 \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset (S_1 \cup S_2)^{\mathrm{c}}$ we can write

$$\begin{aligned}\left(H_{S_1\cup S_2}^{\overline{\boldsymbol{x}}}\right)^{R(\overline{\boldsymbol{x}})} &= \left(H_{S_1}^{\overline{\boldsymbol{x}}}\ H_{S_2}^{\overline{\boldsymbol{x}}\cup S_1}\right)^{R(\overline{\boldsymbol{x}})} = \\ &= \left(H_{S_1}^{\overline{\boldsymbol{x}}}\right)^{R(\overline{\boldsymbol{x}})} \left(H_{S_2}^{\overline{\boldsymbol{x}}\cup S_1}\right)^{R(\overline{\boldsymbol{x}})} = \\ &= \left(H_{S_1}^{\overline{\boldsymbol{x}}}\right)^{R(\overline{\boldsymbol{x}})} \left(H_{S_2}^{\overline{\boldsymbol{x}}\cup S_1}\right)^{R(\overline{\boldsymbol{x}}\cup S_1)}\end{aligned}$$

which concludes the proof. □

Remark IV.7. — We require the function $\boldsymbol{R}$ to be real-valued and strictly positive only in order for the system $\boldsymbol{\mathcal{H}}_{\boldsymbol{R}}$ to be well-defined. But the lemma holds under less restrictive conditions. For example, if the system $\boldsymbol{\mathcal{H}}$ is strictly positive, which is equivalent to say that the corresponding Hamiltonian $\boldsymbol{\mathcal{U}}$ is finite, we can consider $\boldsymbol{R}$ to be any real-valued function, and if the system $\boldsymbol{\mathcal{H}}$ is less or equal than 1 (respectively greater or equal than 1), which is equivalent to say that the Hamiltonian $\boldsymbol{\mathcal{U}}$ is positive (respectively negative), we can allow $\boldsymbol{R}$ to take the value $+\infty$ (respectively $-\infty$). Here and in the sequel we admit that $\alpha^{+\infty} = 0$ for any $0 \leqslant \alpha < 1$, that $\beta^{-\infty} = 0$ for any $\beta > 1$ and that $1^{\pm\infty} = 0^0 = 1$ $\big($note that it is equivalent to admitting that $(\pm\infty)\cdot a = a\cdot(\pm\infty) = \pm\infty$ for any $a > 0$, that $(\pm\infty)\cdot b = b\cdot(\pm\infty) = \mp\infty$ for any $b < 0$ and that $(\pm\infty)\cdot 0 = 0\cdot(\pm\infty) = 0\big)$.

Proposition IV.8. — *Let* $\boldsymbol{\Theta} = \left\{\theta_J^{\overline{\boldsymbol{x}}}, \quad J \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset J^{\mathrm{c}}\right\}$ *be a Gibbsian* Q*-system corresponding to a finite Hamiltonian* $\boldsymbol{\mathcal{U}} = \left\{U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}\right\}$

and $\boldsymbol{R} = \{R(\overline{\boldsymbol{x}}),\ \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu\}$ *be a real-valued function such that* $R(\overline{\boldsymbol{x}}_1) = R(\overline{\boldsymbol{x}}_2)$ *if* $\overline{\boldsymbol{x}}_1 = \overline{\boldsymbol{x}}_2$ *up to a finite number of lattice points. We suppose that the following condition holds: there exist at least one pair* $\boldsymbol{x} \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}$ *such that* $R(\overline{\boldsymbol{x}}) \neq R(\emptyset)$ *and that* $U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \neq 0$*. Then the specification determined by the Q-system* $\boldsymbol{\Theta}_{\boldsymbol{R}}$ *is non-Gibbsian.*

Proof: Since $\boldsymbol{\Theta}$ is Gibbsian, the corresponding H-system $\boldsymbol{\mathcal{H}}$ has the form

$$\boldsymbol{\mathcal{H}} = \Big\{\exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\big), \quad \boldsymbol{x} \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}\Big\}$$

where the Hamiltonian $\boldsymbol{\mathcal{U}}$ is given by some potential $\boldsymbol{\Phi} = \big\{\Phi(J),\ J \in \mathscr{E} \setminus \{\emptyset\}\big\}$. Hence

$$\boldsymbol{\mathcal{H}}_{\boldsymbol{R}} = \Big\{\exp\big(-U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\, R(\overline{\boldsymbol{x}})\big), \quad \boldsymbol{x} \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}\Big\}.$$

We need to show that the specification determined by $\boldsymbol{\mathcal{H}}_{\boldsymbol{R}}$ is non-Gibbsian, *i.e.*, that there exist no convergent potential $\widetilde{\boldsymbol{\Phi}} = \big\{\widetilde{\Phi}(J),\ J \in \mathscr{E} \setminus \{\emptyset\}\big\}$ such that

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\, R(\overline{\boldsymbol{x}}) = \sum_{J\,:\,\emptyset \neq J \subset \boldsymbol{x}} \ \sum_{\widetilde{J} \in \mathscr{E}\,:\,\widetilde{J} \subset \overline{\boldsymbol{x}}} \widetilde{\Phi}\big(J \cup \widetilde{J}\big), \quad \boldsymbol{x} \in \mathscr{E},\ \overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}. \tag{IV.7}$$

Suppose that the contrary is true, *i.e.*, that (IV.7) holds. In this case we would clearly have

$$U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})\, R(\overline{\boldsymbol{x}}) = \lim_{I \uparrow \mathbb{Z}^\nu} U^{\overline{\boldsymbol{x}}_I}(\boldsymbol{x})\, R(\overline{\boldsymbol{x}}_I) = R(\emptyset) \lim_{I \uparrow \mathbb{Z}^\nu} U^{\overline{\boldsymbol{x}}_I}(\boldsymbol{x}) = R(\emptyset)\, U^{\overline{\boldsymbol{x}}}(\boldsymbol{x})$$

for any $\boldsymbol{x} \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}$. But the last relation contradicts with the conditions of the proposition. □

Remarks IV.9. — **1)** Clearly, as in the Lemma IV.6 we can allow $\boldsymbol{R}$ to take the value $+\infty$ or $-\infty$ under suitable conditions.

2) Let us denote $\mathfrak{N} = \big\{\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \;\big|\; \exists\ \boldsymbol{x} \in \mathscr{E}$ such that $\boldsymbol{x} \subset \overline{\boldsymbol{x}}^{\,\mathrm{c}}$ and $U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \neq 0\big\}$. It is not difficult to check that the condition of the Proposition IV.8 holds if and only if the function $\boldsymbol{R} = \{R(\overline{\boldsymbol{x}}),\ \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu\}$ is not constant on $\mathfrak{N}$. The sufficiency is evident. For the proof of necessity note that as we know that there exists a pair $\boldsymbol{x} \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \boldsymbol{x}^{\mathrm{c}}$ such that $R(\overline{\boldsymbol{x}}) \neq R(\emptyset)$ and that $U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \neq 0$, then clearly we have $\overline{\boldsymbol{x}} \in \mathfrak{N}$ and also $\emptyset \in \mathfrak{N}$, since otherwise we would have $U^{\emptyset}(\boldsymbol{x}) = 0$ for all $x \in \mathscr{E}$ which is possible if and only if $\boldsymbol{\Phi} \equiv 0$ on $\mathscr{E} \setminus \{\emptyset\}$ which contradicts with $U^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \neq 0$.

3) If the specification $\boldsymbol{\mathcal{Q}}$ corresponding to the Q-system $\boldsymbol{\Theta}_{\boldsymbol{R}}$ is a conditional distribution of some random field $\mathbf{P}$ and if the function $\boldsymbol{R} = \{R(\overline{\boldsymbol{x}}),\ \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu\}$

is not $\mathbf{P}$-almost surely constant on $\mathfrak{N}$, then this random field $\mathbf{P}$ is clearly non-Gibbsian, *i.e.*, any conditional distribution $\widetilde{\boldsymbol{\mathcal{Q}}}$ of $\mathbf{P}$ is not a Gibbsian specification.

As we see, the Proposition IV.8 is a powerful tool for constructing non-Gibbsian specifications and random fields. Note that non-Gibbsian specifications and random fields constructed this way are not quasilocal. Note also that the Proposition IV.8 can also be very useful for verifying that a given specification or random field is non-Gibbsian. For example, let us verify that the random fields considered in the Example II.4–2 are non-Gibbsian for all $\tau > 0$. For this, let us fix some $\tau > 0$ and calculate the conditional distributions of the random field $\mathbf{P}$ corresponding to this τ. For any $p \in [0,1]$ let us denote by $\mathfrak{I}^p$ the set of all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu$ such that

$$\exists \lim_{I \uparrow \mathbb{Z}^\nu} \frac{|\overline{\boldsymbol{x}}_I|}{|I|} = p(\overline{\boldsymbol{x}}) = p,$$

and put $\overline{\mathfrak{I}} = \mathscr{X}^{\mathbb{Z}^\nu} \Big\backslash \Big(\bigcup\limits_{p \in [0,1]} \mathfrak{I}^p \Big)$. Note that for any $p \in [0,1]$ the set $\mathfrak{I}^p$ has measure 1 with respect to the Bernoulli random field with parameter p and measure 0 with respect to all the other Bernoulli random fields. Hence, each of the sets $\mathfrak{I}^p$ and the set $\overline{\mathfrak{I}}$ have measure 0 with respect to the random field $\mathbf{P}$. Now let us take some $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ such that $\overline{\boldsymbol{x}} \notin \overline{\mathfrak{I}}$ and calculate the limit

$$\boldsymbol{q}^{\overline{\boldsymbol{x}}}_\Lambda(\emptyset) = \lim_{I \uparrow \Lambda^{\mathrm{c}}} \frac{\mathbf{P}_{\Lambda \cup I}(\overline{\boldsymbol{x}}_I)}{\mathbf{P}_I(\overline{\boldsymbol{x}}_I)} = \lim_{I \uparrow \Lambda^{\mathrm{c}}} \frac{|I| + \tau}{|\Lambda| + |I| + \tau} \prod_{i=1}^{|\overline{\boldsymbol{x}}_I|} \frac{|I| + \tau - i}{|\Lambda| + |I| + \tau - i} =$$

$$= \lim_{I \uparrow \Lambda^{\mathrm{c}}} \frac{\prod\limits_{i=1+|\Lambda|}^{|\overline{\boldsymbol{x}}_I|+|\Lambda|} \Big(|\Lambda| + |I| + \tau - i \Big)}{\prod\limits_{i=1}^{|\overline{\boldsymbol{x}}_I|} \Big(|\Lambda| + |I| + \tau - i \Big)} = \lim_{I \uparrow \Lambda^{\mathrm{c}}} \frac{\prod\limits_{i=1+|\overline{\boldsymbol{x}}_I|}^{|\overline{\boldsymbol{x}}_I|+|\Lambda|} \Big(|\Lambda| + |I| + \tau - i \Big)}{\prod\limits_{i=1}^{|\Lambda|} \Big(|\Lambda| + |I| + \tau - i \Big)} =$$

$$= \lim_{J \uparrow \mathbb{Z}^\nu} \frac{\prod\limits_{i=1}^{|\Lambda|} \Big(|J| + \tau - i - |\overline{\boldsymbol{x}}_J| \Big)}{\prod\limits_{i=1}^{|\Lambda|} \Big(|J| + \tau - i \Big)} = \prod_{i=1}^{|\Lambda|} \big(1 - p(\overline{\boldsymbol{x}}) \big) = \big(1 - p(\overline{\boldsymbol{x}}) \big)^{|\Lambda|}.$$

Note that this limit is strictly positive if $0 \leqslant p(\overline{\boldsymbol{x}}) < 1$ and that $\mathbf{P}\big(\mathfrak{I}^1 \cup \overline{\mathfrak{I}}\big) = 0$, and hence putting

$$\theta^{\overline{\boldsymbol{x}}}_\Lambda = \begin{cases} 1 & \text{if } \overline{\boldsymbol{x}} \in \mathfrak{I}^1 \cup \overline{\mathfrak{I}}, \\ \big(1 - p(\overline{\boldsymbol{x}}) \big)^{-|\Lambda|} & \text{otherwise,} \end{cases}$$

we obtain a Q-system $\boldsymbol{\Theta}$ corresponding to some specification $\boldsymbol{\mathcal{Q}}$ which is a conditional distribution of the random field $\mathbf{P}$. In order to write down explicitly this specification $\boldsymbol{\mathcal{Q}}$ let us at first calculate the corresponding H-system $\boldsymbol{\mathcal{H}}$. For $\overline{\boldsymbol{x}} \in \mathfrak{I}^1 \cup \overline{\mathfrak{I}}$ we have clearly $H^{\overline{\boldsymbol{x}}}_{\phi} = 1$ and $H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = 0$ for $\boldsymbol{x} \neq \phi$ or, otherwise speaking, $H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = \mathbb{1}_{\{\boldsymbol{x}=\phi\}}$. For $\overline{\boldsymbol{x}} \notin \mathfrak{I}^1 \cup \overline{\mathfrak{I}}$ we can write

$$\begin{aligned} H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} &= \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \theta^{\overline{\boldsymbol{x}}}_{J} = \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \big(1 - p(\overline{\boldsymbol{x}})\big)^{-|J|} = \\ &= \big(1 - p(\overline{\boldsymbol{x}})\big)^{-|\boldsymbol{x}|} \sum_{J \subset \boldsymbol{x}} (-1)^{|\boldsymbol{x} \setminus J|} \big(1 - p(\overline{\boldsymbol{x}})\big)^{|\boldsymbol{x}|-|J|} = \\ &= \frac{\sum\limits_{J \subset \boldsymbol{x}} \big(p(\overline{\boldsymbol{x}}) - 1\big)^{|\boldsymbol{x} \setminus J|}}{\big(1 - p(\overline{\boldsymbol{x}})\big)^{|\boldsymbol{x}|}} = \left(\frac{p(\overline{\boldsymbol{x}})}{1 - p(\overline{\boldsymbol{x}})}\right)^{|\boldsymbol{x}|} \end{aligned}$$

where we have used the combinatorial version of binomial formula. So the system $\boldsymbol{\mathcal{H}}$ has the form:

$$H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = \begin{cases} \mathbb{1}_{\{\boldsymbol{x}=\phi\}} & \text{if } \overline{\boldsymbol{x}} \in \mathfrak{I}^1 \cup \overline{\mathfrak{I}}, \\ \left(\frac{p(\overline{\boldsymbol{x}})}{1-p(\overline{\boldsymbol{x}})}\right)^{|\boldsymbol{x}|} & \text{otherwise,} \end{cases}$$

and hence the specification $\boldsymbol{\mathcal{Q}}$ is given by

$$\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda}(\boldsymbol{x}) = \frac{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}}{\theta^{\overline{\boldsymbol{x}}}_{\Lambda}} = \begin{cases} \mathbb{1}_{\{\boldsymbol{x}=\phi\}} & \text{if } \overline{\boldsymbol{x}} \in \mathfrak{I}^1 \cup \overline{\mathfrak{I}}, \\ \big(p(\overline{\boldsymbol{x}})\big)^{|\boldsymbol{x}|} \big(1 - p(\overline{\boldsymbol{x}})\big)^{|\Lambda \setminus \boldsymbol{x}|} & \text{otherwise.} \end{cases}$$

Now, let us remark that the system $\boldsymbol{\mathcal{H}}$ can be rewritten in the form $H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = \big(\widetilde{H}^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}\big)^{R(\overline{\boldsymbol{x}})}$ where $\widetilde{H}^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}} = e^{-|\boldsymbol{x}|}$ is the Gibbsian H-system correponding to the potential $\boldsymbol{\Phi} = \Big\{\Phi(J) = \mathbb{1}_{\{|J|=1\}},\ J \in \mathscr{E} \setminus \{\phi\}\Big\}$ and the function $\boldsymbol{R}$ is given by

$$R(\overline{\boldsymbol{x}}) = \begin{cases} +\infty & \text{if } \overline{\boldsymbol{x}} \in \mathfrak{I}^1 \cup \overline{\mathfrak{I}}, \\ -\ln \frac{p(\overline{\boldsymbol{x}})}{1-p(\overline{\boldsymbol{x}})} & \text{otherwise.} \end{cases}$$

Clearly the conditions of the Proposition IV.8 are satisfied and hence the specification $\boldsymbol{\mathcal{Q}}$ is non-Gibbsian. Moreover, according to the Remark IV.9–3 the random field $\mathbf{P}$ is also non-Gibbsian.

IV.3. Generalizations to the case of arbitrary finite state space

The generalization is done just in the same way it was done for H-functions. First off all we note the Theorem IV.4, in the $\{0,1\}$ case shows that one can specify completely a vacuum specification by specifying a Q-system. Clearly this theorem can be reformulated in the terms of H-systems in the following way.

Theorem IV.10. — *A system* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\right\}$ *is a vacuum specification if and only if there exists a consistent H-system* $\boldsymbol{\mathcal{H}}$ *such that for any* $\Lambda \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ *we have*

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}}{\sum\limits_{\boldsymbol{y}\subset\Lambda} H_{\boldsymbol{y}}^{\overline{\boldsymbol{x}}}}, \quad \boldsymbol{x} \subset \Lambda.$$

This version of the theorem is easily generalized to a case of arbitrary finite state space $\mathscr{X}$. That is, in this case one can still specify completely a vacuum specification by specifying a suitably defined consistent H-system.

Let us consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$.

Definition IV.11. — *Let* $\boldsymbol{\mathcal{H}} = \left\{H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}, \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \mathscr{E},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{*K},\ K \subset I^{\mathrm{c}}\right\}$ *be some real-valued function. It is called H-system if* $H_{\boldsymbol{\emptyset}}^{\overline{\boldsymbol{x}}} = 1$ *for all* $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset \mathbb{Z}^\nu$ *and* $H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}} \geqslant 0$ *for all* $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \in \mathscr{E}$, $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset I^{\mathrm{c}}$. *This H-system is called consistent if it satisfies the following condition: for any* $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \in \mathscr{E}$, $\boldsymbol{y} \in \mathscr{X}^{*J}$, $J \in \mathscr{E}$ *such that* $I \cap J = \emptyset$ *and any* $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset (I \cup J)^{\mathrm{c}}$ *we have*

$$H_{\boldsymbol{x}\oplus\boldsymbol{y}}^{\overline{\boldsymbol{x}}} = H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}\, H_{\boldsymbol{y}}^{\overline{\boldsymbol{x}}\oplus\boldsymbol{x}}.$$

Theorem IV.12. — *A system* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}\right\}$ *is a vacuum specification if and only if there exists a consistent H-system* $\boldsymbol{\mathcal{H}}$ *such that for any* $\Lambda \in \mathscr{E}$ *and* $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}$ *we have*

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}}{\sum\limits_{\boldsymbol{y}\in\mathscr{X}^\Lambda} H_{\boldsymbol{y}}^{\overline{\boldsymbol{x}}}}, \quad \boldsymbol{x} \in \mathscr{X}^\Lambda.$$

The proof for this general case is similar to the one corresponding to the {0,1} case.

As in the {0,1} case one can put

$$\theta_\Lambda^{\overline{\boldsymbol{x}}} = \sum_{\boldsymbol{x} \in \mathscr{X}^\Lambda} H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}$$

for all $\Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}$. The system $\theta = \left\{\theta_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E},\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^c}\right\}$ so defined plays again the role of partition function. But unfortunately it no longer determines completely the specification.

All the other results concerning Q-systems and H-systems are easily generalized for this general case.

V. Vacuum specifications and one-point systems

As we have seen in the previous chapter, consistent Q-systems and H-systems are convenient tools for description of vacuum specifications. But this systems are "too rich", since taking a closer look on the consistency condition (IV.4) we see that the information contained in a H-system is redundant, and hence one can think about describing specifications by more simple systems than Q-systems or H-systems. In fact, in this chapter we will show that one can describe vacuum specifications by one-point subsystems of consistent H-systems, and in the next chapter we will consider in more details the case of quasilocal specifications and will show that in this case one can describe specifications by H-functions or Q-functions satisfying some additional conditions. In the first section we consider the $\{0,1\}$ case. In the second section we give a necessary and sufficient condition for a one-point system to be Gibbsian. Finally in the last section we show the way one can generalize the notion of one-point systems to the case of arbitrary finite state space $\mathscr{X}$.

V.1. One-point systems

We start by introducing the following

Definition V.1. — *A system* $\boldsymbol{h} = \left\{h_t^{\overline{\boldsymbol{x}}}, \quad t \in \mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t\right\}$ *is called one-point system if for all* $t \in \mathbb{Z}^\nu$ *and* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ *we have* $h_t^{\overline{\boldsymbol{x}}} \geqslant 0$ *and for all* $s, t \in \mathbb{Z}^\nu$ *and* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \{s,t\}$ *we have*

$$h_s^{\overline{\boldsymbol{x}}}\, h_t^{\overline{\boldsymbol{x}} \cup s} = h_t^{\overline{\boldsymbol{x}}}\, h_s^{\overline{\boldsymbol{x}} \cup t}. \tag{V.1}$$

As shows the following theorem these one-point systems correspond one-to-one to consistent H-systems. In fact they are nothing but one-point subsystems of consistent H-systems and hence, just like H-systems, describe vacuum specifications.

Theorem V.2. — *A system* $\boldsymbol{\mathcal{H}}=\bigl\{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}},\ \boldsymbol{x}\in\mathscr{E}$ *and* $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}\bigr\}$ *is a consistent* H*-system if and only if there exists a one-point system* $\boldsymbol{h}=\bigl\{h^{\overline{\boldsymbol{x}}}_{t},\ t\in\mathbb{Z}^{\nu}$ *and* $\overline{\boldsymbol{x}}\subset\mathbb{Z}^{\nu}\setminus t\bigr\}$ *such that for all* $\boldsymbol{x}\in\mathscr{E}$ *and* $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}$ *we have*

$$H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}=h^{\overline{\boldsymbol{x}}}_{t_1}\,h^{\overline{\boldsymbol{x}}\cup t_1}_{t_2}\,\cdots\,h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_{n-1}}_{t_n} \tag{V.2}$$

where $n=|\boldsymbol{x}|$ *and* $t_1,\dots,t_n$ *is some arbitrary enumeration of elements of the set* $\boldsymbol{x}$*. Particularly, for all* $t\in\mathbb{Z}^{\nu}$ *and* $\overline{\boldsymbol{x}}\subset\mathbb{Z}^{\nu}\setminus t$ *we have* $H^{\overline{\boldsymbol{x}}}_{t}=h^{\overline{\boldsymbol{x}}}_{t}$*.*

Proof: **1)** Necessity. Let $\boldsymbol{\mathcal{H}}=\bigl\{H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}},\quad\boldsymbol{x}\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}\bigr\}$ be a consistent H-system and put $h^{\overline{\boldsymbol{x}}}_{t}=H^{\overline{\boldsymbol{x}}}_{t}\geqslant 0$ for all $t\in\mathbb{Z}^{\nu}$ and $\overline{\boldsymbol{x}}\subset\mathbb{Z}^{\nu}\setminus t$. Since H-system $\boldsymbol{\mathcal{H}}$ is consistent, using the formula (IV.4) we obtain

$$H^{\overline{\boldsymbol{x}}}_{\{s,t\}}=H^{\overline{\boldsymbol{x}}}_{s}\,H^{\overline{\boldsymbol{x}}\cup s}_{t}=h^{\overline{\boldsymbol{x}}}_{s}\,h^{\overline{\boldsymbol{x}}\cup s}_{t}.$$

In the same manner $H^{\overline{\boldsymbol{x}}}_{\{s,t\}}=h^{\overline{\boldsymbol{x}}}_{t}\,h^{\overline{\boldsymbol{x}}\cup t}_{s}$, and hence $\boldsymbol{h}$ is a one-point system. Again using the formula (IV.4) we obtain easily

$$H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}=H^{\overline{\boldsymbol{x}}}_{t_1}\,H^{\overline{\boldsymbol{x}}\cup t_1}_{\{t_2,\dots,t_n\}}=H^{\overline{\boldsymbol{x}}}_{t_1}\,H^{\overline{\boldsymbol{x}}\cup t_1}_{t_2}\,H^{\overline{\boldsymbol{x}}\cup t_1\cup t_2}_{\{t_3,\dots,t_n\}}=\cdots=h^{\overline{\boldsymbol{x}}}_{t_1}\,h^{\overline{\boldsymbol{x}}\cup t_1}_{t_2}\,\cdots\,h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_{n-1}}_{t_n}$$

which concludes the proof of the necessity.

2) Sufficiency. Let $\boldsymbol{h}=\bigl\{h^{\overline{\boldsymbol{x}}}_{t},\quad t\in\mathbb{Z}^{\nu}$ and $\overline{\boldsymbol{x}}\subset\mathbb{Z}^{\nu}\setminus t\bigr\}$ be a one-point system and for all $\boldsymbol{x}\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}$ put

$$H^{\overline{\boldsymbol{x}}}_{\boldsymbol{x}}=h^{\overline{\boldsymbol{x}}}_{t_1}\,h^{\overline{\boldsymbol{x}}\cup t_1}_{t_2}\,\cdots\,h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_{n-1}}_{t_n}\geqslant 0. \tag{V.3}$$

First of all let us verify that this definition is correct, *i.e.*, that it does not depend on the enumeration of the set $\boldsymbol{x}$. For this let us fix some enumeration $t_1,\dots,t_n$ and let $\boldsymbol{\varphi}=\bigl\{\varphi(1),\dots,\varphi(n)\bigr\}$ and $\boldsymbol{\psi}=\bigl\{\psi(1),\dots,\psi(n)\bigr\}$ be two permutations of the set $\{1,\dots,n\}$. We need to show that

$$h^{\overline{\boldsymbol{x}}}_{t_{\varphi(1)}}h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}}_{t_{\varphi(2)}}\cdots h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}\cup\cdots\cup t_{\varphi(n-1)}}_{t_{\varphi(n)}}=h^{\overline{\boldsymbol{x}}}_{t_{\psi(1)}}h^{\overline{\boldsymbol{x}}\cup t_{\psi(1)}}_{t_{\psi(2)}}\cdots h^{\overline{\boldsymbol{x}}\cup t_{\psi(1)}\cup\cdots\cup t_{\psi(n-1)}}_{t_{\psi(n)}}. \tag{V.4}$$

It is well known that any permutation of the set $\{1,\dots,n\}$ can be decomposed in a product of transpositions of nearest neighbours, and hence it suffice to consider only the case where $\boldsymbol{\psi}=\boldsymbol{\varphi}\circ(k,k+1)$ with some $k\in\{1,\dots,n-1\}$, *i.e.*, $\boldsymbol{\psi}=\bigl\{\varphi(1),\dots,\varphi(k-1),\varphi(k+1),\varphi(k),\varphi(k+2),\dots,\varphi(n)\bigr\}$. But in this case the relation (V.4) is reduced to

$$\begin{aligned}h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}\cup\cdots\cup t_{\varphi(k-1)}}_{t_{\varphi(k)}}\,h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}\cup\cdots\cup t_{\varphi(k-1)}\cup t_{\varphi(k)}}_{t_{\varphi(k+1)}}&=\\ =h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}\cup\cdots\cup t_{\varphi(k-1)}}_{t_{\varphi(k+1)}}\,h^{\overline{\boldsymbol{x}}\cup t_{\varphi(1)}\cup\cdots\cup t_{\varphi(k-1)}\cup t_{\varphi(k+1)}}_{t_{\varphi(k)}}&.\end{aligned}$$

which is an evident consequence of (V.1). Now we can finally check the consistency of the H-system $\boldsymbol{H}$. For this let us take some $S_1 = \{t_1, \dots, t_n\} \in \mathscr{E}$ and $S_2 = \{s_1, \dots, s_m\} \in \mathscr{E}$ such that $S_1 \cap S_2 = \emptyset$ and some $\overline{\boldsymbol{x}} \subset (S_1 \cup S_2)^{\mathrm{c}}$. We have $S_1 \cup S_2 = \{t_1, \dots, t_n, s_1, \dots, s_m\}$ and hence using the definition (V.3) of the H-system $\boldsymbol{H}$ we get

$$
\begin{aligned}
H^{\overline{\boldsymbol{x}}}_{\boldsymbol{S}_1} &= h^{\overline{\boldsymbol{x}}}_{t_1}\, h^{\overline{\boldsymbol{x}}\cup t_1}_{t_2}\, \cdots\, h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_{n-1}}_{t_n},\\
H^{\overline{\boldsymbol{x}}\cup S_1}_{\boldsymbol{S}_2} &= h^{\overline{\boldsymbol{x}}\cup S_1}_{s_1}\, h^{\overline{\boldsymbol{x}}\cup S_1\cup s_1}_{s_2}\, \cdots\, h^{\overline{\boldsymbol{x}}\cup S_1\cup s_1\cup\cdots\cup s_{m-1}}_{s_m},\\
H^{\overline{\boldsymbol{x}}}_{\boldsymbol{S}_1\cup S_2} &= h^{\overline{\boldsymbol{x}}}_{t_1}\, \cdots\, h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_{n-1}}_{t_n}\, h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_n}_{s_1}\, \cdots\, h^{\overline{\boldsymbol{x}}\cup t_1\cup\cdots\cup t_n\cup s_1\cup\cdots\cup s_{m-1}}_{s_m},
\end{aligned}
$$

and hence the relation (IV.4) holds. The theorem is proved. □

Note that since $h^{\overline{\boldsymbol{x}}}_t \geqslant 0$ for all $t \in \mathbb{Z}^\nu$ and $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ we can denote $u^{\overline{\boldsymbol{x}}}(t) = -\ln h^{\overline{\boldsymbol{x}}}_t$ permitting the system $\boldsymbol{u} = \left\{u^{\overline{\boldsymbol{x}}}(t), \quad t \in \mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t\right\}$ to take the value $+\infty$. This system is clearly nothing but one-point subsystem of some general Hamiltonian $\boldsymbol{\mathcal{U}}$ including also Gibbsian case.

Let us also note that by properties of one-point systems and H-systems we have

$$
\mathbf{Q}^{\overline{\boldsymbol{x}}}_t(t) = \frac{H^{\overline{\boldsymbol{x}}}_t}{\sum\limits_{\boldsymbol{y}\subset t} H^{\overline{\boldsymbol{x}}}_{\boldsymbol{y}}} = \frac{H^{\overline{\boldsymbol{x}}}_t}{H^{\overline{\boldsymbol{x}}}_\emptyset + H^{\overline{\boldsymbol{x}}}_t} = \frac{h^{\overline{\boldsymbol{x}}}_t}{1 + h^{\overline{\boldsymbol{x}}}_t}\ ,
$$

$$
\mathbf{Q}^{\overline{\boldsymbol{x}}}_t(\emptyset) = \frac{H^{\boldsymbol{x}}_\emptyset}{\sum\limits_{\boldsymbol{y}\subset t} H^{\overline{\boldsymbol{x}}}_{\boldsymbol{y}}} = \frac{H^{\boldsymbol{x}}_\emptyset}{H^{\overline{\boldsymbol{x}}}_\emptyset + H^{\overline{\boldsymbol{x}}}_t} = \frac{1}{1 + h^{\overline{\boldsymbol{x}}}_t}\ ,
$$

and hence

$$
h^{\overline{\boldsymbol{x}}}_t = \frac{\mathbf{Q}^{\overline{\boldsymbol{x}}}_t(t)}{\mathbf{Q}^{\overline{\boldsymbol{x}}}_t(\emptyset)}\ . \tag{V.5}
$$

Using the last formula we see that in fact the Theorem V.2 shows when a system of one-point distributions with boundary conditions is a subsystem consisting of one-point distributions of some specification. This question is an old open problem posed by Dobrushin who, in his paper [8], shows that under some positivity condition (clearly satisfied for the vacuum case) the whole specification can be determined by its subsystem consisting only of one-point distributions, but does not answer the question: "when a given system of one-point distributions with boundary conditions is a subsystem consisting of one-point distributions of some specification". In fact the Theorem V.2 shows that a necessary and

sufficient condition for that is the condition (V.1) or, rewritten in the term of the specification using the formula (V.5), the condition

$$\mathbf{Q}_s^{\overline{\boldsymbol{x}}}(s)\ \mathbf{Q}_t^{\overline{\boldsymbol{x}}\cup s}(t)\ \mathbf{Q}_t^{\overline{\boldsymbol{x}}}(\emptyset)\ \mathbf{Q}_s^{\overline{\boldsymbol{x}}\cup t}(\emptyset) = \mathbf{Q}_t^{\overline{\boldsymbol{x}}}(t)\ \mathbf{Q}_s^{\overline{\boldsymbol{x}}\cup t}(s)\ \mathbf{Q}_s^{\overline{\boldsymbol{x}}}(\emptyset)\ \mathbf{Q}_t^{\overline{\boldsymbol{x}}\cup s}(\emptyset).$$

The problem of description of a specification by its subsystem consisting of one-point distributions is very important because Dobrushin's uniqueness condition is taking into account only one-point distributions. For the same reason, the description of vacuum specifications by one-point systems that we have proposed in this section is very important and interesting. Clearly we can rewrite Dobrushin's uniqueness condition in terms of one-point systems, by substituting $\boldsymbol{\mathcal{Q}}$ by its values expressed by $\boldsymbol{h}$ in the formula (I.9). In the {0,1} case, after obvious simplification this formula rewrites in the following form:

$$\sup_{t\in\mathbb{Z}^\nu} \sum_{s\in\mathbb{Z}^\nu\setminus t} \sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\{s,t\}} \frac{\left|h_t^{\overline{\boldsymbol{x}}} - h_t^{\overline{\boldsymbol{x}}\cup s}\right|}{\left(1+h_t^{\overline{\boldsymbol{x}}}\right)\left(1+h_t^{\overline{\boldsymbol{x}}\cup s}\right)} < 1. \tag{V.6}$$

V.2. Gibbsian one-point systems

Let us at first give some examples of one-point systems.

Examples V.3. — **1)** Let $\boldsymbol{\Phi} = \left\{\Phi(J),\ J\in\mathscr{E}\setminus\{\emptyset\}\right\}$ be a convergent interaction potential. Then the system $\boldsymbol{h} = \left\{\exp\left(-u^{\overline{\boldsymbol{x}}}(t)\right),\quad t\in\mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus t\right\}$ defined by

$$u^{\overline{\boldsymbol{x}}}(t) \ = \sum_{S\in\mathscr{E}\,:\,S\subset\overline{\boldsymbol{x}}} \Phi\big(S\cup t\big).$$

is clearly a one-point system corresponding to Gibbsian specification with the interaction potential $\boldsymbol{\Phi}$. We call such one-point systems Gibbsian.

2) Let $\boldsymbol{h} = \left\{h_t^{\overline{\boldsymbol{x}}},\quad t\in\mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus t\right\}$ be a non-negative system such that $h_t^{\overline{\boldsymbol{x}}_1} = h_t^{\overline{\boldsymbol{x}}_2}$ if $\overline{\boldsymbol{x}}_1 = \overline{\boldsymbol{x}}_2$ up to a finite number of lattice points. Then $\boldsymbol{h}$ is clearly a one-point system.

3) Let $\boldsymbol{h} = \left\{h_t^{\overline{\boldsymbol{x}}},\quad t\in\mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus t\right\}$ be a one-point system and $\boldsymbol{R} = \{R(\overline{\boldsymbol{x}}),\ \overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\}$ be a real-valued strictly positive function such that we have $R(\overline{\boldsymbol{x}}_1) = R(\overline{\boldsymbol{x}}_2)$ if $\overline{\boldsymbol{x}}_1 = \overline{\boldsymbol{x}}_2$ up to a finite number of lattice points. Let us consider the system

$$\boldsymbol{h_R} = \left\{\left(h_t^{\overline{\boldsymbol{x}}}\right)^{R(\overline{\boldsymbol{x}})},\quad t\in\mathbb{Z}^\nu \text{ and } x\subset\mathbb{Z}^\nu\setminus t\right\}.$$

For this system the condition (V.1) is clearly satisfied, and hence this is a one-point system. This system corresponds to H-system considered in the Lemma IV.6. The considerations of the Remark IV.7 hold.

The last example gives us a way to construct from, for example, Gibbsian one-point systems some new one-point systems, and the latter ones can be clearly shown to be non-Gibbsian under some condition analogous to that of the Proposition IV.8 and Remark IV.9. Now let us do better and give a general necessary and sufficient condition for a one-point system to be Gibbsian.

Theorem V.4. — *A one-point system* $\boldsymbol{h} = \big\{h_t^{\overline{\boldsymbol{x}}}, \quad t \in \mathbb{Z}^\nu \text{ and } \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t\big\}$ *is Gibbsian if and only if the following two conditions are satisfied:*

(***h*1**) *for all* $t \in \mathbb{Z}^\nu$ *and* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ *we have* $\lim\limits_{I \uparrow \mathbb{Z}^\nu} h_t^{\overline{\boldsymbol{x}}_I} = h_t^{\overline{\boldsymbol{x}}}$,

(***h*2**) *for all* $t \in \mathbb{Z}^\nu$ *and* $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ *we have* $h_t^{\overline{\boldsymbol{x}}} = 0$ *if there exist some* $T \in \mathscr{E}$ *such that* $h_t^{\overline{\boldsymbol{x}}_T} = 0$.

Proof: **1)** Necessity. We suppose that the one-point system $\boldsymbol{h}$ is Gibbsian, *i.e.*, that for all $t \in \mathbb{Z}^\nu$ and $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ we have $h_t^{\overline{\boldsymbol{x}}} = \exp\big(-u^{\overline{\boldsymbol{x}}}(t)\big)$ with

$$u^{\overline{\boldsymbol{x}}}(t) \;=\; \sum_{S \in \mathscr{E} \,:\, S \subset \overline{\boldsymbol{x}}} \Phi\big(S \cup t\big)$$

where $\boldsymbol{\Phi}$ is some convergent interaction potential. We need to check the conditions (***h*1**) and (***h*2**). The first condition follows obviously from the fact that interaction potential $\boldsymbol{\Phi}$ is convergent. To check the second one let us take some $t \in \mathbb{Z}^\nu$ and $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$ and suppose that there exists some $T \in \mathscr{E}$ such that $h_t^{\overline{\boldsymbol{x}}_T} = 0$. We need to show that $h_t^{\overline{\boldsymbol{x}}} = 0$. We have

$$u^{\overline{\boldsymbol{x}}_T}(t) = -\ln\big(h_t^{\overline{\boldsymbol{x}}_T}\big) = +\infty \;=\; \sum_{S \in \mathscr{E} \,:\, S \subset \overline{\boldsymbol{x}}_T} \Phi\big(S \cup t\big) \;=\; \sum_{S \subset \overline{\boldsymbol{x}}_T} \Phi\big(S \cup t\big).$$

But the last sum contains finite number of summands and hence at least one of them is equal to $+\infty$. This implies that for any $I \in \mathscr{E}$ such that $I \supset T$ we have $u^{\overline{\boldsymbol{x}}_I}(t) = +\infty$, and since $\boldsymbol{\Phi}$ is convergent we have also $u^{\overline{\boldsymbol{x}}}(t) = +\infty$, and hence $h_t^{\overline{\boldsymbol{x}}} = \exp\big(-u^{\overline{\boldsymbol{x}}}(t)\big) = 0$ which concludes the proof of the necessity.

2) Sufficiency. We suppose that the one-point system $\boldsymbol{h}$ satisfies the conditions

(***h*1**) and (***h*2**) and that $\boldsymbol{u}$ is one-point subsystem of the corresponding Hamiltonian. Let us consider the interaction potential $\mathbf{\Phi}$ defined by

$$\Phi(J) = \begin{cases} +\infty & \text{if } \forall\ \xi \in J \text{ we have } u^{J\setminus\xi}(\xi) = +\infty, \\ \displaystyle\sum_{R\subset J\setminus\xi} (-1)^{\left|(J\setminus\xi)\setminus R\right|} u^R(\xi) & \text{if } \exists\ \xi \in J \text{ such that } u^{J\setminus\xi}(\xi) \in \mathbb{R}. \end{cases}$$

Note that the last sum is well defined since the number of summands is finite and by the condition (***h*2**) all the summands are finite. We can also show that this definition is correct, *i.e.*, that if $u^{J\setminus\xi}(\xi), u^{J\setminus\zeta}(\zeta) \in \mathbb{R}$ then

$$\sum_{R\subset J\setminus\xi} (-1)^{\left|(J\setminus\xi)\setminus R\right|} u^R(\xi) = \sum_{R\subset J\setminus\zeta} (-1)^{\left|(J\setminus\zeta)\setminus R\right|} u^R(\zeta).$$

Indeed, we have

$$\begin{aligned} \sum_{R\subset J\setminus\xi} (-1)^{\left|(J\setminus\xi)\setminus R\right|} u^R(\xi) &= \sum_{R\subset J\setminus\{\xi,\zeta\}} (-1)^{\left|(J\setminus\xi)\setminus R\right|} u^R(\xi) + \\ &\qquad + \sum_{R\subset J\setminus\{\xi,\zeta\}} (-1)^{\left|(J\setminus\xi)\setminus(R\cup\zeta)\right|} u^{R\cup\zeta}(\xi) = \\ &= \sum_{R\subset J\setminus\{\xi,\zeta\}} (-1)^{\left|(J\setminus\xi)\setminus R\right|} \big(u^R(\xi) - u^{R\cup\zeta}(\xi)\big), \end{aligned}$$

and in the same manner

$$\sum_{R\subset J\setminus\zeta} (-1)^{\left|(J\setminus\zeta)\setminus R\right|} u^R(\zeta) = \sum_{R\subset J\setminus\{\xi,\zeta\}} (-1)^{\left|(J\setminus\zeta)\setminus R\right|} \big(u^R(\zeta) - u^{R\cup\xi}(\zeta)\big).$$

Since all the terms in these sums are finite and using the condition (V.1) we see that the sums are term by term equal.

It remains to check that the potential $\mathbf{\Phi}$ indeed corresponds to our one-point system $\boldsymbol{h}$, *i.e.*, that

$$u^{\overline{\boldsymbol{x}}}(t) = \sum_{S\in\mathscr{E}\,:\,S\subset\overline{\boldsymbol{x}}} \Phi(S\cup t) \tag{V.7}$$

for all $t \in \mathbb{Z}^\nu$ and $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus t$. Since the condition (***h*1**) holds it is sufficient to verify this relation only in the case when $\overline{\boldsymbol{x}} \in \mathscr{E}$. Let us at first suppose that the l.h.s. of (V.7) is finite. In this case by (***h*1**) we have $u^S(t) < +\infty$ for all $S \subset \overline{\boldsymbol{x}}$. Then by definition of $\mathbf{\Phi}$ we have

$$\Phi(S\cup t) = \sum_{R\subset S} (-1)^{|S\setminus R|} u^R(t),$$

and hence

$$\text{r.h.s. of (V.7)} \;=\sum_{S\subset\overline{\boldsymbol{x}}}\;\sum_{R\subset S}(-1)^{|S\setminus R|}u^{R}(t)=u^{\overline{\boldsymbol{x}}}(t).$$

Now let us consider the case when the l.h.s. of (V.7) is infinite, *i.e.*, when $u^{\overline{\boldsymbol{x}}}(t)=+\infty$. We need to show that the r.h.s. of (V.7) is also infinite. Two cases are possible:

• We have $u^{\emptyset}(t)=+\infty$. In this case by the definition of $\mathbf{\Phi}$ we obtain $\Phi(t)=+\infty$ and since $\Phi(t)$ is one of the summands in the r.h.s. of (V.7) the latter is infinite.

• We have $u^{\emptyset}(t)\in\mathbb{R}$. In this case clearly there exists some $S\subset\overline{\boldsymbol{x}}$ such that $S\neq\emptyset$, $u^{S}(t)=+\infty$, and for all $\xi\in S$ we have $u^{S\setminus\xi}(t)\in\mathbb{R}$. Hence, for all $\xi\in S$ we can write

$$u^{S\setminus\xi}(t)+u^{(S\setminus\xi)\cup t}(\xi)=u^{S\setminus\xi}(\xi)+u^{S}(t)=u^{S\setminus\xi}(\xi)+(+\infty)=+\infty.$$

But $u^{S\setminus\xi}(t)\in\mathbb{R}$, and hence we have $u^{(S\setminus\xi)\cup t}(\xi)=u^{(S\cup t)\setminus\xi}(\xi)=+\infty$ for all $\xi\in S$. Clearly we have also $u^{(S\cup t)\setminus t}(t)=u^{S}(t)=+\infty$. Thus, by definition of $\mathbf{\Phi}$ we have $\Phi(S\cup t)=+\infty$ and hence the r.h.s. of (V.7) is infinite. □

Note that this theorem can obviously be reformulated in terms of H-systems, *i.e.*, a H-system is Gibbsian if and only if the conditions (***h*1**) and (***h*2**) hold. Clearly in this case the conditions (***h*1**) and (***h*2**) can be replaced by equivalent conditions:

(***H*1**) for all $\boldsymbol{x}\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}$ we have $\lim\limits_{I\uparrow\mathbb{Z}^{\nu}}H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}_I}=H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}$,

(***H*2**) for all $\boldsymbol{x}\in\mathscr{E}$ and $\overline{\boldsymbol{x}}\subset\boldsymbol{x}^{\mathrm{c}}$ we have $H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}=0$ if there exist some $T\in\mathscr{E}$ such that $H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}_T}=0$.

Let us finally note here that the Theorem V.4 shows when a vacuum specification has a Gibbs representation. A similar problems were considered in [2], [17], [24] and [12] in less general setup, *e.g.*, for local, quasilocal and/or strictly positive specifications.

V.3. Generalizations to the case of arbitrary finite state space

As shows the preceding sections, in the {0,1}consistent H-systems (and hence vacuum specifications) are completely determined by their one-point subsystems (one-point systems). This assertion generalizes straightforwardly to the case of arbitrary finite state space $\mathscr{X}$. That is, in this case one can still determine completely a H-system by its one-point subsystem.

Let us consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$.

Definition V.5. — *A system*

$$\boldsymbol{h} = \left\{ h_t^{\overline{\boldsymbol{x}}}(x), \quad t \in \mathbb{Z}^\nu,\ x \in \mathscr{X}^*,\ \overline{\boldsymbol{x}} \in \mathscr{X}^{*K},\ K \subset \mathbb{Z}^\nu \setminus t \right\}$$

is called one-point system if for all $t \in \mathbb{Z}^\nu$, $x \in \mathscr{X}^*$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset \mathbb{Z}^\nu \setminus t$ we have $h_t^{\overline{\boldsymbol{x}}}(x) \geqslant 0$ and for all $s,t \in \mathbb{Z}^\nu$, $x,y \in \mathscr{X}^*$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset \mathbb{Z}^\nu \setminus \{s,t\}$ we have

$$h_s^{\overline{\boldsymbol{x}}}(y)\, h_t^{\overline{\boldsymbol{x}} \oplus y_s}(x) = h_t^{\overline{\boldsymbol{x}}}(x)\, h_s^{\overline{\boldsymbol{x}} \oplus x_t}(y).$$

Here and in the sequel x_t denotes a configuration on the set t taking value x in the point t.

Theorem V.6. — *A system $\boldsymbol{\mathcal{H}}$ is a consistent H-system if and only if there exists a one-point system $\boldsymbol{h}$ such that for all $\boldsymbol{x} \in \mathscr{X}^{*I}$, $I \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset I^{\mathrm{c}}$ we have*

$$H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}} = h_{t_1}^{\overline{\boldsymbol{x}}}(x_{t_1})\, h_{t_2}^{\overline{\boldsymbol{x}} \oplus x_{t_1}}(x_{t_2}) \cdots h_{t_n}^{\overline{\boldsymbol{x}} \oplus x_{t_1} \oplus \cdots \oplus x_{t_{n-1}}}(x_{t_n})$$

where $n = |I|$ and $t_1, \ldots, t_n$ is some arbitrary enumeration of elements of the set I. Particularly, for all $t \in \mathbb{Z}^\nu$, $x \in \mathscr{X}^*$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{*K}$, $K \subset \mathbb{Z}^\nu \setminus t$ we have $H_{x_t}^{\overline{\boldsymbol{x}}} = h_t^{\overline{\boldsymbol{x}}}(x)$.

The proof for this general case is just the repetition of the proof corresponding to the {0,1} case. All the other results concerning one-point systems (except the simplified form (V.6) of Dobrushin's uniqueness condition) are easily generalized for this general case.

VI. Description of quasilocal specifications

In this chapter we concentrate on the description of quasilocal specifications since, as we have already seen in the first chapter, they are very important in the theory of random fields. In the first section we will consider the case of vacuum specifications and we will apply the results of the two previous chapters. In the second section we will replace the condition of vacuumness by some slightly different condition and we will show that in this case one can describe specifications by H-functions or Q-functions satisfying some additional conditions.

VI.1. Case of vacuum specifications

Let us at first consider the $\{0,1\}$ case and study how quasilocal vacuum specifications can be described in this case.

Clearly, as before, one can describe them by consistent Q-systems, consistent H-systems and/or one-point systems. Note that it is evident that in this case the specification will be local if and only if corresponding Q-system (H-system, one-point system) is local. Analogously the specification will be quasilocal if and only if corresponding Q-system (H-system, one-point system) is quasilocal with respect to the variable $\overline{\boldsymbol{x}}$, *i.e.*, satisfies corresponding quasilocality condition

$$\alpha_J(I) = \sup_{\overline{\boldsymbol{x}}\subset J^{\mathrm{c}}} \left|\theta_J^{\overline{\boldsymbol{x}}_I} - \theta_J^{\overline{\boldsymbol{x}}}\right| \underset{I\uparrow\mathbb{Z}^\nu}{\longrightarrow} 0, \quad J \in \mathscr{E},$$

$$\beta_{\boldsymbol{x}}(I) = \sup_{\overline{\boldsymbol{x}}\subset \boldsymbol{x}^{\mathrm{c}}} \left|H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}_I} - H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}}\right| \underset{I\uparrow\mathbb{Z}^\nu}{\longrightarrow} 0, \quad \boldsymbol{x} \in \mathscr{E},$$

$$\gamma_t(I) = \sup_{\overline{\boldsymbol{x}}\subset \mathbb{Z}^\nu\setminus t} \left|h_t^{\overline{\boldsymbol{x}}_I} - h_t^{\overline{\boldsymbol{x}}}\right| \underset{I\uparrow\mathbb{Z}^\nu}{\longrightarrow} 0, \quad t \in \mathbb{Z}^\nu.$$

This can be easily proved using the following obvious observation. Since the space $(\Omega, \mathscr{T})$ is compact then any quasilocal function on it is bounded, and if it is strictly positive then it is uniformly strictly positive, *i.e.*, it is greater than some $c > 0$.

Let us mention here that a specification $\boldsymbol{\mathcal{Q}}$ corresponding to some H-system (one-point system) is Gibbsian with uniformly convergent interaction potential if and only if this H-system (one-point system) is quasilocal and strictly positive. Note also, that under the condition of strict positivity of a H-system (one-point system), its quasilocality is clearly equivalent to the quasilocality of its logarithm, *i.e.*, to the quasilocality of Hamiltonian (one-point Hamiltonian).

Note that everything exposed in this section (except Q-systems) generalizes straightforwardly to the case of vacuum specifications with arbitrary finite state space $\mathscr{X}$. Clearly, in this case the quasilocality condition for H-system (one-point system) looks like

$$\beta_{\boldsymbol{x}}(I) = \sup_{\overline{\boldsymbol{x}} \in \mathscr{X}^{I^{\mathrm{c}}}} \left| H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}_I} - H_{\boldsymbol{x}}^{\overline{\boldsymbol{x}}} \right| \underset{I \uparrow \mathbb{Z}^\nu}{\longrightarrow} 0, \quad \boldsymbol{x} \in \mathscr{X}^{*I},\ I \in \mathscr{E},$$

$$\gamma_{t,x}(I) = \sup_{\overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu \setminus t}} \left| h_t^{\overline{\boldsymbol{x}}_I}(x) - h_t^{\overline{\boldsymbol{x}}}(x) \right| \underset{I \uparrow \mathbb{Z}^\nu}{\longrightarrow} 0, \quad t \in \mathbb{Z}^\nu,\ x \in \mathscr{X}^*.$$

VI.2. Quasilocal specifications, Q-functions and H-functions

Now let us propose an alternative approach towards description of quasilocal specifications based not on Q-systems, H-systems and/or one-point systems, but on Q-functions and H-functions. For instance we consider the $\{0,1\}$ case.

Theorem VI.1. — *Let* $\boldsymbol{\mathcal{Q}} = \left\{ \mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}} \right\}$ *be a quasilocal specification satisfying*

(**Q1**) $\mathbf{Q}_\Lambda^{\emptyset}(\emptyset) > 0$ *for all* $\Lambda \in \mathscr{E}$,

(**Q2**) $\mathbf{Q}_\Lambda^{\emptyset}(\boldsymbol{x}) + \mathbf{Q}_\Lambda^{\emptyset}(\boldsymbol{x} \cup t) > 0$ *for all* $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$, $t \in \Lambda$ *and* $\boldsymbol{x} \subset \Lambda \setminus t$.

Then there exists a H*-function* $\boldsymbol{H} = \{H_{\boldsymbol{x}},\ \boldsymbol{x} \in \mathscr{E}\}$ *satisfying*

(***H*1**) $H_{\boldsymbol{x}} + H_{\boldsymbol{x} \cup t} > 0$ *for all* $\boldsymbol{x} \in \mathscr{E}$ *and* $t \notin \boldsymbol{x}$,

(***H*2**) *for all* $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ *and* $\boldsymbol{x} \subset \Lambda$ *there exists uniformly on* $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ *the limit*

$$\lim_{I \uparrow \mathbb{Z}^\nu} \frac{H_{\boldsymbol{x} \cup \overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z} \subset \Lambda} H_{\boldsymbol{z} \cup \overline{\boldsymbol{x}}_I}},$$

and such that for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ *and all* $\overline{\boldsymbol{y}} \in \mathscr{E}$ *such that* $\overline{\boldsymbol{y}} \subset \Lambda^{\mathrm{c}}$ *we have*

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{y}}}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x} \cup \overline{\boldsymbol{y}}}}{\sum\limits_{\boldsymbol{z} \subset \Lambda} H_{\boldsymbol{z} \cup \overline{\boldsymbol{y}}}}, \quad \boldsymbol{x} \subset \Lambda. \tag{VI.1}$$

Conversely, if $\boldsymbol{H}$ *is a* H*-function satisfying* ($\boldsymbol{H1}$) *and* ($\boldsymbol{H2}$)*, then one can find a quasilocal specification* $\boldsymbol{\mathcal{Q}}$ *satisfying* ($\mathbf{Q1}$), ($\mathbf{Q2}$) *and (VI.1).*

Proof: **1)** Necessity. Let $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\right\}$ be a quasilocal specification satisfying ($\mathbf{Q1}$) and ($\mathbf{Q2}$). The system $\mathbf{Q} = \left\{\mathbf{Q}_\Lambda^{\emptyset}, \quad \Lambda \in \mathscr{E}\right\}$ is clearly a system of probability distributions consistent in Dobrushin's sense, and hence by the Theorems III.5 and III.2 there exists some H-function $\boldsymbol{H}$ corresponding to some Q-function $\boldsymbol{\theta}$ such that for all $\Lambda \in \mathscr{E}$ and $\boldsymbol{x} \subset \Lambda$ we have

$$\mathbf{Q}_\Lambda^{\emptyset}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x}}}{\theta_\Lambda}.$$

Further, for all $\boldsymbol{x} \in \mathscr{E}$ and $t \notin \boldsymbol{x}$ we can write

$$H_{\boldsymbol{x}} + H_{\boldsymbol{x}\cup t} = \theta_{\boldsymbol{x}\cup t}\ \mathbf{Q}_{\boldsymbol{x}\cup t}^{\emptyset}(\boldsymbol{x}) + \theta_{\boldsymbol{x}\cup t}\ \mathbf{Q}_{\boldsymbol{x}\cup t}^{\emptyset}(\boldsymbol{x}\cup t) =$$

$$= \theta_{\boldsymbol{x}\cup t}\left(\mathbf{Q}_{\boldsymbol{x}\cup t}^{\emptyset}(\boldsymbol{x}) + \mathbf{Q}_{\boldsymbol{x}\cup t}^{\emptyset}(\boldsymbol{x}\cup t)\right) > 0,$$

and hence the condition ($\boldsymbol{H1}$) holds. In order to verify (VI.1) and the condition ($\boldsymbol{H2}$) let us at first note that by ($\boldsymbol{H1}$) for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ and all $\overline{\boldsymbol{y}} \in \mathscr{E}$ such that $\overline{\boldsymbol{y}} \subset \Lambda^{\mathrm{c}}$ we have

$$\sum_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}} \geqslant H_{\overline{\boldsymbol{y}}} + H_{\overline{\boldsymbol{y}}\cup t} > 0$$

where we have chosen some $t \in \Lambda$, and hence

$$\left(\mathbf{Q}_{\Lambda\cup\overline{\boldsymbol{y}}}^{\emptyset}\right)_{\overline{\boldsymbol{y}}}(\overline{\boldsymbol{y}}) = \sum_{\boldsymbol{z}\subset\Lambda} \mathbf{Q}_{\Lambda\cup\overline{\boldsymbol{y}}}^{\emptyset}(\boldsymbol{z}\cup\overline{\boldsymbol{y}}) = \sum_{\boldsymbol{z}\subset\Lambda} \frac{H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}}}{\theta_{\Lambda\cup\overline{\boldsymbol{y}}}} = \frac{1}{\theta_{\Lambda\cup\overline{\boldsymbol{y}}}} \sum_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}} > 0.$$

Now, since $\boldsymbol{\mathcal{Q}}$ is a specification we can write

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{y}}}(\boldsymbol{x}) = \frac{\mathbf{Q}_{\Lambda\cup\overline{\boldsymbol{y}}}^{\emptyset}(\boldsymbol{x}\cup\overline{\boldsymbol{y}})}{\left(\mathbf{Q}_{\Lambda\cup\overline{\boldsymbol{y}}}^{\emptyset}\right)_{\overline{\boldsymbol{y}}}(\overline{\boldsymbol{y}})} = \frac{\dfrac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{y}}}}{\theta_{\Lambda\cup\overline{\boldsymbol{y}}}}}{\dfrac{1}{\theta_{\Lambda\cup\overline{\boldsymbol{y}}}} \displaystyle\sum_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}}} = \frac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{y}}}}{\displaystyle\sum_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}}},$$

and hence (VI.1) holds. Condition ($\boldsymbol{H2}$) holds obviously since the specification $\boldsymbol{\mathcal{Q}}$ is quasilocal. The necessity is proved.

2) Sufficiency. Let $\boldsymbol{H} = \{H_{\boldsymbol{x}},\ \boldsymbol{x} \in \mathscr{E}\}$ be a H-function satisfying the conditions ($\boldsymbol{H1}$) and ($\boldsymbol{H2}$) and $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ be the corresponding Q-function. First of all, let us note that by ($\boldsymbol{H1}$) the denominators in (VI.1) and ($\boldsymbol{H2}$) are strictly positive. Now, for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ and all $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ we can put

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) = \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\displaystyle\sum_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}} \geqslant 0,$$

and for $\Lambda = \emptyset$ as always we consider $\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\emptyset}(\emptyset) = 1$ for all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu$. Clearly (VI.1) is satisfied. Further, for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ and all $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$ we have

$$\sum_{\boldsymbol{x}\subset\Lambda} \mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda}(\boldsymbol{x}) = \sum_{\boldsymbol{x}\subset\Lambda} \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}} = \lim_{I\uparrow\mathbb{Z}^\nu} \frac{\sum\limits_{\boldsymbol{x}\subset\Lambda} H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}} = 1,$$

and for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$, $\widetilde{\Lambda} \in \mathscr{E}$ such that $\Lambda \cap \widetilde{\Lambda} = \emptyset$ and all $\boldsymbol{x} \subset \Lambda$, $\boldsymbol{y} \subset \widetilde{\Lambda}$ and $\overline{\boldsymbol{x}} \subset \big(\Lambda \cup \widetilde{\Lambda}\big)^{\mathrm{c}}$ we can write

$$\begin{aligned}
\mathbf{Q}^{\overline{\boldsymbol{x}}\cup\boldsymbol{y}}_{\Lambda}(\boldsymbol{x}) \big(\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}\big)_{\widetilde{\Lambda}}(\boldsymbol{y}) &= \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{\boldsymbol{x}\cup(\overline{\boldsymbol{x}}\cup\boldsymbol{y})_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup(\overline{\boldsymbol{x}}\cup\boldsymbol{y})_I}} \times \sum_{\boldsymbol{z}\subset\Lambda} \mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}(\boldsymbol{z}\cup\boldsymbol{y}) \\
&= \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{(\boldsymbol{x}\cup\boldsymbol{y})\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{(\boldsymbol{z}\cup\boldsymbol{y})\cup\overline{\boldsymbol{x}}_I}} \times \sum_{\boldsymbol{z}\subset\Lambda} \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{(\boldsymbol{z}\cup\boldsymbol{y})\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{R\subset\Lambda\cup\widetilde{\Lambda}} H_{R\cup\overline{\boldsymbol{x}}_I}} = \\
&= \lim_{I\uparrow\mathbb{Z}^\nu} \frac{H_{(\boldsymbol{x}\cup\boldsymbol{y})\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{R\subset\Lambda\cup\widetilde{\Lambda}} H_{R\cup\overline{\boldsymbol{x}}_I}} = \mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda\cup\widetilde{\Lambda}}(\boldsymbol{x}\cup\boldsymbol{y})
\end{aligned}$$

where we suppose I to be sufficiently grand, so that $I \supset \boldsymbol{y}$. Thus, the system $\boldsymbol{\mathcal{Q}} = \big\{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda}, \quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\big\}$ is a specification. Its quasilocality follows obviously from its definition and from the condition (***H*2**). It remains to verify the conditions (**Q1**) and (**Q2**). For all $\Lambda \in \mathscr{E}$ we have

$$\mathbf{Q}^{\emptyset}_{\Lambda}(\emptyset) = \frac{H_{\emptyset}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}}} = \frac{1}{\theta_\Lambda} > 0,$$

and for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$, $t \in \Lambda$ and $\boldsymbol{x} \subset \Lambda \setminus t$ we can write

$$\mathbf{Q}^{\emptyset}_{\Lambda}(\boldsymbol{x}) + \mathbf{Q}^{\emptyset}_{\Lambda}(\boldsymbol{x}\cup t) = \frac{H_{\boldsymbol{x}}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}}} + \frac{H_{\boldsymbol{x}\cup t}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}}} = \frac{1}{\theta_\Lambda}\left(H_{\boldsymbol{x}} + H_{\boldsymbol{x}\cup t}\right) > 0.$$

The theorem is proved. □

Note that this theorem can be reformulated in terms of Q-functions in the following way.

Corollary VI.2. — *Let $\boldsymbol{\mathcal{Q}} = \big\{\mathbf{Q}^{\overline{\boldsymbol{x}}}_{\Lambda}, \quad \Lambda \in \mathscr{E}$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}\big\}$ be a quasilocal specification satisfying the conditions* (**Q1**) *and* (**Q2**)*. Then there exists a Q-function $\boldsymbol{\theta} = \{\theta_J,\ J \in \mathscr{E}\}$ satisfying*

($\boldsymbol{\theta 1}$) $\sum_{S\subset \boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus S|}\theta_{S\cup t}>0$ *for all* $\boldsymbol{x}\in\mathscr{E}$ *and* $t\notin\boldsymbol{x}$,

($\boldsymbol{\theta 2}$) *for all* $\Lambda\in\mathscr{E}\setminus\{\emptyset\}$ *and* $J\subset\Lambda$ *there exists uniformly on* $\overline{\boldsymbol{x}}\subset\Lambda^{\mathrm{c}}$ *the limit*

$$\lim_{I\uparrow\mathbb{Z}^\nu}\frac{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I}(-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{J\cup S}}{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I}(-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{\Lambda\cup S}}\ ,$$

and such that for all $\Lambda\in\mathscr{E}\setminus\{\emptyset\}$ *and all* $\overline{\boldsymbol{y}}\in\mathscr{E}$ *such that* $\overline{\boldsymbol{y}}\subset\Lambda^{\mathrm{c}}$ *we have*

$$\mathbf{Q}_\Lambda^{\overline{\boldsymbol{y}}}(\boldsymbol{x})=\frac{\sum\limits_{R\subset\boldsymbol{x}\cup\overline{\boldsymbol{y}}}(-1)^{\left|(\boldsymbol{x}\cup\overline{\boldsymbol{y}})\setminus R\right|}\theta_R}{\sum\limits_{S\subset\overline{\boldsymbol{y}}}(-1)^{|\overline{\boldsymbol{y}}\setminus S|}\theta_{\Lambda\cup S}}\ ,\quad \boldsymbol{x}\subset\Lambda. \tag{VI.2}$$

Conversely, if $\boldsymbol{\theta}$ *is a* Q*-function satisfying* ($\boldsymbol{\theta 1}$) *and* ($\boldsymbol{\theta 2}$)*, then one can find a quasilocal specification* $\boldsymbol{\mathcal{Q}}$ *satisfying* (**Q1**), (**Q2**) *and (VI.2).*

Proof: First of all, let us note that if $\boldsymbol{\theta}$ is some Q-function and $\boldsymbol{H}$ is the corresponding H-function, then using (III.3) we get

$$\begin{aligned}\sum_{\boldsymbol{z}\subset\Lambda}H_{\boldsymbol{z}\cup\overline{\boldsymbol{y}}}&=\sum_{\boldsymbol{z}\subset\Lambda}\sum_{J\subset\boldsymbol{z}}(-1)^{|\boldsymbol{z}\setminus J|}\sum_{S\subset\overline{\boldsymbol{y}}}(-1)^{|\overline{\boldsymbol{y}}\setminus S|}\theta_{J\cup S}=\\&=\sum_{S\subset\overline{\boldsymbol{y}}}(-1)^{|\overline{\boldsymbol{y}}\setminus S|}\sum_{\boldsymbol{z}\subset\Lambda}\sum_{J\subset\boldsymbol{z}}(-1)^{|\boldsymbol{z}\setminus J|}\theta_{J\cup S}=\sum_{S\subset\overline{\boldsymbol{y}}}(-1)^{|\overline{\boldsymbol{y}}\setminus S|}\theta_{\Lambda\cup S}\end{aligned} \tag{VI.3}$$

for all $\Lambda\in\mathscr{E}$ and all $\overline{\boldsymbol{y}}\in\mathscr{E}$ such that $\overline{\boldsymbol{y}}\subset\Lambda^{\mathrm{c}}$.

1) Necessity. By the preceding theorem there exists a H-function $\boldsymbol{H}$ satisfying ($\boldsymbol{H1}$), ($\boldsymbol{H2}$) and (VI.1). Let $\boldsymbol{\theta}$ be the corresponding Q-function, and let us verify that it satisfies ($\boldsymbol{\theta 1}$), ($\boldsymbol{\theta 2}$) and (VI.2). For all $\boldsymbol{x}\in\mathscr{E}$ and $t\notin\boldsymbol{x}$ we have

$$\sum_{S\subset\boldsymbol{x}}(-1)^{|\boldsymbol{x}\setminus S|}\theta_{S\cup t}=H_{\boldsymbol{x}}+H_{\boldsymbol{x}\cup t}$$

where we have used the formula (VI.3) with $\Lambda=t$ and $\overline{\boldsymbol{y}}=\boldsymbol{x}$. Hence the condition ($\boldsymbol{\theta 1}$) is equivalent to the condition ($\boldsymbol{H1}$). Again by (VI.3) we get

$$\frac{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I}(-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{J\cup S}}{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I}(-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{\Lambda\cup S}}=\frac{\sum\limits_{\boldsymbol{x}\subset J}H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda}H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}}=\sum_{\boldsymbol{x}\subset J}\frac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda}H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}},$$

and hence the condition ($\boldsymbol{\theta 2}$) follows from the condition ($\boldsymbol{H2}$). The relation (VI.2) is clearly equivalent to (VI.1) using (III.3) and (VI.3).

2) SUFFICIENCY. Let $\boldsymbol{H} = \{H_{\boldsymbol{x}},\ \boldsymbol{x} \in \mathscr{E}\}$ be the H-system corresponding to the Q-system $\boldsymbol{\theta}$. Let us verify that $\boldsymbol{H}$ satisfies the conditions ($\boldsymbol{H1}$) and ($\boldsymbol{H2}$). For the first one see the proof of necessity. For the second one, using (III.3) and (VI.3) we can write

$$\frac{H_{\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}} = \frac{\sum\limits_{R\subset\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I} (-1)^{\left|(\boldsymbol{x}\cup\overline{\boldsymbol{x}}_I)\setminus R\right|}\theta_R}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}} = \frac{\sum\limits_{J\subset\boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|} \sum\limits_{S\subset\overline{\boldsymbol{x}}_I} (-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{J\cup S}}{\sum\limits_{\boldsymbol{z}\subset\Lambda} H_{\boldsymbol{z}\cup\overline{\boldsymbol{x}}_I}}$$

$$= \sum_{J\subset\boldsymbol{x}} (-1)^{|\boldsymbol{x}\setminus J|}\ \frac{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I} (-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{J\cup S}}{\sum\limits_{S\subset\overline{\boldsymbol{x}}_I} (-1)^{|\overline{\boldsymbol{x}}_I\setminus S|}\theta_{\Lambda\cup S}}$$

and hence the condition ($\boldsymbol{H2}$) follows from the condition ($\boldsymbol{\theta 2}$). Thus, by the preceding theorem there exists a quasilocal specification $\boldsymbol{\mathcal{Q}}$ satisfying ($\mathbf{Q1}$), ($\mathbf{Q2}$) and (VI.1). Hence it satisfies (VI.2) too, which concludes the proof. □

Let us note here that the class of specifications that we have considered in this section, *i.e.*, the class of all quasilocal specifications satisfying the conditions ($\mathbf{Q1}$) and ($\mathbf{Q2}$), includes the class of Gibbsian specifications with uniformly convergent interaction potentials as the particular case when we have $\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) > 0$ for all $\Lambda \in \mathscr{E}$, $\boldsymbol{x} \subset \Lambda$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$.

Finally let us turn to consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$. The generalization of the Theorem VI.1 to this case is quite straightforward.

THEOREM VI.3. — *Let* $\boldsymbol{\mathcal{Q}} = \left\{\mathbf{Q}_\Lambda^{\overline{\boldsymbol{x}}},\quad \Lambda \in \mathscr{E} \text{ and } \overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}\right\}$ *be a quasilocal specification satisfying*

($\mathbf{Q1}$) $\mathbf{Q}_\Lambda^{\boldsymbol{\emptyset}}(\boldsymbol{\emptyset}) > 0$ *for all* $\Lambda \in \mathscr{E}$,

($\mathbf{Q2}$) $\mathbf{Q}_\Lambda^{\boldsymbol{\emptyset}}(\boldsymbol{x}) + \sum\limits_{y\in\mathscr{X}^*} \mathbf{Q}_\Lambda^{\boldsymbol{\emptyset}}(\boldsymbol{x}\oplus y_t) > 0$ *for all* $\Lambda \in \mathscr{E}\setminus\{\emptyset\}$, $t \in \Lambda$ *and* $\boldsymbol{x} \in \mathscr{X}^{\Lambda\setminus t}$.

Then there exists a H*-function* $\boldsymbol{H} = \left\{H_{\boldsymbol{x}},\quad \boldsymbol{x} \in \mathscr{X}^{*J},\ J \in \mathscr{E}\right\}$ *satisfying*

($\boldsymbol{H1}$) $H_{\boldsymbol{x}} + \sum\limits_{y\in\mathscr{X}^*} H_{\boldsymbol{x}\oplus y_t} > 0$ *for all* $\boldsymbol{x} \in \mathscr{X}^{*J}$, $J \in \mathscr{E}$ *and* $t \notin J$,

(***H*2**) *for all* $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ *and* $\boldsymbol{x} \in \mathscr{X}^{\Lambda}$ *there exists uniformly on* $\overline{\boldsymbol{x}} \in \mathscr{X}^{\Lambda^{\mathrm{c}}}$ *the limit*

$$\lim_{I \uparrow \mathbb{Z}^{\nu}} \frac{H_{\boldsymbol{x} \oplus \overline{\boldsymbol{x}}_I}}{\sum\limits_{\boldsymbol{z} \in \mathscr{X}^{\Lambda}} H_{\boldsymbol{z} \oplus \overline{\boldsymbol{x}}_I}} \ ,$$

and such that for all $\Lambda \in \mathscr{E} \setminus \{\emptyset\}$ *and all* $\overline{\boldsymbol{y}} \in \mathscr{X}^{*J}$, $J \in \mathscr{E}$ *such that* $J \subset \Lambda^{\mathrm{c}}$ *we have*

$$\mathbf{Q}_{\Lambda}^{\overline{\boldsymbol{y}}}(\boldsymbol{x}) = \frac{H_{\boldsymbol{x} \oplus \overline{\boldsymbol{y}}}}{\sum\limits_{\boldsymbol{z} \in \mathscr{X}^{\Lambda}} H_{\boldsymbol{z} \oplus \overline{\boldsymbol{y}}}} \ , \quad \boldsymbol{x} \in \mathscr{X}^{\Lambda}. \tag{VI.4}$$

Conversely, if $\boldsymbol{H}$ *is a* H*-function satisfying* (***H*1**) *and* (***H*2**), *then one can find a quasilocal specification* $\boldsymbol{\mathcal{Q}}$ *satisfying* (**Q1**), (**Q2**) *and (VI.4).*

Part II

Identification of random fields

VII. Parametric estimation

In the preceding chapters we have seen different approaches towards description of random fields (P-functions, Q-functions, Q-systems, H-systems and one-point systems). In the remaining part of this work we will consider the problem of statistical identification of random fields. More precisely, we will concentrate on the random fields specified through translation invariant (stationary) one-point systems, since the latter ones provide a parametrization of random fields suitable for statistical inference.

In this chapter we consider the problem of estimation of local one-point systems. The problem is clearly parametric in this case. In the next chapter we will consider the nonparametric problem of estimation of one-point systems in the case they are quasilocal.

For simplicity of notation we will consider the $\{0,1\}$ case but, as we will mention in the last section, the results holds in the case of arbitrary finite state space $\mathscr{X}$. We will construct an estimator as a ratio of some empirical conditional frequencies and prove its exponential consistency and its $\mathbf{L}^p$-consistency for all $p \in (0,\infty)$.

Let us note here, that for maximum likelihood estimators F. Comets in [3] also gets exponential consistency using the theory of large deviations.

Note also, that in general the problem of estimation for Gibbs random fields is complicated by such classical phenomenons of Gibbs random fields theory as non-uniqueness ($|\mathscr{G}| > 1$) and translation invariance breaking. In our work the results are established irrespectively of this aspects of Gibbs random fields theory, since they hold uniformly on $\mathscr{G}$, independently of $|\mathscr{G}| = 1$ or not.

Finally, let us remark that the problem of estimation for Gibbs random fields is very interesting and important, since the results can be used in so-called "image processing". Parametric statistical inference for Gibbs random fields is now quite well developed in classical Gibbsian setup. The actual state of the theory is well presented in the monograph by X. Guyon [14] and the references therein. For

more information on image processing and parametric statistical inference for Gibbs random fields, the interested reader can also see [3], [11], [15], [21], [22] and [26] – [112].

VII.1. Statistical model

We consider vacuum specifications with state space $\mathscr{X} = \{0,1\}$ specified through one-point systems. Let us at first note that a vacuum specification $\boldsymbol{\mathcal{Q}}$ is translation invariant if and only if the corresponding one-point system $\boldsymbol{h}$ is translation invariant, *i.e.*, if we have

$$h_t^{\overline{\boldsymbol{x}}} = h_{t+s}^{\overline{\boldsymbol{x}}+s}$$

for all $t, s \in \mathbb{Z}^\nu$. In this case, clearly one needs to know only the subsystem $\big\{h^{\overline{\boldsymbol{x}}},\ \overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}\big\}$, where $h^{\overline{\boldsymbol{x}}} = h_{\mathbf{0}}^{\overline{\boldsymbol{x}}}$ and $\mathbf{0}$ is the origin of $\mathbb{Z}^\nu$. This subsystem will be the object of statistical interest in the remaining part of this work. Since it determines the whole one-point system we will use the same notation $\boldsymbol{h}$ for it. Condition of the quasilocality in this case will be written in the form

$$\gamma(I) = \sup_{\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}} \big| h^{\overline{\boldsymbol{x}}_I} - h^{\overline{\boldsymbol{x}}} \big| \underset{I \uparrow \mathbb{Z}^\nu}{\longrightarrow} 0.$$

We denote $\mathscr{H} = \big\{\boldsymbol{h} : \boldsymbol{h} \text{ is quasilocal and translation invariant}\big\}$. To any $\boldsymbol{h} \in \mathscr{H}$ we associate some specification $\boldsymbol{\mathcal{Q}}$ and hence some sets $\mathscr{G}(\boldsymbol{h}) = \mathscr{G}(\boldsymbol{\mathcal{Q}})$ and $\mathscr{G}_{\text{t.i.}}(\boldsymbol{h}) = \mathscr{G}_{\text{t.i.}}(\boldsymbol{\mathcal{Q}})$ of random fields described by the Theorem I.8. Recall that non-uniqueness and translation invariance breaking are possible. Note that if $\boldsymbol{h} \in \mathscr{H}$ is strictly positive, then $\boldsymbol{\mathcal{Q}}$ is Gibbsian (for some uniformly convergent potential), and hence we have $\mathscr{G}(\boldsymbol{h}_1) \cap \mathscr{G}(\boldsymbol{h}_2) = \emptyset$ if $\boldsymbol{h}_1 \neq \boldsymbol{h}_2$, which is nothing but identifiability condition for our model.

In this chapter we consider the subclass

$$\mathscr{H}_{\text{loc}} = \big\{\boldsymbol{h} : \boldsymbol{h} \text{ is local and translation invariant}\big\} \subset \mathscr{H}.$$

Suppose $\boldsymbol{h} \in \mathscr{H}_{\text{loc}}$ is some unknown one-point system. As we already know, $\boldsymbol{h}$ induces a set $\mathscr{G}(\boldsymbol{h})$ of Gibbs random fields. In the sequel Λ_n will denote the symmetric cube with the side size n centred at the origin $\mathbf{0}$ of $\mathbb{Z}^\nu$. Here without loss of generality we assume that n is odd. We observe a realisation of some random field $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ in the observation window Λ_n. That is, based on the data $\boldsymbol{x}_n = \boldsymbol{x}_{\Lambda_n} \subset \Lambda_n$ generated by some random field $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ we want to estimate $\boldsymbol{h}$. More formally, the statistical model is

$$\big\{\Omega,\ \mathscr{F},\ \mathbf{P} \in \mathscr{G}(\boldsymbol{h}),\ \boldsymbol{h} \in \mathscr{H}_{A,B}^{V}\big\}$$

where $0 < A \leqslant B < \infty$ are some constants, $\mathbf{0} \in V \in \mathscr{E}$ is some fixed finite set, and $\mathscr{H}_{A,B}^{V}$ is the space of one-point systems satisfying the following conditions.

(C1) $\boldsymbol{h} \in \mathscr{H}_{\mathrm{loc}}$, *i.e.*, $\boldsymbol{h}$ is local and translation invariant.

(C2) For all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu} \setminus \mathbf{0}$ we have $A \leqslant h^{\overline{\boldsymbol{x}}} \leqslant B$.

(C3) The "neighbourhood of locality" is included in $V^{*} = V \setminus \mathbf{0}$, *i.e.*,

$$\sup_{\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu} \setminus \mathbf{0}} \left| h^{\overline{\boldsymbol{x}}_I} - h^{\overline{\boldsymbol{x}}} \right| = 0$$

if $I \supset V^{*}$.

Let us remark that our statistical model is a bit unusual, in the sense that the probability measure $\mathbf{P}$ is not determined by the parameter $\boldsymbol{h}$. Rather, $\boldsymbol{h}$ determines some set $\mathscr{G}(\boldsymbol{h})$ of probability measures. The observations come from an arbitrary element of this set but we are not interested in this element, the only object of interest is the parameter $\boldsymbol{h}$ itself. That is, we want to identify the class $\mathscr{G}(\boldsymbol{h})$ corresponding to (unknown) one-point system $\boldsymbol{h}$, and not a particular element of this class. In fact, this is the reason for which our results hold irrespectively of non-uniqueness and translation invariance breaking. In some sense, if $|\mathscr{G}(\boldsymbol{h})| > 1$, then $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ can be viewed as $\mathbf{P} = \mathbf{P}(\boldsymbol{h}, \mu)$, and only $\boldsymbol{h}$ is the parameter of interest (something like semiparametric statistical problem), while all our considerations will be performed on conditional distributions, the latter ones depending only on $\boldsymbol{h}$, and not on μ.

Remark also, that since **(C1)** and **(C2)** imply that we are in the Gibbsian case, by the Theorem I.8–4 our model is identifiable: $\mathscr{G}(\boldsymbol{h}_1) \cap \mathscr{G}(\boldsymbol{h}_2) = \emptyset$ for $\boldsymbol{h}_1 \neq \boldsymbol{h}_2$. Finally note, that this identifiability will not be used explicitly in establishing our results.

Any real-valued random function $\overline{\boldsymbol{h}}_n = \left\{ \overline{h}_n^{\overline{\boldsymbol{x}}}, \quad \overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu} \setminus \mathbf{0} \right\}$ constructed from $\boldsymbol{x}_n$ is said to be an *estimator* of $\boldsymbol{h}$. The distance between the estimator $\overline{\boldsymbol{h}}_n$ and the true value $\boldsymbol{h}$ is measured in the supremum norm:

$$\left\| \overline{\boldsymbol{h}}_n - \boldsymbol{h} \right\| = \sup_{\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu} \setminus \mathbf{0}} \left| \overline{h}_n^{\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}} \right|.$$

The estimator $\overline{\boldsymbol{h}}_n$ is said to be *consistent*, if for any $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$ we have $\left\| \overline{\boldsymbol{h}}_n - \boldsymbol{h} \right\| \underset{n \to \infty}{\longrightarrow} 0$ in probability, uniformly over $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$, *i.e.*, if for any $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$

and any $\varepsilon > 0$ we have

$$\sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\|\overline{\boldsymbol{h}}_n - \boldsymbol{h}\big\| > \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

The estimator $\overline{\boldsymbol{h}}_n$ is said to be *uniformly consistent*, if it is consistent uniformly on $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$, *i.e.*, if for any $\varepsilon > 0$ we have

$$\sup_{\boldsymbol{h}\in\mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\|\overline{\boldsymbol{h}}_n - \boldsymbol{h}\big\| > \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 0.$$

The estimator $\overline{\boldsymbol{h}}_n$ is said to be $\mathbf{L}^p$*-consistent* for some $p \in (0,\infty)$, if for any $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$ we have $\big\|\overline{\boldsymbol{h}}_n - \boldsymbol{h}\big\| \underset{n\to\infty}{\longrightarrow} 0$ in $\mathbf{L}^p$, uniformly over $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$, *i.e.*, if for any $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$ we have

$$\sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{E}\,\big\|\overline{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p \underset{n\to\infty}{\longrightarrow} 0.$$

The estimator $\overline{\boldsymbol{h}}_n$ is said to be *uniformly* $\mathbf{L}^p$*-consistent* for some $p \in (0,\infty)$, if it is $\mathbf{L}^p$-consistent uniformly on $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$, *i.e.*, if we have

$$\sup_{\boldsymbol{h}\in\mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{E}\,\big\|\overline{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p \underset{n\to\infty}{\longrightarrow} 0.$$

Let us finally note here, that if the random field corresponding to a one-point system $\boldsymbol{h}$ is unique, then all the statistical model, the identifiability and all the notions of consistency regain their classical statistical sense. To guaranty uniqueness one can suppose, for example, that $\boldsymbol{h}$ satisfies the Dobrushin's uniqueness condition.

VII.2. Construction of the estimator

Let us at first note that by (V.5) we have

$$h^{\overline{\boldsymbol{x}}} = h_{\mathbf{0}}^{\overline{\boldsymbol{x}}} = \frac{\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(\mathbf{0})}{\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(\emptyset)} = \frac{\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(1)}{\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(0)}\ . \tag{VII.1}$$

Further, we see that the conditional probabilities $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x)$, $x \in \{0,1\}$, are equal to $\mathbf{P}_{\mathbf{0}|V^*}\big(x \mid \overline{\boldsymbol{x}}_{V^*}\big) = \mathbf{P}\big(\xi_{\mathbf{0}} = x \mid \xi_{V^*} = \overline{\boldsymbol{x}}_{V^*}\big)$. In fact, using total probability formula and the condition **(C3)** we get

$$\begin{aligned}\mathbf{P}_{\mathbf{0}|V^*}\big(x \mid \overline{\boldsymbol{x}}_{V^*}\big) &= \int\limits_{\mathscr{X}^{V^{\mathrm{c}}}} \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}\cup\overline{\boldsymbol{y}}}(x)\ \mathbf{P}_{V^{\mathrm{c}}|V^*}\big(\mathrm{d}\overline{\boldsymbol{y}} \mid \overline{\boldsymbol{x}}_{V^*}\big) = \\ &= \int\limits_{\mathscr{X}^{V^{\mathrm{c}}}} \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x)\ \mathbf{P}_{V^{\mathrm{c}}|V^*}\big(\mathrm{d}\overline{\boldsymbol{y}} \mid \overline{\boldsymbol{x}}_{V^*}\big) = \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x).\end{aligned}$$

Now, if n is large enough, then $\mathbf{P}_{\mathbf{0}|V^*}\big(x \bigm| \overline{\boldsymbol{x}}_{V^*}\big)$ can be estimated by the "empirical conditional frequency" of the value x observed in some point $t \in \Lambda_n$ given that $\overline{\boldsymbol{x}}_{V^*} + t$ is observed on the set $V^* + t$.

More precisely, let $\boldsymbol{x}(n)$ be the periodization on $\mathbb{Z}^\nu$ of the observation $\boldsymbol{x}_n$, that is, $\big(\boldsymbol{x}(n)\big)_{\Lambda_n + n\,t} = \boldsymbol{x}_n + n\,t$ for all $t \in \mathbb{Z}^\nu$. Note that equivalently periodization can be viewed as wrapping the observation $\boldsymbol{x}_n$ on a torus. Now, for every $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$, let us put

$$A^1 = \big\{\boldsymbol{y} \subset \mathbb{Z}^\nu \,:\, \boldsymbol{y}_V = \overline{\boldsymbol{x}}_{V^*} \cup \mathbf{0}\big\} \quad \text{and} \quad A^0 = \big\{\boldsymbol{y} \subset \mathbb{Z}^\nu \,:\, \boldsymbol{y}_V = \overline{\boldsymbol{x}}_{V^*}\big\}.$$

Let us also put

$$N^1 = \sum_{t \in \Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n) - t \in A^1\}} \qquad \text{and} \qquad N^0 = \sum_{t \in \Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n) - t \in A^0\}}.$$

Clearly, N^1 and N^0 are the total numbers of subconfigurations of $\boldsymbol{x}_n$ of the "form" V and equal to $\overline{\boldsymbol{x}}_{V^*} \cup \mathbf{0}$ and $\overline{\boldsymbol{x}}_{V^*}$ respectively.

Now we define our estimator $\widehat{\boldsymbol{h}}_n$ by

$$\widehat{h}_n^{\overline{\boldsymbol{x}}} = \begin{cases} N^1 \big/ N^0 & \text{if } N^0 > 0 \text{ and } N^1 > 0, \\ A & \text{if } N^1 = 0, \\ B & \text{if } N^0 = 0 \text{ (and } N^1 > 0). \end{cases}$$

Note that the cases $N^0 = 0$ and $N^1 = 0$ are asymptotically not important. Moreover, we could have not considered at all the second case, that is, we could have put the estimator still to be $N^1 \big/ N^0 = 0$. Our definition of the estimator pursues rather practical aims, and is motivated by the following reasons: $N^0 = 0$ means that $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(0) \approx 0$ and hence $h^{\overline{\boldsymbol{x}}}$ is "large", while $N^1 = 0$ means that $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(1) \approx 0$ and hence $h^{\overline{\boldsymbol{x}}}$ is "small"; but we know *a priori* that $A \leqslant h^{\overline{\boldsymbol{x}}} \leqslant B$.

Let us note here, that the idea of using empirical conditional frequencies to construct estimators, as well as some results on consistency of estimators of such type for parametric models in the classical Gibbsian setup, can be found in [21], [22], [11], [15] and [14].

VII.3. Asymptotic study of the estimator

In this section we will show the uniform exponential consistency of our estimator, as well as its uniform $\mathbf{L}^p$-consistency. The first one is given by the following

Theorem VII.1 [Uniform exponential consistency of the estimator]**.** — *Assume that $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$ and $\widehat{\boldsymbol{h}}_n$ is our estimator. Then there exist some positive constants $C, \alpha > 0$ such that*

$$\sup_{\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P} \in \mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\| \widehat{\boldsymbol{h}}_n - \boldsymbol{h} \big\| > \varepsilon \Big) \leqslant C \, \mathrm{e}^{-\alpha \, \varepsilon^2 \, n^{\nu}}$$

for all $\varepsilon \in (0\,, 1/2)$ and all $n \in \mathbb{N}$, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly exponentially consistent.

Proof: All throughout the proof C and α denote generic positive constants which can differ from formula to formula (and even in the same formula).

The first component of the proof is the following lemma, giving us a uniform lower bound for the conditional probabilities $\mathbf{Q}_{\Lambda}^{\overline{\boldsymbol{x}}}(\boldsymbol{x})$ and for the probabilities $\mathbf{P}_{\Lambda}(\boldsymbol{x})$.

Lemma VII.2. — *Let $P \in \mathscr{G}(\boldsymbol{h})$ for some $\boldsymbol{h}$ satisfying the condition* **(C2)**. *Then, uniformly on $\boldsymbol{x} \subset \Lambda$ and $\overline{\boldsymbol{x}} \subset \Lambda^{\mathrm{c}}$, we have*

$$\mathbf{Q}_{\Lambda}^{\overline{\boldsymbol{x}}}(\boldsymbol{x}) \geqslant \mathrm{e}^{-b^{\star} \, |\Lambda|} \qquad \textit{and} \qquad \mathbf{P}_{\Lambda}(\boldsymbol{x}) \geqslant \mathrm{e}^{-b^{\star} \, |\Lambda|}$$

where $b^{\star} = \max\big\{ \ln(1+B)\,, \ln(1+B) - \ln A \big\}$.

Proof: The second assertion clearly follows from the first one using the total probability formula. By the same formula and properties of conditional distributions the first assertion clearly can be derived from the bound $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x) \geqslant \mathrm{e}^{-b^{\star}}$ for all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^{\nu} \setminus \mathbf{0}$ and $x \in \{0,1\}$. But by **(C2)** we have

$$\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(1) = \frac{h^{\overline{\boldsymbol{x}}}}{1 + h^{\overline{\boldsymbol{x}}}} \geqslant \frac{A}{1+B} \qquad \text{and} \qquad \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(0) = \frac{1}{1 + h^{\overline{\boldsymbol{x}}}} \geqslant \frac{1}{1+B}$$

and hence

$$\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x) \geqslant \min\left\{ \frac{A}{1+B} \, , \, \frac{1}{1+B} \right\} = \mathrm{e}^{\min\{\ln A - \ln(1+B), -\ln(1+B)\}} = \mathrm{e}^{-b^{\star}}.$$

The lemma is proved. □

Now, let us decompose Λ_n in the following way. We denote $\gamma = \sup_{t \in V} \|t\|$ and, for technical reasons, we suppose that $n = m\,(3\,\gamma + 1)$ for some $m \in \mathbb{N}$. Then Λ_n is partitioned into $m^{\nu} = n^{\nu}/(3\,\gamma+1)^{\nu}$ cubes $D_1, \ldots, D_{m^{\nu}}$ with side $3\,\gamma + 1$. Each D_i contains $(3\,\gamma+1)^{\nu}$ lattice sites. We order sites of each D_i in the

same arbitrary way. Hence, every $t \in \Lambda_n$ can be referred to as a pair (i,j), $i = 1,\dots,m^\nu,\ j = 1,\dots,(3\,\gamma+1)^\nu$, which means j-th site in the cube D_i. In the sequel we will use both the notations t and (i,j) for points of Λ_n.

If we define

$$Y^0_{ij} = \mathbb{1}_{\{\boldsymbol{x}(n)-(i,j)\in A^0\}} \qquad \text{and} \qquad Y^1_{ij} = \mathbb{1}_{\{\boldsymbol{x}(n)-(i,j)\in A^1\}}$$

and

$$N^0_j = \sum_{i=1}^{m^\nu} Y^0_{ij} \qquad \text{and} \qquad N^1_j = \sum_{i=1}^{m^\nu} Y^1_{ij},$$

then N^0 and N^1 from the definition of the estimator will have the form

$$N^0 = \sum_{j=1}^{(3\,\gamma+1)^\nu} N^0_j \qquad \text{and} \qquad N^1 = \sum_{j=1}^{(3\,\gamma+1)^\nu} N^1_j.$$

Note that all Y^0_{ij}, Y^1_{ij}, N^0_j, N^1_j, N^0 and N^1 depend on n, on $\overline{\boldsymbol{x}}_{V^*}$ and on the observation $\boldsymbol{x}_n$.

Now, for any $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$, we can write

$$\begin{aligned}
\left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}}\right| &= \left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}_{V^*}}\right| = \\
&= \mathbb{1}_{\{N^0=0 \text{ or } N^1=0\}}\left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}_{V^*}}\right| + \mathbb{1}_{\{N^0>0,\ N^1>0\}}\left|\frac{\sum_{j=1}^{(3\,\gamma+1)^\nu} N^1_j}{N^0} - h^{\overline{\boldsymbol{x}}_{V^*}}\right| \leqslant \\
&\leqslant \mathbb{1}_{\{N^0=0\}}\left|B - h^{\overline{\boldsymbol{x}}_{V^*}}\right| + \mathbb{1}_{\{N^1=0\}}\left|A - h^{\overline{\boldsymbol{x}}_{V^*}}\right| + \\
&\qquad + \mathbb{1}_{\{N^0>0,\ N^1>0\}} \sum_{j=1}^{(3\,\gamma+1)^\nu}\left|\frac{N^1_j}{N^0} - \frac{N^0_j}{N^0}\,h^{\overline{\boldsymbol{x}}_{V^*}}\right| = \\
&= \mathbb{1}_{\{N^0=0\}}\left|B - h^{\overline{\boldsymbol{x}}_{V^*}}\right| + \mathbb{1}_{\{N^1=0\}}\left|A - h^{\overline{\boldsymbol{x}}_{V^*}}\right| + \\
&\qquad + \sum_{j=1}^{(3\,\gamma+1)^\nu} \mathbb{1}_{\{N^0>0,\ N^1>0,\ N^0_j=0\}}\,\frac{N^1_j}{N^0} + \\
&\qquad + \sum_{j=1}^{(3\,\gamma+1)^\nu} \mathbb{1}_{\{N^0_j>0,\ N^1>0\}}\,\frac{1}{N^0}\left|N^1_j - N^0_j\,h^{\overline{\boldsymbol{x}}_{V^*}}\right| = \\
&= D^1_n(\overline{\boldsymbol{x}}) + D^2_n(\overline{\boldsymbol{x}}) + D^3_n(\overline{\boldsymbol{x}}) + D^4_n(\overline{\boldsymbol{x}}) \qquad \text{(VII.2)}
\end{aligned}$$

with evident notations.

To estimate this four summands we need the following

Lemma VII.3. — *Denote* $\Gamma = \mathrm{e}^{-b^\star\,|V|}$, *let* $\lambda_n = \Gamma\, m^\nu$ *and fix some* $r \in \{0,1\}$. *Then there exist some positive constant* $\alpha > 0$ *such that*

$$\mathbf{P}\left(\frac{N_j^r}{\lambda_n} < 1-\varepsilon\right) \leqslant \mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu},$$

uniformly on $\varepsilon \in (0,1)$, $n \in \mathbb{N}$, $j = 1,\dots,(3\,\gamma+1)^\nu$ *and* $\overline{\boldsymbol{x}}_{V^*} \in \mathscr{X}^{V^*}$.

Proof: For definiteness let us take $r = 0$. We denote by V_{ij} a cube with side $2\,\gamma+1$ centred at (i,j), $i = 1,\dots,m^\nu$, $j = 1,\dots,(3\,\gamma+1)^\nu$, and let $\mathcal{V}_j = \mathbb{Z}^\nu \setminus (V_{1j} \cup \dots \cup V_{m^\nu j})$. Note that Y_{ij}^0 depends only on the restriction of our periodized observation $\boldsymbol{x}(n)$ on the set V_{ij}, and that for $i_1 \neq i_2$ we have $\rho(V_{i_1 j}\,, V_{i_2 j}) \geqslant \gamma+1 > \gamma$. So, the restrictions of our random field on $V_{i_1 j}$ and on $V_{i_2 j}$ are conditionally independent, and hence, for any $\lambda > 0$, we have

$$\mathbf{E}\Big(\mathrm{e}^{-\lambda\,N_j^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) = \prod_{i=1}^{m^\nu} \mathbf{E}\Big(\mathrm{e}^{-\lambda\,Y_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big). \tag{VII.3}$$

Clearly, using the Lemma VII.2, definition of Y_{ij}^0 and total probability formula, we have

$$\mathbf{E}\big(Y_{ij}^0 \;\big|\; \boldsymbol{x}_{\mathcal{V}_j}\big) \geqslant \mathrm{e}^{-b^\star\,|V|} = \Gamma.$$

Furthermore, using Taylor expansion formula, we get

$$\begin{aligned}\mathbf{E}\Big(\mathrm{e}^{-\lambda\,Y_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) &= \mathrm{e}^{-\lambda\,\mathbf{E}\left(Y_{ij}^0 \,|\, \boldsymbol{x}_{\mathcal{V}_j}\right)}\; \mathbf{E}\Big(\mathrm{e}^{-\lambda\,\left(Y_{ij}^0 - \mathbf{E}\left(Y_{ij}^0 \,|\, \boldsymbol{x}_{\mathcal{V}_j}\right)\right)} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) \leqslant \\ &\leqslant \mathrm{e}^{-\lambda\,\Gamma}\left(1 + \frac{\lambda^2}{2}\,\mathrm{e}^{\lambda}\right) \leqslant \exp\Big(-\lambda\Big(\Gamma - \frac{\lambda}{2}\,\mathrm{e}^{\lambda}\Big)\Big).\end{aligned} \tag{VII.4}$$

Finally, combining (VII.3), (VII.4), and using Chebychev's inequality and total probability formula, we get

$$\begin{aligned}\mathbf{P}\left(\frac{N_j^0}{\lambda_n} < 1-\varepsilon\right) &\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\lambda_n}\; \mathbf{E}\,\mathrm{e}^{-\lambda\,N_j^0} \leqslant \\ &\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\,\Gamma\,m^\nu}\; \exp\left(-\lambda\left(\Gamma - \frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right) m^\nu\right) = \\ &= \exp\left(-\lambda\,m^\nu\left(\varepsilon\,\Gamma - \frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right)\right).\end{aligned}$$

Now, choosing $\lambda = \varepsilon\,\Gamma/\mathrm{e} < 1$, we get

$$\mathbf{P}\left(\frac{N_j^0}{\lambda_n} < 1-\varepsilon\right) \leqslant \exp\left(-\frac{\varepsilon\,\Gamma}{\mathrm{e}}\,m^\nu\left(\varepsilon\,\Gamma - \frac{\varepsilon\,\Gamma}{2}\right)\right) = \mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}$$

with $\alpha = \dfrac{\Gamma^2}{2\,\mathrm{e}\,(3\,\gamma+1)^\nu}$. The lemma is proved. □

Using this lemma we clearly get

$$\mathbf{P}\big(N_j^r = 0\big) \leqslant \mathbf{P}\bigg(\frac{N_j^r}{\lambda_n} < 1-\varepsilon\bigg) \leqslant \mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}$$

for all $j = 1, \dots, (3\,\gamma+1)^\nu$ and $r \in \{0,1\}$. Therefore we have

$$\begin{aligned}\mathbf{P}\Big(\big\|D_n^1(\cdot)\big\| > \varepsilon/4\Big) = \mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^1(\overline{\boldsymbol{x}})\big| > \varepsilon/4\Big) \leqslant \\ \leqslant \sum_{\overline{\boldsymbol{x}}_{V^*}\in\mathscr{X}^{V^*}} \mathbf{P}\big(N^0 = 0\big) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}\end{aligned} \tag{VII.5}$$

where we take into account that N^0 depends only on $\overline{\boldsymbol{x}}_{V^*}$, and hence the supremum over $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$ is in fact a maximum over $\overline{\boldsymbol{x}}_{V^*} \in \mathscr{X}^{V^*}$, *i.e.*, a maximum over $2^{|V^*|} = C$ elements.

In exactly the same way we have

$$\mathbf{P}\Big(\big\|D_n^2(\cdot)\big\| > \varepsilon/4\Big) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}, \tag{VII.6}$$

and similarly we get

$$\begin{aligned}\mathbf{P}\Big(\big\|D_n^3(\cdot)\big\| > \varepsilon/4\Big) = \mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^3(\overline{\boldsymbol{x}})\big| > \varepsilon/4\Big) \\ \leqslant \sum_{\overline{\boldsymbol{x}}_{V^*}\in\mathscr{X}^{V^*}} \sum_{j=1}^{(3\,\gamma+1)^\nu} \mathbf{P}\big(N_j^0 = 0\big) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}.\end{aligned} \tag{VII.7}$$

Finally, the last summand is estimated by the following lemma.

Lemma VII.4. — *There exist some positive constants $C, \alpha > 0$ such that*

$$\mathbf{P}\Big(\big\|D_n^4(\cdot)\big\| > \varepsilon/4\Big) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu} \tag{VII.8}$$

for all $\varepsilon \in (0\,,1/2)$ and all $n \in \mathbb{N}$.

Proof: As before, it is sufficient to show that

$$\mathbf{P}\bigg(N_j^0 > 0, \frac{1}{N^0}\Big|N_j^1 - N_j^0\,h^{\overline{\boldsymbol{x}}_{V^*}}\Big| > \frac{\varepsilon}{4\,(3\,\gamma+1)^\nu}\bigg) \leqslant C\,\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}.$$

We have obviously

$$\mathbf{P}\left(N_j^0>0,\frac{1}{N^0}\left|N_j^1-N_j^0\,h^{\overline{\boldsymbol{x}}_{V^*}}\right|>\frac{\varepsilon}{4\,(3\,\gamma+1)^\nu}\right)\leqslant$$

$$\leqslant\mathbf{P}\left(\left|\sum_{i=1}^{m^\nu}\left(Y_{ij}^1-Y_{ij}^0\,h^{\overline{\boldsymbol{x}}_{V^*}}\right)\right|>\frac{\varepsilon\,N^0}{4\,(3\,\gamma+1)^\nu}\right)\leqslant$$

$$\leqslant\mathbf{P}\left(\sum_{j=1}^{(3\,\gamma+1)^\nu}\frac{N_j^0}{\lambda_n}\leqslant(1-\varepsilon)\,(3\,\gamma+1)^\nu\right)+\mathbf{P}\left(\left|\sum_{i=1}^{m^\nu}W_{ij}\right|\geqslant\tau\,\lambda_n\right)$$

where $\tau=\varepsilon\,(1-\varepsilon)/4$ and $W_{ij}=Y_{ij}^1-Y_{ij}^0\,h^{\overline{\boldsymbol{x}}_{V^*}}$. The estimate of the first term easily follows from the preceding lemma. To estimate the second one let us at first note that using translation invariance, total probability formula, the formulas (I.5), (VII.1) and the condition **(C3)** we have

$$\mathbf{E}\Big(Y_{ij}^0\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)\,h^{\overline{\boldsymbol{x}}_{V^*}}=\mathbf{P}_{V\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{V^*}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\,h^{\overline{\boldsymbol{x}}_{V^*}}=$$

$$=\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(0)\;\mathbf{P}_{V^*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{V^*}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times$$

$$\times\,\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}}(1)\Big/\,\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}}(0)=$$

$$=\mathbf{P}_{V^*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{V^*}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\;\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}}(1)$$

$$=\mathbf{P}_{V^*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{V^*}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times$$

$$\times\,\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{V^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(1)=$$

$$=\mathbf{E}\Big(Y_{ij}^1\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big).$$

This implies that

$$\mathbf{E}\Big(W_{ij}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)=\mathbf{E}\Big(Y_{ij}^1\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)-\mathbf{E}\Big(Y_{ij}^0\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)\,h^{\overline{\boldsymbol{x}}_{V^*}}=0$$

and hence, for any $\lambda>0$, using the fact that $|W_{ij}|\leqslant B'=\max\{1,B\}$ and Taylor expansion formula, we get

$$\mathbf{E}\Big(\mathrm{e}^{\lambda\,W_{ij}}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)\leqslant1+\frac{\lambda^2\,B'^2}{2}\,\mathrm{e}^{\lambda\,B'}\leqslant\exp\Big(\frac{\lambda^2\,B'^2}{2}\,\mathrm{e}^{\lambda\,B'}\Big).$$

Finally, using Chebychev's inequality and total probability formula, we get

$$\mathbf{P}\left(\sum_{i=1}^{m^\nu}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant\mathrm{e}^{-\lambda\,\tau\,\lambda_n}\;\mathbf{E}\exp\left(\lambda\sum_{i=1}^{m^\nu}W_{ij}\right)=$$

$$= \mathrm{e}^{-\lambda\,\tau\,\Gamma\,m^\nu}\ \mathbf{E}\left(\prod_{i=1}^{m^\nu}\mathbf{E}\Big(\mathrm{e}^{\lambda\,W_{ij}}\ \Big|\ \xi_{\mathcal{V}_j}\Big)\right)\leqslant$$

$$\leqslant \exp\left(-\lambda\,m^\nu\left(\tau\,\Gamma-\frac{B'^2}{2}\,\lambda\,\mathrm{e}^{\lambda\,B'}\right)\right).$$

Now, choosing $\lambda=\dfrac{\tau\,\Gamma}{B'^2\,\mathrm{e}^{B'}}<1$, we get

$$\mathbf{P}\left(\sum_{i=1}^{m^\nu}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant\exp\left(-\frac{\tau\,\Gamma}{B'^2\,\mathrm{e}^{B'}}\,m^\nu\left(\tau\,\Gamma-\frac{\tau\,\Gamma}{2}\right)\right)\leqslant\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}$$

with $\alpha=\dfrac{\Gamma^2}{128\,B'^2\,\mathrm{e}^{B'}\,(3\,\gamma+1)^\nu}$.

By the same argument we have

$$\mathbf{P}\left(-\sum_{i=1}^{n^{\nu/2}}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant\mathrm{e}^{-\alpha\,\varepsilon^2\,n^\nu}$$

which concludes the proof of the lemma. □

Now, combining (VII.5), (VII.6), (VII.7), (VII.8), and taking into account the inequality (VII.2), we get the assertion of the theorem. The uniformity with respect to $\mathbf{P}\in\mathscr{G}(\boldsymbol{h})$ and $h\in\mathscr{H}^V_{A,B}$ is trivial. The Theorem VII.1 is proved. □

Let us note, that taking a closer look on the proof we can give some explicit constants C and α, even if they are not necessarily the optimal ones. For example, one can take

$$C=2^{|V^*|}\left((3\,\gamma+1)^\nu+1\right)\left((3\,\gamma+1)^\nu+2\right)\quad\text{and}\quad\alpha=\frac{\Gamma^2}{128\,B'^2\,\mathrm{e}^{B'}\,(3\,\gamma+1)^\nu}.$$

Now let us turn to $\mathbf{L}^p$-consistency. The uniform $\mathbf{L}^p$-consistency of our estimator is given by the following

Theorem VII.5 [Uniform $\mathbf{L}^p$-consistency of the estimator]**.** — *Assume that $\boldsymbol{h}\in\mathscr{H}^V_{A,B}$, $\widehat{\boldsymbol{h}}_n$ is our estimator, and fix some $p\in(0,\infty)$. Then, for sufficiently large values of n, we have*

$$\sup_{\boldsymbol{h}\in\mathscr{H}^V_{A,B}}\ \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})}\left(\mathbf{E}\left\|\widehat{\boldsymbol{h}}_n-\boldsymbol{h}\right\|^p\right)^{1/p}\leqslant n^{-(\nu/2-\sigma)}$$

where σ is an arbitrary small positive constant, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly $\mathbf{L}^p$-consistent.

Proof: Let us consider $\varepsilon_n = n^{-(\nu/2-\sigma)}$ with an arbitrary small positive constant σ. Using the preceding theorem we get

$$\begin{aligned}\mathbf{E}\,\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p = \int\limits_{\|\widehat{\boldsymbol{h}}_n-\boldsymbol{h}\|>\varepsilon_n} \big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p \,\mathrm{d}\,\mathbf{P} \quad + \int\limits_{\|\widehat{\boldsymbol{h}}_n-\boldsymbol{h}\|\leqslant\varepsilon_n} \big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p \,\mathrm{d}\,\mathbf{P} \leqslant \\ \leqslant (\max\{n^\nu,B\}+B)^p\ \mathbf{P}\Big(\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\| > \varepsilon_n\Big) + \varepsilon_n^p \leqslant \\ \leqslant C\, n^{\nu\, p}\, \mathrm{e}^{-\alpha\, \varepsilon_n^2\, n^\nu} + \varepsilon_n^p = \\ = C\, n^{\nu\, p}\, \mathrm{e}^{-\alpha\, n^{2\,\sigma}} + n^{-(\nu/2-\sigma)\, p} \leqslant C\, n^{-(\nu/2-\sigma)\, p}\end{aligned}$$

for sufficiently large values of n, where we use the fact that $\boldsymbol{h}$ is bounded by B and $\widehat{\boldsymbol{h}}$ by $\max\{n^\nu,B\}$. The assertion of the theorem follows trivially. □

Remark VII.6. — Note also, that if one enlarges the class $\mathscr{H}_{A,B}^V$ to the class $\mathscr{H}^V$ by replacing the condition **(C2)** by a weaker condition of strict positivity

(C2′) for all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$ we have $h^{\overline{\boldsymbol{x}}} > 0$,

then for any $\boldsymbol{h} \in \mathscr{H}^V$ there exist some constants $A = A(\boldsymbol{h})$ and $B = B(\boldsymbol{h})$ such that the condition **(C2)** is satisfied, and hence one can still obtain (no longer uniform) exponential and $\mathbf{L}^p$ consistencies of our estimator. Clearly, in this setup the definition of our estimator needs to be slightly modified for the cases $N^1 = 0$ and $N^0 = 0$. For example, we can put the estimator to be equal to some arbitrary fixed $\widetilde{h} > 0$ in this cases.

VII.4. Generalizations to the case of arbitrary finite state space

Now let us consider the case of arbitrary finite state space $\mathscr{X}$. As always we suppose that there is some fixed element $\emptyset \in \mathscr{X}$ which is called vacuum and we denote $\mathscr{X}^* = \mathscr{X} \setminus \{\emptyset\}$.

As in the {0,1} case, we consider subsystems $\big\{h^{\overline{\boldsymbol{x}}}(x), \quad x \in \mathscr{X}^*,\ \overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu\setminus\mathbf{0}}\big\}$, where $h^{\overline{\boldsymbol{x}}}(x) = h^{\overline{\boldsymbol{x}}}_{\mathbf{0}}(x)$, of translation invariant one-point systems. The statistical model is

$$\big\{\Omega,\ \mathscr{F},\ \mathbf{P} \in \mathscr{G}(\boldsymbol{h}),\ \boldsymbol{h} \in \mathscr{H}_{A,B}^V\big\}$$

where $0 < A \leqslant B < \infty$ are some constants, $\mathbf{0} \in V \in \mathscr{E}$ is some fixed finite set, and $\mathscr{H}_{A,B}^{V}$ is the space of one-point systems satisfying the following conditions.

(C1) $\boldsymbol{h} \in \mathscr{H}_{\mathrm{loc}}$, *i.e.*, $\boldsymbol{h}$ is local and translation invariant.

(C2) For all $x \in \mathscr{X}^*$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu \setminus \mathbf{0}}$ we have $A \leqslant h^{\overline{\boldsymbol{x}}}(x) \leqslant B$.

(C3) The "neighbourhood of locality" is included in $V^* = V \setminus \mathbf{0}$, *i.e.*,

$$\sup_{x \in \mathscr{X}^*} \sup_{\overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu \setminus \mathbf{0}}} \left| h^{\overline{\boldsymbol{x}}_I}(x) - h^{\overline{\boldsymbol{x}}}(x) \right| = 0$$

if $I \supset V^*$.

The distance between the estimator $\overline{\boldsymbol{h}}_n$ and the true value $\boldsymbol{h}$ is measured in the supremum norm:

$$\left\| \overline{\boldsymbol{h}}_n - \boldsymbol{h} \right\| = \sup_{x \in \mathscr{X}^*} \sup_{\overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu \setminus \mathbf{0}}} \left| \overline{h}_n^{\overline{\boldsymbol{x}}}(x) - h^{\overline{\boldsymbol{x}}}(x) \right|.$$

As before, we let $\boldsymbol{x}(n)$ be the periodization on $\mathbb{Z}^\nu$ of the observation $\boldsymbol{x}_n$, and for every $x \in \mathscr{X}^*$ and $\overline{\boldsymbol{x}} \in \mathscr{X}^{\mathbb{Z}^\nu \setminus \mathbf{0}}$ we put

$$A^x = \left\{ \boldsymbol{y} \in \mathscr{X}^{\mathbb{Z}^\nu} \,:\, \boldsymbol{y}_{\Lambda_k} = \overline{\boldsymbol{x}}_{\Lambda_k^*} \oplus x_{\mathbf{0}} \right\} \quad \text{and} \quad A^{\emptyset} = \left\{ \boldsymbol{y} \in \mathscr{X}^{\mathbb{Z}^\nu} \,:\, \boldsymbol{y}_{\Lambda_k} = \overline{\boldsymbol{x}}_{\Lambda_k^*} \right\}.$$

We also put

$$N^x = \sum_{t \in \Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n) - t \in A^x\}} \qquad \text{and} \qquad N^{\emptyset} = \sum_{t \in \Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n) - t \in A^{\emptyset}\}}.$$

Now we define our estimator $\widehat{\boldsymbol{h}}_n$ by

$$\widehat{h}_n^{\overline{\boldsymbol{x}}}(x) = \begin{cases} N^x / N^{\emptyset} & \text{if } N^{\emptyset} > 0 \text{ and } N^x > 0, \\ A & \text{if } N^x = 0, \\ B & \text{if } N^{\emptyset} = 0 \text{ (and } N^x > 0). \end{cases}$$

In this setup, the theorems corresponding to the {0,1} case hold in the general case without reformulation. That is, we have the following theorems.

Theorem VII.7 [Uniform exponential consistency of the estimator]**.** — *Assume that $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$ and $\widehat{\boldsymbol{h}}_n$ is our estimator. Then there exist some positive constants $C, \alpha > 0$ such that*

$$\sup_{\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P} \in \mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\| \widehat{\boldsymbol{h}}_n - \boldsymbol{h} \big\| > \varepsilon \Big) \leqslant C\, \mathrm{e}^{-\alpha\, \varepsilon^2\, n^\nu}$$

for all $\varepsilon \in (0\,, 1/2)$ and all $n \in \mathbb{N}$, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly exponentially consistent.

Theorem VII.8 [Uniform $\mathbf{L}^p$-consistency of the estimator]**.** — *Assume that* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{V}$, $\widehat{\boldsymbol{h}}_n$ *is our estimator, and fix some* $p \in (0,\infty)$*. Then, for sufficiently large values of* n*, we have*

$$\sup_{\boldsymbol{h}\in\mathscr{H}_{A,B}^{V}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \left(\mathbf{E}\left\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\right\|^p\right)^{1/p} \leqslant n^{-(\nu/2-\sigma)}$$

where σ *is an arbitrary small positive constant, i.e., the estimator* $\widehat{\boldsymbol{h}}_n$ *is uniformly* $\mathbf{L}^p$*-consistent.*

Let us note, that here again one can give some explicit constants C and α. They will be given by the same formulas as in the $\{0,1\}$ case, except that the first term in the expression for C will be equal to $|\mathscr{X}|^{|V^*|}$, and in the Lemma VII.2 we will have $b^\star = \max\bigl\{\ln(1+|\mathscr{X}^*|\,B), \ln(1+|\mathscr{X}^*|\,B) - \ln A\bigr\}$.

Finally note, that the considerations of the Remark VII.6 clearly hold in this general case.

VIII. Nonparametric estimation

In this chapter we consider the problem of nonparametric estimation of quasilocal one-point systems. We construct an estimator by combining the ideas of the previous chapter with the main idea of the method of sieves (introduced by U. Grenander [13]): approximation of infinite-dimensional parameter by finite-dimensional ones. We prove exponential consistency and $\mathbf{L}^p$-consistency, for all $p \in (0,\infty)$, of our sieve estimator in different setups.

Let us note here, that unlike parametric statistical inference for Gibbs random fields, the nonparametric one seems to be less investigated. We can mention here a preprint by C. Ji [15]. He considers a classical Gibbsian setup where the random field is described by an exponentially decreasing pair-interaction potential. For this model he studies the sieve estimator of "local characteristics". The proof presented there needs some rectifications. Our work is similar to [15] in that our one-point system is in fact something similar to local characteristics, and in that we study the sieve estimator. But unlike [15], our setup is much more general and in our case we estimate the object (one-point system) which itself describes the random field.

Let us finally note here, that though we consider in this chapter only the $\{0,1\}$ case, in the setup of the last section of the previous chapter all the results of this chapter are generalized to the case of arbitrary finite state space $\mathscr{X}$ without reformulation.

VIII.1. Statistical model

We adopt here all the notations of the Section VII.1.

Suppose $\boldsymbol{h} \in \mathscr{H}$ is some unknown translation invariant quasilocal one-point system. As we already know, $\boldsymbol{h}$ induces a set $\mathscr{G}(\boldsymbol{h})$ of Gibbs random fields. As before, we observe a realisation of some random field $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ in the observation window Λ_n. That is, based on the data $\boldsymbol{x}_n = \boldsymbol{x}_{\Lambda_n} \subset \Lambda_n$ generated by some

random field $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ we want to estimate $\boldsymbol{h}$. More formally, the statistical model is

$$\left\{\Omega,\ \mathscr{F},\ \mathbf{P} \in \mathscr{G}(\boldsymbol{h}),\ \boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}\right\}$$

where $0 < A \leqslant B < \infty$ are some constants and $\mathscr{H}_{A,B}^{\exp}$ is the space of one-point systems satisfying the following conditions.

(C4) $\boldsymbol{h} \in \mathscr{H}$, *i.e.*, $\boldsymbol{h}$ is quasilocal and translation invariant.

(C2) For all $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$ we have $A \leqslant h^{\overline{\boldsymbol{x}}} \leqslant B$.

(C5) The "rate of quasilocality" is exponential in the sense that

$$\gamma(I) = \sup_{\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}} \left| h^{\overline{\boldsymbol{x}}_I} - h^{\overline{\boldsymbol{x}}} \right| \leqslant c\, \mathrm{e}^{-a\, \rho(I^c \setminus \mathbf{0}\,,\, \mathbf{0})^{\nu+\delta}}$$

where c, a and δ are some positive constants.

Note that c, a and δ are not supposed to be known *a priori* and may differ for different $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$.

Sometimes we would rather use the equivalent form of the condition **(C5)**

$$\varphi(d) = \sup_{I\,:\,\rho(I^c \setminus \mathbf{0}\,,\,\mathbf{0}) \geqslant d}\ \sup_{\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}} \left| h^{\overline{\boldsymbol{x}}_I} - h^{\overline{\boldsymbol{x}}} \right| \leqslant c\, \mathrm{e}^{-a\, d^{\nu+\delta}}$$

and we will call the function $\varphi(\cdot)$ *rate of quasilocality*.

Note that **(C4)** and **(C2)** imply that we are in the Gibbsian case, and hence by the Theorem I.8–4 we have identifiability: $\mathscr{G}(\boldsymbol{h}_1) \cap \mathscr{G}(\boldsymbol{h}_2) = \emptyset$ for $\boldsymbol{h}_1 \neq \boldsymbol{h}_2$. Finally note, that as before this identifiability will not be used explicitly in our demonstrations.

VIII.2. Construction of the sieve estimator

The main idea of the estimator is to take some $k = k(n)$ and approximate $h^{\overline{\boldsymbol{x}}}$ by the ratio of the conditional probabilities with condition in the volume Λ_k^*. For this we use the formula (VII.1) and we approximate the conditional probabilities $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x)$, $x \in \{0,1\}$, by $\mathbf{P}_{\mathbf{0}|\Lambda_k^*}\big(x \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\big)$ where Λ_k is called *sieve* and $k = k(n)$

is called *sieve size* and is supposed to grow fast enough. In fact, using total probability formula and quasilocality condition, we have

$$\mathbf{P}_{\mathbf{0}|\Lambda_k^*}\bigl(x \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr) = \int\limits_{\mathscr{X}^{\Lambda_k^{\mathrm{c}}}} \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}\cup\overline{\boldsymbol{y}}}(x)\, \mathbf{P}_{\Lambda_k^{\mathrm{c}}|\Lambda_k^*}\bigl(\mathrm{d}\overline{\boldsymbol{y}} \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr) \approx$$

$$\approx \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x) \int\limits_{\mathscr{X}^{\Lambda_k^{\mathrm{c}}}} \mathbf{P}_{\Lambda_k^{\mathrm{c}}|\Lambda_k^*}\bigl(\mathrm{d}\overline{\boldsymbol{y}} \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr) = \mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x).$$

On the other hand, if k grows much slower than n, then $\mathbf{P}_{\mathbf{0}|\Lambda_k^*}\bigl(x \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr)$ in its turn can be estimated as before by empirical conditional frequency of the value x observed in some point $t \in \Lambda_n$ given that $\overline{\boldsymbol{x}}_{\Lambda_k^*} + t$ is observed on the set $\Lambda_k^* + t$.

More precisely, we define

$$A^1 = A_k^1 = \bigl\{\boldsymbol{y} \subset \mathbb{Z}^\nu \,:\, \boldsymbol{y}_{\Lambda_k} = \overline{\boldsymbol{x}}_{\Lambda_k^*} \cup \mathbf{0}\bigr\} \quad \text{and} \quad A^0 = A_k^0 = \bigl\{\boldsymbol{y} \subset \mathbb{Z}^\nu \,:\, \boldsymbol{y}_{\Lambda_k} = \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr\}.$$

Further, just as in the parametric case, we put

$$N^1 = \sum_{t\in\Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n)-t\in A^1\}} \qquad \text{and} \qquad N^0 = \sum_{t\in\Lambda_n} \mathbb{1}_{\{\boldsymbol{x}(n)-t\in A^0\}},$$

and finally we define the *sieve estimator* $\widehat{\boldsymbol{h}}_n$ by

$$\widehat{h}_n^{\overline{\boldsymbol{x}}} = \begin{cases} N^1/N^0 & \text{if } N^0 > 0 \text{ and } N^1 > 0, \\ A & \text{if } N^1 = 0, \\ B & \text{if } N^0 = 0 \text{ (and } N^1 > 0). \end{cases}$$

VIII.3. Asymptotic study of the sieve estimator

Note that the definition of the sieve estimator depends on the choice of k. Choosing k too large may result in insufficient number of repetitions of the subconfiguration $\overline{\boldsymbol{x}}_{\Lambda_k^*}$ in $\boldsymbol{x}_n$, *i.e.*, one can have too often $N^0 = 0$ or $N^1 = 0$. On the other hand, choosing k too small may result in poor quality of the approximation $\mathbf{Q}_{\mathbf{0}}^{\overline{\boldsymbol{x}}}(x) \approx \mathbf{P}_{\mathbf{0}|\Lambda_k^*}\bigl(x \bigm| \overline{\boldsymbol{x}}_{\Lambda_k^*}\bigr)$. The following theorem shows a "good" choice of k. As before, we denote $b^\star = \max\bigl\{\ln(1+B)\,,\, \ln(1+B) - \ln A\bigr\}$. We denote also $d^\star = \nu/(2\,b^\star)$.

Theorem VIII.1 [Exponential consistency of the sieve estimator]**.** — *Assume that* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$ *and* $\widehat{\boldsymbol{h}}_n$ *is the sieve estimator with* $k = \big[(d \ln n)^{1/\nu}\big]$ *and* $d \in (0,d^\star)$*. Then, for any* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$ *and any* $\varepsilon > 0$*, there exist some positive constant* $\alpha > 0$ *and some* $n_0 \in \mathbb{N}$ *such that*

$$\sup_{\mathbf{P} \in \mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\| \widehat{\boldsymbol{h}}_n - \boldsymbol{h} \big\| > \varepsilon \Big) \leqslant \mathrm{e}^{-\alpha\, n^{\nu - 2\,d\,b^\star} / \ln n}$$

for all $n > n_0$*, i.e., the estimator* $\widehat{\boldsymbol{h}}_n$ *is exponentially consistent.*

Proof: All throughout the proof C, α and n_0 denote generic positive constants which can differ from formula to formula (and even in the same formula).

The main component of the proof of the theorem is the so-called "conditional mixing lemma".

Lemma VIII.2 [Conditional mixing]**.** — *Let* $\mathbf{P} \in \mathscr{G}(\boldsymbol{h})$ *for some* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$ *and let* $\varphi(\cdot)$ *be the corresponding rate of quasilocality. Let also* $L = L(n) \in \mathbb{N}$ *and let the sets* $R_1 = R_1(n), \dots, R_L = R_L(n)$ *be finite subsets of* $\mathbb{Z}^\nu$ *such that* $\rho(R_{\ell_1}, R_{\ell_2}) \geqslant \beta_n$ *for* $\ell_1 \neq \ell_2$ *where* $\beta_n \underset{n\to\infty}{\longrightarrow} \infty$ *and*

$$\lim_{n\to\infty} \max_{1 \leqslant \ell \leqslant L} |R_\ell|\, \varphi(\beta_n) = 0.$$

Denote $\mathcal{R} = \mathbb{Z}^\nu \setminus (R_1 \cup \dots \cup R_L)$ *and suppose* $u_\ell : \mathscr{X}^{R_\ell} \longrightarrow \mathbb{R}$*,* $\ell = 1, \dots, L$*, are some bounded measurable functions. Then*

$$\mathbf{E}_{R_1 \cup \dots \cup R_L | \mathcal{R}} \left(\prod_{\ell=1}^{L} u_\ell\big(\boldsymbol{x}_{R_\ell}\big) \,\middle|\, \boldsymbol{x}_{\mathcal{R}} \right) = \left(\prod_{\ell=1}^{L} \mathbf{E}_{R_\ell | \mathcal{R}} \Big(u_\ell\big(\boldsymbol{x}_{R_\ell}\big) \,\Big|\, \boldsymbol{x}_{\mathcal{R}} \Big) \right) (1 + \delta_n)^L \tag{VIII.1}$$

where $\mathbf{E}_{R_\ell | \mathcal{R}}$ *is the expectation with respect to* $\mathbf{P}_{R_\ell | \mathcal{R}}$ *and*

$$\delta_n = O\Big(\max_{1 \leqslant \ell \leqslant L} |R_\ell|\, \varphi(\beta_n) \Big). \tag{VIII.2}$$

Proof: First of all let as note that if $\boldsymbol{x}_t = \boldsymbol{y}_t$ for all t such that $\rho(t, \mathbf{0}) \geqslant d$ then by **(C2)** and **(C5)** we have

$$\left| \ln \frac{h^{\boldsymbol{y}}}{h^{\boldsymbol{x}}} \right| = \big| \ln h^{\boldsymbol{y}} - \ln h^{\boldsymbol{x}} \big| \leqslant C \big| h^{\boldsymbol{y}} - h^{\boldsymbol{x}} \big| \leqslant C\, \varphi(d).$$

Now suppose $K_1 = K_1(n)$, $K_2 = K_2(n)$ and $K_3 = K_3(n)$ form a disjoint decomposition of $\mathbb{Z}^\nu$ such that $K_1 \in \mathscr{E}$ and $\rho(K_1,K_2) \geqslant \beta_n$. Then, using

translation invariance and the formula (V.2), for all $\boldsymbol{x}, \boldsymbol{x}' \subset \mathbb{Z}^\nu$ we easily get

$$\left| \ln \frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} \right| \leqslant C \, |\boldsymbol{x}_{K_1}| \, \varphi(\beta_n) \leqslant C \, |K_1| \, \varphi(\beta_n).$$

If, moreover, $|K_1| \, \varphi(\beta_n) \underset{n \to \infty}{\longrightarrow} 0$, then clearly

$$\left| \frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} - 1 \right| = O\Big(|K_1| \, \varphi(\beta_n) \Big).$$

Now we can see that for all $\boldsymbol{x}, \boldsymbol{x}' \subset \mathbb{Z}^\nu$

$$\frac{\mathbf{Q}_{K_1}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}(\boldsymbol{x}_{K_1})}{\mathbf{Q}_{K_1}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}(\boldsymbol{x}_{K_1})} = \frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}} \sum\limits_{S \subset K_1} H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}} \sum\limits_{J \subset K_1} H_J^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} =$$

$$= \frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} \sum_{S \subset K_1} \frac{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{\sum\limits_{J \subset K_1} H_J^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} \frac{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}}{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} =$$

$$= \left(\left(\frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} - 1 \right) + 1 \right) \sum_{S \subset K_1} \mathbf{Q}_{K_1}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}(\boldsymbol{x}_{K_3}) \left(\left(\frac{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}}{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} - 1 \right) + 1 \right) =$$

$$= \left(\frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} - 1 \right) \sum_{S \subset K_1} \mathbf{Q}_{K_1}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}(\boldsymbol{x}_{K_3}) \left(\frac{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}}{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} - 1 \right) +$$

$$+ \left(\frac{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}}{H_{\boldsymbol{x}_{K_1}}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}} - 1 \right) + \sum_{S \subset K_1} \mathbf{Q}_{K_1}^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}(\boldsymbol{x}_{K_3}) \left(\frac{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}_{K_2}}}{H_S^{\boldsymbol{x}_{K_3} \cup \boldsymbol{x}'_{K_2}}} - 1 \right) + 1 =$$

$$= \Delta_n + 1 \tag{VIII.3}$$

where $\Delta_n = O\Big(|K_1| \, \varphi(\beta_n) \Big)$. Using the last formula and the total probability formula we get for all $\ell = 1, \ldots, n$

$$\mathbf{P}_{R_\ell | \mathcal{R} \cup R_1 \cup \cdots \cup R_{\ell-1}} \big(\boldsymbol{x}_{R_\ell} \bigm| \boldsymbol{x}_{\mathcal{R}} \cup \boldsymbol{x}_{R_1} \cup \cdots \cup \boldsymbol{x}_{R_{\ell-1}} \big) = \mathbf{P}_{R_\ell | \mathcal{R}} \big(\boldsymbol{x}_{R_\ell} \bigm| \boldsymbol{x}_{\mathcal{R}} \big) \, (1 + \delta_n)$$

where δ_n satisfies (VIII.2). Multiplying this relations over $\ell = 1, \ldots, n$ we get

$$\mathbf{P}_{R_1 \cup \cdots \cup R_L | \mathcal{R}} \big(\boldsymbol{x}_{R_1} \cup \cdots \cup \boldsymbol{x}_{R_L} \mid \boldsymbol{x}_{\mathcal{R}} \big) = \left(\prod_{\ell=1}^{L} \mathbf{P}_{R_\ell | \mathcal{R}} \big(\boldsymbol{x}_{R_\ell} \bigm| \boldsymbol{x}_{\mathcal{R}} \big) \right) (1 + \delta_n)^L$$

which implies (VIII.1). The lemma is proved. □

In order to use the conditional mixing lemma, let us decompose Λ_n in the following way. For technical reasons suppose $n = 2\,m\,k$ for some $m \in \mathbb{N}$. Then Λ_n is partitioned into $m^\nu = n^\nu/(2\,k)^\nu$ cubes $D_1, \dots, D_{m^\nu}$ with side $2\,k$. Each D_i contains $(2\,k)^\nu$ lattice sites. We order sites of each D_i in the same arbitrary way. Hence, every $t \in \Lambda_n$ can be referred to as a pair (i,j), $i = 1, \dots, m^\nu$, $j = 1, \dots, (2\,k)^\nu$, which means j-th site in the cube D_i. In the sequel we will use both the notations t and (i,j) for points of Λ_n.

If we define

$$Y_{ij}^0 = \mathbb{1}_{\{\boldsymbol{x}(n)-(i,j)\in A^0\}} \qquad \text{and} \qquad Y_{ij}^1 = \mathbb{1}_{\{\boldsymbol{x}(n)-(i,j)\in A^1\}}$$

and

$$N_j^0 = \sum_{i=1}^{m^\nu} Y_{ij}^0 \qquad \text{and} \qquad N_j^1 = \sum_{i=1}^{m^\nu} Y_{ij}^1,$$

then N^0 and N^1 from the definition of the sieve estimator will have the form

$$N^0 = \sum_{j=1}^{(2\,k)^\nu} N_j^0 \qquad \text{and} \qquad N^1 = \sum_{j=1}^{(2\,k)^\nu} N_j^1.$$

Note that all Y_{ij}^0, Y_{ij}^1, N_j^0, N_j^1, N^0 and N^1 depend on n, on $\overline{\boldsymbol{x}}_{\Lambda_k^*}$ and on the observation $\boldsymbol{x}_n$.

Now, for any $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$, we can write

$$\begin{aligned}
\left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}}\right| &\leqslant \left|h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} - h^{\overline{\boldsymbol{x}}}\right| + \left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| = \\
&= \left|h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} - h^{\overline{\boldsymbol{x}}}\right| + \mathbb{1}_{\{N^0=0 \text{ or } N^1=0\}}\left|\widehat{h}_n^{\,\overline{\boldsymbol{x}}} - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| + \\
&\qquad + \mathbb{1}_{\{N^0>0,\ N^1>0\}}\left|\frac{\sum_{j=1}^{(2\,k)^\nu} N_j^1}{N^0} - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| \leqslant \\
&\leqslant \left|h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} - h^{\overline{\boldsymbol{x}}}\right| + \mathbb{1}_{\{N^0=0\}}\left|B - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| + \mathbb{1}_{\{N^1=0\}}\left|A - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| + \\
&\qquad + \mathbb{1}_{\{N^0>0,\ N^1>0\}} \sum_{j=1}^{(2\,k)^\nu} \left|\frac{N_j^1}{N^0} - \frac{N_j^0}{N^0}\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| = \\
&= \left|h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} - h^{\overline{\boldsymbol{x}}}\right| + \mathbb{1}_{\{N^0=0\}}\left|B - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| + \mathbb{1}_{\{N^1=0\}}\left|A - h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| +
\end{aligned}$$

$$+\sum_{j=1}^{(2\,k)^\nu}\mathbb{1}_{\{N^0>0,\ N^1>0,\ N_j^0=0\}}\,\frac{N_j^1}{N^0}+$$

$$+\sum_{j=1}^{(2\,k)^\nu}\mathbb{1}_{\{N_j^0>0,\ N^1>0\}}\,\frac{1}{N^0}\left|N_j^1-N_j^0\,h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right|=$$

$$=D_n^1(\overline{\boldsymbol{x}})+D_n^2(\overline{\boldsymbol{x}})+D_n^3(\overline{\boldsymbol{x}})+D_n^4(\overline{\boldsymbol{x}})+D_n^5(\overline{\boldsymbol{x}}) \tag{VIII.4}$$

with evident notations.

First of all, by **(C5)** we have

$$\left\|D_n^1(\cdot)\right\|=\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\left|h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}-h^{\overline{\boldsymbol{x}}}\right|\leqslant\varphi(k)\leqslant c\,\mathrm{e}^{-a\,k^{\nu+\delta}}\underset{n\to\infty}{\longrightarrow}0$$

and hence

$$\mathbf{P}\Big(\left\|D_n^1(\cdot)\right\|>\varepsilon/5\Big)=0 \tag{VIII.5}$$

for $n\geqslant n_0$.

To estimate the remaining summands we need the following

Lemma VIII.3. — *Denote* $\Gamma(n)=n^{-d\,b^\star}$, *let* $\lambda_n=\Gamma(n)\,m^\nu=n^{\nu-d\,b^\star}/(2\,k)^\nu$ *and fix some* $r\in\{0,1\}$. *Then, for any* $\varepsilon\in(0,1)$, *there exist some positive constant* $\alpha>0$ *and some* $n_0\in\mathbb{N}$ *such that*

$$\mathbf{P}\left(\frac{N_j^r}{\lambda_n}<1-\varepsilon\right)\leqslant\mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n},$$

uniformly on $n>n_0$, $j=1,\dots,(2\,k)^\nu$ *and* $\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}$.

Proof: For definiteness let us take $r=0$. We denote by V_{ij} a cube with side k centred at (i,j), $i=1,\dots,m^\nu$, $j=1,\dots,(2\,k)^\nu$, and let $\mathcal{V}_j=\mathbb{Z}^\nu\setminus(V_{1j}\cup\dots\cup V_{m^\nu j})$. Note that Y_{ij}^0 depends only on the restriction of our periodized observation $\boldsymbol{x}(n)$ on the set V_{ij} and that for $i_1\neq i_2$ we have $\rho(V_{i_1j}\,,V_{i_2j})\geqslant 2\,k-k=k$. So, for any $\lambda>0$, it follows from the conditional mixing lemma that

$$\mathbf{E}\Big(\mathrm{e}^{-\lambda\,N_j^0}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)=(1+\delta_n)^{m^\nu}\prod_{i=1}^{m^\nu}\mathbf{E}\Big(\mathrm{e}^{-\lambda\,Y_{ij}^0}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big) \tag{VIII.6}$$

with $\delta_n=O\big(k^\nu\,\varphi(k)\big)=O\Big(d\,\ln n\,c\,\mathrm{e}^{-a\,k^{\nu+\delta}}\Big)=o\big(n^{-\beta}\big)$ for all $\beta>0$.

Clearly, using the Lemma VII.2, definition of Y_{ij}^0 and total probability formula, we have

$$\mathbf{E}\big(Y_{ij}^0\;\big|\;\boldsymbol{x}_{\mathcal{V}_j}\big)\geqslant\mathrm{e}^{-b^\star|\Lambda_k|}\geqslant\mathrm{e}^{-b^\star d\,\ln n}=\Gamma(n).$$

Furthermore, using Taylor expansion formula, we get

$$\begin{aligned}\mathbf{E}\Big(\mathrm{e}^{-\lambda\, Y_{ij}^0}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big) &= \mathrm{e}^{-\lambda\;\mathbf{E}\left(Y_{ij}^0\middle|\boldsymbol{x}_{\mathcal{V}_j}\right)}\;\mathbf{E}\Big(\mathrm{e}^{-\lambda\left(Y_{ij}^0-\mathbf{E}\left(Y_{ij}^0\middle|\boldsymbol{x}_{\mathcal{V}_j}\right)\right)}\;\Big|\;\boldsymbol{x}_{\mathcal{V}_j}\Big)\leqslant\\ &\leqslant \mathrm{e}^{-\lambda\,\Gamma(n)}\left(1+\frac{\lambda^2}{2}\,\mathrm{e}^{\lambda}\right)\leqslant\exp\!\left(-\lambda\Big(\Gamma(n)-\frac{\lambda}{2}\,\mathrm{e}^{\lambda}\Big)\right).\end{aligned}\tag{VIII.7}$$

Finally, combining (VIII.6), (VIII.7), and using Chebychev's inequality and total probability formula, for sufficiently large values of n we get

$$\begin{aligned}\mathbf{P}\left(\frac{N_j^0}{\lambda_n}<1-\varepsilon\right)&\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\lambda_n}\;\mathbf{E}\,\mathrm{e}^{-\lambda\,N_j^0}\leqslant\\ &\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\,\Gamma(n)\,m^\nu}\;\exp\!\left(-\lambda\left(\Gamma(n)-\frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right)m^\nu\right)(1+\delta_n)^{m^\nu}\leqslant\\ &\leqslant C\,\exp\!\left(-\lambda\,m^\nu\left(\varepsilon\,\Gamma(n)-\frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right)\right).\end{aligned}$$

Now, choosing $\lambda=\varepsilon\,\Gamma(n)/\mathrm{e}=\varepsilon\,n^{-d\,b^\star}/\mathrm{e}<1$, for sufficiently large values of n we get

$$\begin{aligned}\mathbf{P}\left(\frac{N_j^0}{\lambda_n}<1-\varepsilon\right)&\leqslant C\,\exp\!\left(-\frac{\varepsilon\,n^{-d\,b^\star}}{\mathrm{e}}\,\frac{n^\nu}{2^\nu\,d\,\ln n}\left(\varepsilon\,n^{-d\,b^\star}-\frac{\varepsilon\,n^{-d\,b^\star}}{2}\right)\right)\leqslant\\ &\leqslant \mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n}\end{aligned}$$

with an arbitrary $\alpha<\dfrac{\varepsilon^2}{2^{\nu+1}\,\mathrm{e}\,d}$. The lemma is proved. □

Using this lemma we clearly get

$$\mathbf{P}\big(N_j^r=0\big)\leqslant\mathbf{P}\left(\frac{N_j^r}{\lambda_n}<1-\varepsilon\right)\leqslant\mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n}$$

for all $j=1,\dots,(2\,k)^\nu$, $r\in\{0,1\}$ and for sufficiently large values of n. Therefore, for sufficiently large values of n, we have

$$\begin{aligned}\mathbf{P}\Big(\big\|D_n^2(\cdot)\big\|>\varepsilon/5\Big)&=\mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^2(\overline{\boldsymbol{x}})\big|>\varepsilon/5\Big)\leqslant\\ &\leqslant\sum_{\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}}\mathbf{P}\big(N^0=0\big)\leqslant\mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n}\end{aligned}\tag{VIII.8}$$

where we take into account that N^0 depends only on $\overline{\boldsymbol{x}}_{\Lambda_k^*}$, and hence the supremum over $\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}$ is in fact a maximum over $\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}$, *i.e.*, a maximum over $2^{|\Lambda_k^*|}\leqslant 2^{d\,\ln n}$ elements.

In exactly the same way we have

$$\mathbf{P}\Big(\big\|D_n^3(\cdot)\big\| > \varepsilon/5\Big) \leqslant \mathrm{e}^{-\alpha\, n^{\nu-2\,d\,b^\star}/\ln n}, \tag{VIII.9}$$

and similarly we get

$$\begin{aligned}\mathbf{P}\Big(\big\|D_n^4(\cdot)\big\| > \varepsilon/5\Big) = \mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^4(\overline{\boldsymbol{x}})\big| > \varepsilon/5\Big) \leqslant \\ \leqslant \sum_{\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}}\sum_{j=1}^{(2\,k)^\nu}\mathbf{P}\big(N_j^0=0\big) \leqslant \mathrm{e}^{-\alpha\, n^{\nu-2\,d\,b^\star}/\ln n}.\end{aligned} \tag{VIII.10}$$

Finally, the last summand is estimated by the following lemma.

Lemma VIII.4. — *For any $\varepsilon\in(0,1)$ there exist some positive constant $\alpha>0$ and some $n_0\in\mathbb{N}$ such that*

$$\mathbf{P}\Big(\big\|D_n^5(\cdot)\big\| > \varepsilon/5\Big) \leqslant \mathrm{e}^{-\alpha\, n^{\nu-2\,d\,b^\star}/\ln n} \tag{VIII.11}$$

for all $n>n_0$

Proof: As before, it is sufficient to show that

$$\mathbf{P}\left(N_j^0>0, \frac{1}{N^0}\left|N_j^1 - N_j^0\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| > \frac{\varepsilon}{5\,(2\,k)^\nu}\right) \leqslant \mathrm{e}^{-\alpha\, n^{\nu-2\,d\,b^\star}/\ln n}.$$

We have obviously

$$\begin{aligned}\mathbf{P}\left(N_j^0>0, \frac{1}{N^0}\left|N_j^1 - N_j^0\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| > \frac{\varepsilon}{5\,(2\,k)^\nu}\right) \leqslant \\ \leqslant \mathbf{P}\left(\left|\sum_{i=1}^{m^\nu}\Big(Y_{ij}^1 - Y_{ij}^0\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\Big)\right| > \frac{\varepsilon\, N^0}{5\,(2\,k)^\nu}\right) \leqslant \\ \leqslant \mathbf{P}\left(\sum_{j=1}^{(2\,k)^\nu}\frac{N_j^0}{\lambda_n} \leqslant (1-\varepsilon)\,(2\,k)^\nu\right) + \mathbf{P}\left(\left|\sum_{i=1}^{m^\nu} W_{ij}\right| \geqslant \tau\,\lambda_n\right)\end{aligned}$$

where $\tau=\varepsilon\,(1-\varepsilon)/5$ and $W_{ij}=Y_{ij}^1 - Y_{ij}^0\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}$. The estimate of the first term easily follows from the preceding lemma. To estimate the second one let us at

first note that using translation invariance, total probability formula and the formulas (I.5), (VII.1) and (VIII.3) we have

$$\begin{aligned}
\mathbf{E}\Big(Y_{ij}^0 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} &= \mathbf{P}_{\Lambda_k\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} = \\
&= \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(0)\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times \\
&\qquad \times \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(1)\Big/\, \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(0) = \\
&= (1+\rho_n)\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\; \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(1) \\
&= (1+\rho_n)^2\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times \\
&\qquad \times \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(1) = \\
&= \mathbf{E}\Big(Y_{ij}^1 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\,(1+\rho_n)
\end{aligned}$$

where $\rho_n = O\big(\varphi(k)\big) = O\Big(c\,\mathrm{e}^{-a\,k^{\nu+\delta}}\Big) = o\big(n^{-\beta}\big)$ for all $\beta > 0$.

This implies that

$$\mathbf{E}\Big(W_{ij} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) = \mathbf{E}\Big(Y_{ij}^1 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) - \mathbf{E}\Big(Y_{ij}^0 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} = O(\rho_n)$$

and hence, for any $\lambda > 0$, using the fact that $-B \leqslant W_{ij} \leqslant 1$ and Taylor expansion formula, we get

$$\begin{aligned}
\mathbf{E}\Big(\mathrm{e}^{\lambda\, W_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) &= \mathrm{e}^{\lambda\; \mathbf{E}\left(W_{ij}^0\left|\boldsymbol{x}_{\mathcal{V}_j}\right.\right)}\; \mathbf{E}\Big(\mathrm{e}^{\lambda\,\left(W_{ij}^0-\mathbf{E}\left(W_{ij}^0\left|\boldsymbol{x}_{\mathcal{V}_j}\right.\right)\right)} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) \leqslant \\
&\leqslant \mathrm{e}^{\lambda\, O(\rho_n)}\left(1+\frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right) \leqslant \\
&\leqslant \exp\left(\lambda\, O(\rho_n) + \frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right).
\end{aligned}$$

Finally, using Chebychev's inequality, total probability formula and conditional mixing lemma, we get

$$\begin{aligned}
\mathbf{P}\left(\sum_{i=1}^{m^\nu} W_{ij} \geqslant \tau\,\lambda_n\right) &\leqslant \mathrm{e}^{-\lambda\,\tau\,\lambda_n}\; \mathbf{E}\exp\left(\lambda\,\sum_{i=1}^{m^\nu} W_{ij}\right) \leqslant \\
&\leqslant \mathrm{e}^{-\lambda\,\tau\,\Gamma(n)\,m^\nu}\; \mathbf{E}\left(\prod_{i=1}^{m^\nu}\mathbf{E}\Big(\lambda\,\mathrm{e}^{W_{ij}} \;\Big|\; \xi_{\mathcal{V}_j}\Big)\right)(1+\delta_n)^{m^\nu} \leqslant
\end{aligned}$$

$$\leqslant C\,\mathrm{e}^{-\lambda\,\tau\,n^{-d\,b^\star}\,m^\nu}\left(\exp\left(\lambda\,O(\rho_n)+\frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right)\right)^{m^\nu}\leqslant$$

$$\leqslant C\,\exp\left(-\lambda\,m^\nu\left(\tau\,n^{-d\,b^\star}-\frac{(B+1)^2}{2}\,\lambda\,\mathrm{e}^{\lambda\,(B+1)}-O(\rho_n)\right)\right).$$

Now, choosing $\lambda=\dfrac{\tau\,n^{-d\,b^\star}}{(B+1)^2\,\mathrm{e}^{B+1}}<1$, we get

$$\mathbf{P}\left(\sum_{i=1}^{m^\nu}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant C\,\exp\left(-\frac{\tau\,n^{-d\,b^\star}}{(B+1)^2\,\mathrm{e}^{B+1}}\,\frac{n^\nu}{2^\nu\,d\,\ln n}\,\frac{\tau\,n^{-d\,b^\star}}{2}\right)\leqslant$$

$$\leqslant\mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n}$$

with an arbitrary $\alpha<\dfrac{\tau^2}{2^{\nu+1}\,(B+1)^2\,\mathrm{e}^{B+1}\,d}$.

By the same argument we have

$$\mathbf{P}\left(-\sum_{i=1}^{m^\nu}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant\mathrm{e}^{-\alpha\,n^{\nu-2\,d\,b^\star}/\ln n}$$

which concludes the proof of the lemma. □

Now, combining (VIII.5), (VIII.8), (VIII.9), (VIII.10), (VIII.11) and taking into account the inequality (VIII.4), we get the assertion of the theorem. The uniformity on $\mathbf{P}\in\mathscr{G}(\boldsymbol{h})$ is trivial. The Theorem VIII.1 is proved. □

Let us note, that from the details of the proof it clearly follows some explicit expression for the constant α. For example, if $\varepsilon\in(0,1)$, then one can take an arbitrary

$$\alpha<\frac{\tau^2}{2^{\nu+1}\,(B+1)^2\,\mathrm{e}^{B+1}\,d}.$$

Note also, that taking a closer look on the proof we can give a "more precise" bound on the rate of consistency, showing explicitly the dependence of the rate on ε. That is, for $\varepsilon\in(0\,,1/2)$, we have the bound

$$\mathbb{1}_{\left\{6\,c\,n^{-a\,d\,(d\,\ln n)^{\delta/\nu}}>\varepsilon\right\}}+\psi_n\,\exp\left\{-\alpha\,\varepsilon^2\,n^{\nu-2\,d\,b^\star}/\ln n+O(\rho_n)\,\beta\,\varepsilon\,n^{\nu-d\,b^\star}/\ln n\right\}$$

where $\alpha=\dfrac{1}{25\cdot2^{\nu+3}\,(B+1)^2\,\mathrm{e}^{B+1}\,d}$, $\beta=\dfrac{1}{5\cdot2^{\nu+1}\,(B+1)^2\,\mathrm{e}^{B+1}\,d}$ and the sequence ψ_n is given by $\psi_n=2^{d\,\ln n}\left(2^\nu\,d\,\ln n+1\right)\left(2^\nu\,d\,\ln n+2\right)$.

Using this last bound, just as in the parametric case, we easily obtain the following

Theorem VIII.5 [$\mathbf{L}^p$-consistency of the sieve estimator]**.** — *Assume that* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$, $\widehat{\boldsymbol{h}}_n$ *is the sieve estimator with* $k = \big[(d \ln n)^{1/\nu}\big]$ *and* $d \in (0,d^\star)$*, and fix some* $p \in (0,\infty)$*. Then, for any* $\boldsymbol{h} \in \mathscr{H}_{A,B}^{\exp}$ *and for sufficiently large values of* n*, we have*

$$\sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \Big(\mathbf{E}\,\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\|^p\Big)^{1/p} \leqslant n^{-(\nu/2 - d\,b^\star - \sigma)}$$

where σ *is an arbitrary small positive constant, i.e., the estimator* $\widehat{\boldsymbol{h}}_n$ *is* $\mathbf{L}^p$*-consistent.*

Remark, that unlike the parametric case, the condition **(C2)** is really important here, that is, the considerations of the Remark VII.6 do not hold in this case. Indeed, the constants A and B are present in the rates of consistency (under the form of $b^\star$) and even in the definition of the estimator (under the form of $d^\star$).

Let us finally note here, that the consistencies of the sieve estimator proved in the Theorems VIII.1 and VIII.5 can be trivially straightened to be uniform, if we consider a narrower class of one-point systems by fixing not only the constants A and B from the condition **(C2)**, but also the constants a, c, and δ from the condition **(C5)**. More precisely, let $\widetilde{\mathscr{H}} = \widetilde{\mathscr{H}}\big(A, B, a, c, \delta\big)$ be the class of one-point systems satisfying **(C4)**, **(C2)** and **(C5)** with some *a priori* fixed constants A, B, a, c and δ. Then the following theorems hold.

Theorem VIII.6 [Uniform exponential consistency of the sieve estimator]**.** — *Assume that* $\boldsymbol{h} \in \widetilde{\mathscr{H}}$ *and* $\widehat{\boldsymbol{h}}_n$ *is the sieve estimator with* $k = \big[(d \ln n)^{1/\nu}\big]$ *and* $d \in (0,d^\star)$*. Then there exist some positive constant* $\alpha > 0$ *and some* $n_0 \in \mathbb{N}$ *such that*

$$\sup_{\boldsymbol{h}\in\widetilde{\mathscr{H}}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{P}\Big(\big\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\big\| > \varepsilon\Big) \leqslant \mathrm{e}^{-\alpha\, n^{\nu - 2\,d\,b^\star}/\ln n}$$

for all $\varepsilon \in (0\,, 1/2)$ *and all* $n > n_0$*, i.e., the estimator* $\widehat{\boldsymbol{h}}_n$ *is uniformly exponentially consistent.*

Theorem VIII.7 [Uniform $\mathbf{L}^p$-consistency of the sieve estimator]**.** — *Assume that* $\boldsymbol{h} \in \widetilde{\mathscr{H}}$, $\widehat{\boldsymbol{h}}_n$ *is the sieve estimator with* $k = \big[(d \ln n)^{1/\nu}\big]$ *and* $d \in (0,d^\star)$*,*

and fix some $p \in (0,\infty)$. Then, for sufficiently large values of n, we have

$$\sup_{\boldsymbol{h}\in\widetilde{\mathscr{H}}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \Bigl(\mathbf{E}\,\bigl\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\bigr\|^p\Bigr)^{1/p} \leqslant n^{-(\nu/2 - d\,b^\star - \sigma)}$$

where σ is an arbitrary small positive constant, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly $\mathbf{L}^p$-consistent.

VIII.4. About a different choice of the sieve size

Let us note that all the bounds on the rates of consistency obtained in the previous section are "slowered" by the constant d from the definition of the sieve size k. Hence, one can think about getting rid of the terms containing d by slightly modifying the choice of the sieve size k. In fact, we will show below that in the case of the space $\widetilde{\mathscr{H}}$, by putting $k = \bigl[(\ln n)^{1/(\nu+\delta/2)}\bigr]$, one can get almost the same bounds on the rates of consistency as in parametric case. Note that we no longer put d in the definition of k. The reason for this is the fact that even if we have put it, it would not be present in the rates of consistency.

As before, we denote $b^\star = \max\bigl\{\ln(1+B), \ln(1+B) - \ln A\bigr\}$. We also denote

$$\Gamma(n) = n^{-b^\star\,(\ln n)^{-\delta/(2\,\nu+\delta)}}, \quad \gamma(n) = (\ln n)^{1-\delta/(2\,\nu+\delta)} \quad \text{and} \quad \varkappa(n) = \frac{\Gamma^2(n)}{\gamma(n)}.$$

One can easily verify that the functions $\Gamma(n)$, $\gamma(n)$ and $\varkappa(n)$ are *slowly varying* (in the sense of Karamata), *i.e.*, for any $c > 0$ we have, for example, $\varkappa(c\,n)/\varkappa(n) \underset{n\to\infty}{\longrightarrow} 1$. Moreover, we have $\Gamma(n) \underset{n\to\infty}{\longrightarrow} 0$ and $\varkappa(n) \underset{n\to\infty}{\longrightarrow} 0$. Let us note here, that since $\Gamma(n)$ and $\varkappa(n)$ are slowly varying functions, then they tend to 0 slower than $n^{-\beta}$ for all $\beta > 0$. Similarly, we have $\gamma(n) \underset{n\to\infty}{\longrightarrow} \infty$, and this convergence is slower than n^β for all $\beta > 0$.

Theorem VIII.8 [Uniform exponential consistency of the sieve estimator]**.** — *Assume that $\boldsymbol{h} \in \widetilde{\mathscr{H}}$ and $\widehat{\boldsymbol{h}}_n$ is the sieve estimator with $k = \bigl[(\ln n)^{1/(\nu+\delta/2)}\bigr]$. Then, for any $\varepsilon > 0$, there exist some positive constant $\alpha > 0$ and some $n_0 \in \mathbb{N}$ such that*

$$\sup_{\boldsymbol{h}\in\widetilde{\mathscr{H}}} \sup_{\mathbf{P}\in\mathscr{G}(\boldsymbol{h})} \mathbf{P}\Bigl(\bigl\|\widehat{\boldsymbol{h}}_n - \boldsymbol{h}\bigr\| > \varepsilon\Bigr) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}$$

for all $n > n_0$, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly exponentially consistent.

Proof: All throughout the proof C, α and n_0 denote generic positive constants which can differ from formula to formula (and even in the same formula).

As in the proof of the theorem VIII.1, we apply the conditional mixing lemma by doing the same decomposition of Λ_n as before. The inequality (VIII.4) and the estimate (VIII.5) of the first summand are clearly still valid.

To estimate the remaining summands we need the following

Lemma VIII.9. — *Let $\lambda_n = \Gamma(n)\, m^\nu$ and fix some $r \in \{0,1\}$. Then, for any $\varepsilon \in (0,1)$, there exist some positive constant $\alpha > 0$ and some $n_0 \in \mathbb{N}$ such that*

$$\mathbf{P}\left(\frac{N_j^r}{\lambda_n} < 1-\varepsilon\right) \leqslant \mathrm{e}^{-\alpha\, \varkappa(n)\, n^\nu},$$

uniformly on $n > n_0$, $j = 1, \ldots, (2\,k)^\nu$ and $\overline{\boldsymbol{x}}_{\Lambda_k^*} \in \mathscr{X}^{\Lambda_k^*}$.

Proof: For definiteness let us take $r = 0$. We denote by V_{ij} a cube with side k centred at (i,j), $i = 1, \ldots, m^\nu$, $j = 1, \ldots, (2\,k)^\nu$, and let $\mathcal{V}_j = \mathbb{Z}^\nu \setminus (V_{1j} \cup \cdots \cup V_{m^\nu j})$. Note that Y_{ij}^0 depends only on the restriction of our periodized observation $\boldsymbol{x}(n)$ on the set V_{ij} and that for $i_1 \neq i_2$ we have $\rho(V_{i_1 j}\,, V_{i_2 j}) \geqslant 2\,k - k = k$. So, for any $\lambda > 0$, it follows from the conditional mixing lemma that

$$\mathbf{E}\Big(\mathrm{e}^{-\lambda\, N_j^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) = (1+\delta_n)^{m^\nu} \prod_{i=1}^{m^\nu} \mathbf{E}\Big(\mathrm{e}^{-\lambda\, Y_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) \tag{VIII.12}$$

with $\delta_n = O\big(k^\nu\, \varphi(k)\big) = O\Big(\gamma(n)\, c\, \mathrm{e}^{-a\, k^{\nu+\delta}}\Big) = o\big(n^{-\beta}\big)$ for all $\beta > 0$.

Clearly, using the Lemma VII.2, definition of Y_{ij}^0 and total probability formula, we have

$$\mathbf{E}\big(Y_{ij}^0 \;\big|\; \boldsymbol{x}_{\mathcal{V}_j}\big) \geqslant \mathrm{e}^{-b^\star\, |\Lambda_k|} \geqslant \mathrm{e}^{-b^\star\, \gamma(n)} = \Gamma(n).$$

Furthermore, using Taylor expansion formula, we get

$$\begin{aligned} \mathbf{E}\Big(\mathrm{e}^{-\lambda\, Y_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) &= \mathrm{e}^{-\lambda\, \mathbf{E}\left(Y_{ij}^0 \,|\, \boldsymbol{x}_{\mathcal{V}_j}\right)}\, \mathbf{E}\left(\mathrm{e}^{-\lambda\, \left(Y_{ij}^0 - \mathbf{E}\left(Y_{ij}^0 \,|\, \boldsymbol{x}_{\mathcal{V}_j}\right)\right)} \;\middle|\; \boldsymbol{x}_{\mathcal{V}_j}\right) \leqslant \\ &\leqslant \mathrm{e}^{-\lambda\, \Gamma(n)} \left(1 + \frac{\lambda^2}{2}\, \mathrm{e}^{\lambda}\right) \leqslant \exp\left(-\lambda\Big(\Gamma(n) - \frac{\lambda}{2}\, \mathrm{e}^{\lambda}\Big)\right). \end{aligned} \tag{VIII.13}$$

Finally, combining (VIII.12), (VIII.13), and using Chebychev's inequality and

total probability formula, for sufficiently large values of n we get

$$\begin{aligned}\mathbf{P}\left(\frac{N_j^0}{\lambda_n} < 1-\varepsilon\right) &\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\lambda_n}\ \mathbf{E}\,\mathrm{e}^{-\lambda\,N_j^0} \leqslant \\ &\leqslant \mathrm{e}^{\lambda\,(1-\varepsilon)\,\Gamma(n)\,m^\nu}\ \exp\left(-\lambda\left(\Gamma(n)-\frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right)m^\nu\right)(1+\delta_n)^{m^\nu} \leqslant \\ &\leqslant C\,\exp\left(-\lambda\,m^\nu\left(\varepsilon\,\Gamma(n)-\frac{\lambda}{2}\,\mathrm{e}^{\lambda}\right)\right).\end{aligned}$$

Now, choosing $\lambda = \varepsilon\,\Gamma(n)/\mathrm{e} < 1$, for sufficiently large values of n we get

$$\begin{aligned}\mathbf{P}\left(\frac{N_j^0}{\lambda_n} < 1-\varepsilon\right) &\leqslant C\,\exp\left(-\frac{\varepsilon\,\Gamma(n)}{\mathrm{e}}\,\frac{n^\nu}{2^\nu\,\gamma(n)}\left(\varepsilon\,\Gamma(n)-\frac{\varepsilon\,\Gamma(n)}{2}\right)\right) \leqslant \\ &\leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}\end{aligned}$$

with an arbitrary $\alpha < \dfrac{\varepsilon^2}{2^{\nu+1}\,\mathrm{e}}$. The lemma is proved. □

Using this lemma we clearly get

$$\mathbf{P}\big(N_j^r = 0\big) \leqslant \mathbf{P}\left(\frac{N_j^r}{\lambda_n} < 1-\varepsilon\right) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}$$

for all $j = 1,\ldots,(2\,k)^\nu$, $r \in \{0{,}1\}$ and for sufficiently large values of n. Therefore, for sufficiently large values of n, we have

$$\begin{aligned}\mathbf{P}\Big(\big\|D_n^2(\cdot)\big\| > \varepsilon/5\Big) &= \mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^2(\overline{\boldsymbol{x}})\big| > \varepsilon/5\Big) \leqslant \\ &\leqslant \sum_{\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}} \mathbf{P}\big(N^0 = 0\big) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}\end{aligned} \tag{VIII.14}$$

where we take into account that N^0 depends only on $\overline{\boldsymbol{x}}_{\Lambda_k^*}$, and hence the supremum over $\overline{\boldsymbol{x}} \subset \mathbb{Z}^\nu \setminus \mathbf{0}$ is in fact a maximum over $\overline{\boldsymbol{x}}_{\Lambda_k^*} \in \mathscr{X}^{\Lambda_k^*}$, *i.e.*, a maximum over $2^{|\Lambda_k^*|} \leqslant 2^{\gamma(n)}$ elements.

In exactly the same way we have

$$\mathbf{P}\Big(\big\|D_n^3(\cdot)\big\| > \varepsilon/5\Big) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}, \tag{VIII.15}$$

and similarly we get

$$\mathbf{P}\Big(\big\|D_n^4(\cdot)\big\| > \varepsilon/5\Big) = \mathbf{P}\Big(\sup_{\overline{\boldsymbol{x}}\subset\mathbb{Z}^\nu\setminus\mathbf{0}}\big|D_n^4(\overline{\boldsymbol{x}})\big| > \varepsilon/5\Big) \leqslant$$

$$\leqslant \sum_{\overline{\boldsymbol{x}}_{\Lambda_k^*}\in\mathscr{X}^{\Lambda_k^*}} \sum_{j=1}^{(2\,k)^\nu} \mathbf{P}\big(N_j^0 = 0\big) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}. \tag{VIII.16}$$

Finally, the last summand is estimated by the following lemma.

Lemma VIII.10. — *For any $\varepsilon \in (0,1)$ there exist some positive constant $\alpha > 0$ and some $n_0 \in \mathbb{N}$ such that*

$$\mathbf{P}\Big(\big\|D_n^5(\cdot)\big\| > \varepsilon/5\Big) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu} \tag{VIII.17}$$

for all $n > n_0$

Proof: As before, it is sufficient to show that

$$\mathbf{P}\left(N_j^0 > 0, \frac{1}{N^0}\left|N_j^1 - N_j^0\,h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| > \frac{\varepsilon}{5\,(2\,k)^\nu}\right) \leqslant \mathrm{e}^{-\alpha\,\varkappa(n)\,n^\nu}.$$

We have obviously

$$\mathbf{P}\left(N_j^0 > 0, \frac{1}{N^0}\left|N_j^1 - N_j^0\,h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\right| > \frac{\varepsilon}{5\,(2\,k)^\nu}\right) \leqslant$$

$$\leqslant \mathbf{P}\left(\left|\sum_{i=1}^{m^\nu}\Big(Y_{ij}^1 - Y_{ij}^0\,h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}\Big)\right| > \frac{\varepsilon\,N^0}{5\,(2\,k)^\nu}\right) \leqslant$$

$$\leqslant \mathbf{P}\left(\sum_{j=1}^{(2\,k)^\nu}\frac{N_j^0}{\lambda_n} \leqslant (1-\varepsilon)\,(2\,k)^\nu\right) + \mathbf{P}\left(\left|\sum_{i=1}^{m^\nu} W_{ij}\right| \geqslant \tau\,\lambda_n\right)$$

where $\tau = \varepsilon\,(1-\varepsilon)/5$ and $W_{ij} = Y_{ij}^1 - Y_{ij}^0\,h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}$. The estimate of the first term easily follows from the preceding lemma. To estimate the second one let us at first note that using translation invariance, total probability formula and the

formulas (I.5), (VII.1) and (VIII.3) we have

$$\begin{aligned}
\mathbf{E}\Big(Y_{ij}^0 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} &= \mathbf{P}_{\Lambda_k\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} = \\
&= \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(0)\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times \\
&\qquad \times \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(1)\Big/\, \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(0) = \\
&= (1+\rho_n)\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\; \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}}(1) \\
&= (1+\rho_n)^2\; \mathbf{P}_{\Lambda_k*\left|\mathcal{V}_j-(i,j)\right.}\Big(\overline{\boldsymbol{x}}_{\Lambda_k^*} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}-(i,j)\Big)\times \\
&\qquad \times \mathbf{Q_0}^{\overline{\boldsymbol{x}}_{\Lambda_k^*}\cup\left(\boldsymbol{x}_{\mathcal{V}_j}-(i,j)\right)}(1) = \\
&= \mathbf{E}\Big(Y_{ij}^1 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\,(1+\rho_n)
\end{aligned}$$

where $\rho_n = O\big(\varphi(k)\big) = O\Big(c\,\mathrm{e}^{-a\,k^{\nu+\delta}}\Big) = o\big(n^{-\beta}\big)$ for all $\beta > 0$.

This implies that

$$\mathbf{E}\Big(W_{ij} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) = \mathbf{E}\Big(Y_{ij}^1 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) - \mathbf{E}\Big(Y_{ij}^0 \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big)\, h^{\overline{\boldsymbol{x}}_{\Lambda_k^*}} = O(\rho_n)$$

and hence, for any $\lambda > 0$, using the fact that $-B \leqslant W_{ij} \leqslant 1$ and Taylor expansion formula, we get

$$\begin{aligned}
\mathbf{E}\Big(\mathrm{e}^{\lambda\, W_{ij}^0} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) &= \mathrm{e}^{\lambda\; \mathbf{E}\left(W_{ij}^0\middle|\boldsymbol{x}_{\mathcal{V}_j}\right)}\; \mathbf{E}\Big(\mathrm{e}^{\lambda\left(W_{ij}^0-\mathbf{E}\left(W_{ij}^0\middle|\boldsymbol{x}_{\mathcal{V}_j}\right)\right)} \;\Big|\; \boldsymbol{x}_{\mathcal{V}_j}\Big) \leqslant \\
&\leqslant \mathrm{e}^{\lambda\, O(\rho_n)}\left(1+\frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right) \leqslant \\
&\leqslant \exp\left(\lambda\, O(\rho_n) + \frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right).
\end{aligned}$$

Finally, using Chebychev's inequality, total probability formula and conditional mixing lemma, we get

$$\begin{aligned}
\mathbf{P}\left(\sum_{i=1}^{m^\nu} W_{ij} \geqslant \tau\,\lambda_n\right) &\leqslant \mathrm{e}^{-\lambda\,\tau\,\lambda_n}\; \mathbf{E}\exp\left(\lambda\sum_{i=1}^{m^\nu} W_{ij}\right) \leqslant \\
&\leqslant \mathrm{e}^{-\lambda\,\tau\,\Gamma(n)\,m^\nu}\; \mathbf{E}\left(\prod_{i=1}^{m^\nu}\mathbf{E}\Big(\lambda\,\mathrm{e}^{W_{ij}} \;\Big|\; \xi_{\mathcal{V}_j}\Big)\right)(1+\delta_n)^{m^\nu} \leqslant
\end{aligned}$$

$$\leqslant C\,\mathrm{e}^{-\lambda\,\tau\,\Gamma(n)\,m^{\nu}}\left(\exp\left(\lambda\,O(\rho_n)+\frac{\lambda^2\,(B+1)^2}{2}\,\mathrm{e}^{\lambda\,(B+1)}\right)\right)^{m^{\nu}}\leqslant$$

$$\leqslant C\,\exp\left(-\lambda\,m^{\nu}\left(\tau\,\Gamma(n)-\frac{(B+1)^2}{2}\,\lambda\,\mathrm{e}^{\lambda\,(B+1)}-O(\rho_n)\right)\right).$$

Now, choosing $\lambda=\dfrac{\tau\,\Gamma(n)}{(B+1)^2\,\mathrm{e}^{B+1}}<1$, we get

$$\mathbf{P}\left(\sum_{i=1}^{m^{\nu}}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant C\,\exp\left(-\frac{\tau\,\Gamma(n)}{(B+1)^2\,\mathrm{e}^{B+1}}\,\frac{n^{\nu}}{2^{\nu}\,\gamma(n)}\,\frac{\tau\,\Gamma(n)}{2}\right)\leqslant$$

$$\leqslant\mathrm{e}^{-\alpha\,\varkappa(n)\,n^{\nu}}$$

with an arbitrary $\alpha<\dfrac{\tau^2}{2^{\nu+1}\,(B+1)^2\,\mathrm{e}^{B+1}}$.

By the same argument we have

$$\mathbf{P}\left(-\sum_{i=1}^{m^{\nu}}W_{ij}\geqslant\tau\,\lambda_n\right)\leqslant\mathrm{e}^{-\alpha\,\varkappa(n)\,n^{\nu}}$$

which concludes the proof of the lemma. □

Now, combining (VIII.5), (VIII.14), (VIII.15), (VIII.16), (VIII.17) and taking into account the inequality (VIII.4), we get the assertion of the theorem. The uniformity with respect to $\mathbf{P}\in\mathscr{G}(\boldsymbol{h})$ and $h\in\widetilde{\mathscr{H}}$ is trivial. The Theorem VIII.8 is proved. □

Let us note, that from the details of the proof it clearly follows some explicit expression for the constant α. For example, if $\varepsilon\in(0{,}1)$, then one can take an arbitrary

$$\alpha<\frac{\tau^2}{2^{\nu+1}\,(B+1)^2\,\mathrm{e}^{B+1}}.$$

Note also, that taking a closer look on the proof we can give a "more precise" bound on the rate of consistency, showing explicitly the dependence of the rate on ε. That is, for $\varepsilon\in(0\,,1/2)$, we have the bound

$$\mathbb{1}_{\left\{6\,c\,n^{-a\,(\ln n)^{\delta/(2\,\nu+\delta)}}>\varepsilon\right\}}+\psi_n\,\exp\left\{-\alpha\,\varkappa(n)\,\varepsilon^2\,n^{\nu}+O(\rho_n)\,\beta\,\frac{\Gamma(n)}{\gamma(n)}\,\varepsilon\,n^{\nu}\right\}$$

where $\alpha = \dfrac{1}{25 \cdot 2^{\nu+3} (B+1)^2 \, \mathrm{e}^{B+1}}$, $\beta = \dfrac{1}{5 \cdot 2^{\nu+1} (B+1)^2 \, \mathrm{e}^{B+1}}$ and the sequence ψ_n is given by $\psi_n = 2^{\gamma(n)} \bigl(2^\nu \, \gamma(n) + 1\bigr) \bigl(2^\nu \, \gamma(n) + 2\bigr)$.

Using this last bound, as before, we easily obtain the following

Theorem VIII.11 [Uniform $\mathbf{L}^p$-consistency of the sieve estimator]**.** — *Assume that $\boldsymbol{h} \in \widetilde{\mathscr{H}}$, $\widehat{\boldsymbol{h}}_n$ is the sieve estimator with $k = \bigl[(\ln n)^{1/(\nu+\delta/2)}\bigr]$, and fix some $p \in (0,\infty)$. Then, for sufficiently large values of n, we have*

$$\sup_{\boldsymbol{h} \in \widetilde{\mathscr{H}}} \sup_{\mathbf{P} \in \mathscr{G}(\boldsymbol{h})} \Bigl(\mathbf{E} \, \bigl\| \widehat{\boldsymbol{h}}_n - \boldsymbol{h} \bigr\|^p \Bigr)^{1/p} \leqslant n^{-(\nu/2-\sigma)}$$

where σ is an arbitrary small positive constant, i.e., the estimator $\widehat{\boldsymbol{h}}_n$ is uniformly $\mathbf{L}^p$-consistent.

Let us finally note here, that only the constant δ is important in the definition of the sieve estimator. Hence we can apply the considerations of the Remark VII.6 by "releasing" the constants A, B, a and c, *i.e.*, by enlarging the class $\widetilde{\mathscr{H}}$ to the class $\widetilde{\mathscr{H}}^{\,\delta}$ defined by the conditions **(C4)**, **(C2′)** and **(C5)** with some *a priory* fixed constant δ. We will still obtain (no longer uniform) exponential and $\mathbf{L}^p$ consistencies of the sieve estimator. The problem with this approach is that the slowly varying function $\varkappa(n)$ present in the bounds on the rates of consistency will depend on $\boldsymbol{h}$ (by the way of $b^\star$). To avoid this, one can "release" only the constants a and c, *i.e.*, consider the class $\widetilde{\mathscr{H}}^{\,\delta}_{A,B}$ defined by the conditions **(C4)**, **(C2)** and **(C5)** with some *a priory* fixed constants A, B and δ. In this case we still obtain (no longer uniform) exponential and $\mathbf{L}^p$ consistencies of the sieve estimator.

Références

Références bibliographiques

[1] R. V. Ambartzumian, H. S. Sukiasian, "*Inclusion-exclusion and point processes*", Acta Appl. Math. **22**, pp. 15–31, 1991.

[2] M. B. Averintsev, "*A way of describing random fields with a discrete argument*", Probl. Inform. Transmiss. **6**, pp. 169–175, 1970.

[3] F. Comets, "*On consistency of a class of estimators for exponential families of Markov random fields on the lattice*", Ann. Stat. **20**, pp. 455–468, 1992.

[4] S. Dachian, "*About one method of description of random fields*", degree thesis, Department of Mathematics, Yerevan State University, 1995.

[5] S. Dachian, "*Nonparametric estimation for Gibbs random fields specified through one-point systems*", Université du Maine, preprint no. 98–6, Mai 1998, submitted for publication.

[6] S. Dachian, B. S. Nahapetian, "*Inclusion-exclusion description of random fields*", J. Contemp. Math. Anal. Armen. Acad. Sci. **30**, no. 6, pp. 50–61, 1995.

[7] S. Dachian, B. S. Nahapetian, "*An approach towards description of random fields*", Université du Maine, preprint no. 98–5, Mai 1998, submitted for publication.

[8] R. L. Dobrushin, "*The description of a random field by means of conditional probabilities and conditions of its regularity*", Theory of Probability and Applications **13**, pp. 197–224, 1968.

[9] R. L. Dobrushin, "*Gibbsian random fields for lattice systems with pairwise interaction*", Funct. Anal. Appl. **2**, pp. 292–301, 1968.

[10] R. L. Dobrushin, "*The problem of uniqueness of a Gibbs random field and the problem of phase transitions*", Funct. Anal. Appl. **2**, pp. 302–312, 1968.

[11] S. Geman, C. Graffigne, "*Markov random fields image models and their applications to computer vision*", in "Proc. of the Intern. Cong. of Math. 1986", ed.: A. M. Gleason, American Math. Society, NY, 1987.

[12] H.-O. Georgii, "Gibbs measures and phase transitions", *de Gruyter, Berlin*, 1988.

[13] U. Grenander, "Abstract inference", *Wiley, New York*, 1981.

[14] X. Guyon, "Random fields on a network. Modelling, statistics and applications", *Springer-Verlag, New York*, 1995.

[15] C. Ji, "*Sieve estimators for pair-interaction potentials and local characteristics in Gibbs random fields*", University of North Carolina, preprint no. 2037, October 1990.

[**16**] R. B. Israel, "*Banach algebras and Kadanoff transformations*", Random Fields (Estergom 1979), North-Holland, **7**, pp. 593–608, 1981.

[**17**] O. K. Kozlov, "*Gibbsian description of a system of random quantities*", Probl. Inform. Transmiss. **10**, pp. 258–265, 1974.

[**18**] J. L. Lebowitz, C. Maes, "*The effect of an external field on an interface, entropic repulsion*", J. Stat. Phys. **46**, pp. 39–49, 1987.

[**19**] V. A. Malyshev, "*Cluster expansion in lattice models of statistical physics and the quantum theory of fields*", Russ. Math. Surv. **35**, no. 2, pp. 3–53, 1980.

[**20**] V. A. Malyshev, R. A. Minlos, "Gibbs random fields. Method of cluster expansions" (in Russian), *Nauka, Moscow*, 1985.

[**21**] A. Possolo, "*Estimation of binary Markov random fields*", Tech. report no. 77, Dept. of Stat., Univ. of Washington, 1986.

[**22**] A. Possolo, "*Subsampling a random field*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[**23**] R. H. Schonmann, "*Projections of Gibbs measures may be non-Gibbsian*", Commun. Math. Phys. **124**, pp. 1–7, 1989.

[**24**] W. G. Sullivan, "*Potentials for almost Markovian random fields*", Commun. Math. Phys. **33**, pp. 61–74, 1973.

[**25**] A. van Enter, R. Fernandez, A. Sokal, "*Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory*", J. Stat. Phys. **72**, no. 5–6, pp. 879–1152, 1993.

Bibliographie

[**26**] L. S. Andersen, "*Inference for hidden Markov models*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[**27**] R. Azencott, "*Image analysis and Markov fields*", in "Proc. of the 1-st Intern. Conf. on App. Math., Paris", SIAM Philad., eds.: J. Mckenna and R. Temen, pp. 53–61, 1987.

[**28**] M. Almeida, B. Gidas, "*A variational method for estimating the parameters of Markov random field from complete or incomplete data* ", Ann. Appl. Prob. **3**, no. 1, pp. 103–136, 1993.

[**29**] J. Besag, "*Spatial interaction and the statistical analysis of lattice systems (with discussions)*", J. Roy. Stat. Soc. **36B**, pp. 192–236, 1974.

[**30**] J. Besag, "*Spatial analysis of non lattice data*", The Stat. **24**, pp. 179–195, 1975.

[**31**] J. Besag, "*Efficiency of pseudo-likelihood estimation for simple Gaussian field*", Biometrika **64**, pp. 616–618, 1977.

[**32**] J. Besag, "*Some methods of statistical analysis for spatial data*", Bull. Int. Statist. Inst. **47**, no. 2, pp. 77–92, 1978.

[33] J. Besag, "*On the statistical analysis of dirty pictures (with discussions)*", J. Roy. Stat. Soc. **48B**, pp. 259–302, 1986.

[34] J. Besag, P. A. P. Moran, "*On the estimation and testing of spatial interaction in Gaussian lattice processes*", Biometrika **62**, pp. 555–562, 1975.

[35] J. Besag, J. York, A. Mollie, "*Bayesian image restoration, with two applications in spatial statistics (with discussions)*", Ann. Inst. Stat. Math. **43**, pp. 1–59, 1991.

[36] B. Chalmond, "*Image restoration using an estimated Markov model*", Signal Processing **15**, 1988.

[37] R. Chellappa, R. L. Kashyap, "*Digital image restoration using spatial interaction models*", IEEE Trans. on Accoust., Speech., Signal Proc., pp. 461–472, 1982.

[38] M. Clyde, D. Strauss, "*Logistic regression for spatial pair potential models*", in "Spat. Stat. And Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[39] F. S. Cohen, D. B. Cooper, "*Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markovian random fields*", IEEE Trans. PAMI **9**, pp. 195–219, 1987.

[40] F. Comets, "*Detecting phase transitions for Markov random fields*", in "Trans. of the 12-th Prague Conference", 1994.

[41] F. Comets, "*Detecting phase transition for Gibbs measures*", Ann. Applied Prob. **7**, no. 2, pp. 545–563, 1997.

[42] F. Comets, B. Gidas, "*Asymptotics of maximum likelihood estimator for the Curie-Weiss model*", Ann. Stat.**19**, pp. 557–578, 1991.

[43] F. Comets, B. Gidas, "*Parameter estimation for Gibbs distributions from partially observed data*", Ann. Applied Prob. **2**, pp. 142–170, 1992.

[44] G. R. Cross, A. K. Jain, "*Markov random fields texture models*", IEEE Trans. PAMI **5**, pp. 25–40, 1983.

[45] J. Demongeot, "*Asymptotic inference for Markov random fields on* $\mathbb{Z}^2$", in "Numerical Met. of Study of Crit. Phen., Proc. Symp. Carry-le-Rouet, France", Springer Series in Synergetics **9**, eds.: J. Della, et. al., 1981.

[46] H. Derin, H. Elliott, "*Modeling and segmentation of noisy and textured images using Gibbs random fields*", IEEE Trans. PAMI **9**, pp. 39–55, 1987.

[47] H. Derin, H. Elliott, R. Cristi, D. Geman, "*Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields*", IEEE Trans. PAMI **6**, pp. 707–720, 1984.

[48] H. Derin, C. S. Won, "*Estimating the parameters of a class of Gibbs distributions*", Conf. on Inform. Sci. And Syst., pp. 222–228, 1988.

[49] P. J. Diggle, "*On parameter estimator and goodnes-on-fit testing for spatial patterns*", Biometrics **35**, pp. 87–101, 1979.

[50] P. J. Diggle, "*Statistical analysis of spatial point patterns*", Acad. Press, London, 1983.

[51] P. J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, M. Tanemura, "*On parameter estimation for pairwise interaction point processes*", Intern. Stat. Rev. **62**, pp. 99–117, 1994.

[52] J. M. Dinten, X. Guyon, J. F. Yao, "*On the choice of the regularization parameter: the case of binary images in the Bayesian restoration framework*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[53] R. J. Elliot, L. Aggoun, "*MAP estimation using measure change for continuous-state random fields*", Syst. Cont. Lett. **26**, no. 4, pp. 239–244, 1995.

[54] T. Fiksel, "*Estimation of parametrized pair potentials of marked and non-marked Gibbsian point processes*", Elektron. Inform. Kybernet. **20**, pp. 270–278, 1984.

[55] T. Fiksel, "*Estimation of interaction potentials of Gibbsian point processes*", Math. Operationsforsch. Stat. Ser. Stat. **19**, pp. 77–86, 1988.

[56] A. Frigessi, M. Piccioni, "*Parameter estimation for two-dimensional Ising fields corrupted by noise*", Stoch. Proc. & App. **34**, pp. 297–311, 1990.

[57] A. Frigessi, M. Piccioni, "*Consistent parameter estimation for 2–D Ising fields corrupted by noise: numerical experiments*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[58] D. Geman, "*Random fields and inverse problems in imaging*", Lect. Notes in Math. **1472**, Springer-Verlag, 1991.

[59] D. Geman, S. Geman, "*Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*", IEEE Trans. PAMI **6**, pp. 721–741, 1984.

[60] D. Geman, S. Geman, "*Bayesian image analysis*", in "Disordered Systems and Biological Organization", Springer-Verlag, eds.: E. Bienenstock, et. al., 1986.

[61] D. Geman, S. Geman, C. Graffigne, "*Locating texture and object boundaries*", in "Pattern Recognition Theory and Applications", NATO ASI Series, Springer-Verlag, eds.: P. Devijver, et. al., 1986.

[62] D. Geman, S. Geman, C. Graffigne, P. Dong, "*Boundary detection by constraint optimization*", IEEE Trans. PAMI **12**, pp. 609–628, 1990.

[63] S. Geman, D. E. McClure, "*Bayesian image analysis: an application to single photon emission tomography*", in "Proc. of the American Statistical Association, Statistical Computing Section", pp. 12–18, 1985.

[64] B. Gidas, "*Consistency of maximum likelihood and maximum pseudo-likelihood estimators for Gibbs distributions*", in "Proc. of the Workshop on Stoch. Diff. Syst. with App. in Elec. Comp. Eng.", IMA, University of Minnesota, 1986, Springer-Verlag, 1987.

[65] B. Gidas, "*A renormalization group approach to image processing problems*", IEEE Trans. PAMI **11**, pp. 164–180, 1989.

[66] B. Gidas, "*Parameter estimation for Gibbs distributions from fully observed data*", in "Markov Random Fields, Theory and Applications", Acad. Press, eds.: R. Chellappa and A. Jain, 1993.

[67] E. Glötzl, B. Rauschenschwandtner, "*On the statistics of Gibbsian processes*", in "The 1-st Pannonian Symp. on Math. Stat., Bad Tatzmansdorf, Austria, 1979", Lect. Notes in Stat. **8**, eds.: P. Revesz, et. al., pp. 83–93, 1981.

[68] F. Godtliebsen, C. K. Chu, "*Estimation of the number of true gray levels, their values and relative frequencies in a noisy image*", J. Amer. Statist. Assoc. **90**, pp. 890–899, 1995.

[69] D. M. Greig, B. T. Porteous, A. H. Seheult, "*Exact MAP estimation for binary images*", J. Roy. Stat. Soc. **51B**, pp. 271–279, 1987.

[70] X. Guyon, "*Estimation d'un champ par pseudo-vraisemblence conditionnelle : étude asymptotique et application en cas Markovien*", in "Actes de la 6-ème Rencontre Franco-Belge des Stat., Brux.", also in "Spat. Proc. and Spac. Time Series Analysis", ed.: F. Droesbeke, 1987.

[71] X. Guyon, "*Méthodes de pseudo-vraisemblence et de codage pour le processus ponctuels de Gibbs*", prébub. de SAMOS (Univ. Paris 1), no. 6, 1991.

[72] X. Guyon, H. R. Künsch, "*Asymptotic comparison of estimators in the Ising model*", in "Stoch. Models, Stat. Met. And Algorithms in Image Analysis", Lecture Notes in Statistics **74**, Springer-Verlag, eds.: P. Barone, A. Frigessi and M. Piccioni, 1992.

[73] M. Janžura, "*Estimating interactions in binary lattice data with nearest-neighbour property*", Kybernetika **23**, no. 2, pp. 136–142, 1987.

[74] M. Janžura, "*Statistical analysis of Gibbs random fields*", in "Trans. of the 10-th Prague Conf. on Inf. Theory, Stat. Dec. Fct., Rand. Proc. 1986", 1988.

[75] M. Janžura, "*Asymptotic theory of parameter estimation for Gauss-Markov random fields*", Kybernetika **24**, no. 3, pp. 161–176, 1988.

[76] M. Janžura, "*Divergences of Gauss-Markov random fields with applications to statistical inference*", Kybernetika **24**, no. 6, pp. 401–412, 1988.

[77] M. Janžura, "*LAN for Gibbs random fields*", in "Proc. of the 4-th Prague Symp. on Asympt. Stat.", eds.: P. Mandl and M. Huskova, 1989.

[78] M. Janžura, "*Asymptotic normality of the maximum pseudo-likelihood estimate of parameters for Gibbs random fields*", in "Trans. of the 12-th Prague Conference", 1994.

[79] J. L. Jensen, "*Asymptotic normality of estimates in spatial point processes*", Scand. J. of Stat. **20**, pp. 97–109, 1993.

[80] J. L. Jensen, H. R. Künsch, "*On asymptotic normality of pseudo-likelihood estimates for pairwise interaction processes*", Ann. Inst. Stat. Math. **46**, no. 3, pp. 475–486, 1994.

[81] J. L. Jensen, J. Møller, "*Pseudo-likelihood for exponential family models of spatial processes*", Ann. Appl. Prob. **1**, pp. 445–461, 1991.

[82] C. Ji, "*Some estimation problems for Gibbs states*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.

[**83**] R. L. Kashyap, R. Chellappa, "*Estimation and choice of neighbours in spatial interaction models of images*", IEEE Transactions on Information Theory **29**, pp. 60–72, 1983.

[**84**] H. R. Künsch, "*Thermodynamics and statistical analysis of Gaussian random fields*", Z. Wahrsch. Verw. Gibiete **58**, pp. 407–422, 1981.

[**85**] S. Mase, "*LAN of Gibbs models on a lattice*", Advances in Applied Probability **16**, pp. 585–602, 1984.

[**86**] S. Mase, "*Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions*", Prob. Theory and Rel. Fields **92**, pp. 51–67, 1992.

[**87**] S. Mase, "*Consistency of the maximum pseudo-likelihood estimator of continuos state space Gibbsian processes*", Ann. App. Prob. **5**, no. 3, pp. 603–612, 1995.

[**88**] S. Mase, Y. Ogata, M. Tanemura, "*Statistical analysis of mapped point patterns — present condition of theory and applications*", Amer. Math. Soc. Transl. **161**, pp. 95–108, 1994.

[**89**] U. V. Naik-Nimbalkar, M. B. Rajarshi, "*Towards semi-parametric image processing*", Indian J. of Pure and Appl. Math. **25**, pp. 239–245, 1994.

[**90**] Y. Ogata, M. Tanemura, "*Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure*", Ann. Inst. Stat. Math. **33**, pp. 315–338, 1981.

[**91**] Y. Ogata, M. Tanemura, "*Likelihood analysis of spatial point patterns*", J. Roy. Stat. Soc. **46B**, pp. 496–518, 1984.

[**92**] Y. Ogata, M. Tanemura, "*Estimation of interaction potentials of marked spatial point process through the maximum likelihood method*", Biometrics **41**, pp. 421–433, 1985.

[**93**] Y. Ogata, M. Tanemura, "*Likelihood estimation of interaction potentials and external fields of inhomogenius spatial point patterns*", in "Proc. Pacif. Stat. Cong.", eds.: I. S. Francis, B. F. G. Manly and F. C. Lam, 1986.

[**94**] Y. Ogata, M. Tanemura, "*Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns*", Ann. Inst. Stat. Math. **41**, pp. 583–600, 1989.

[**95**] A. Penttinen, "*Modeling interactions in spatial point patterns: parameter estimation by the maximum likelihood method*", Jyväskylä Studies in Comp. Sci., Econ. And Stat. **7**, 1984.

[**96**] M. Piccioni, A. Ramponi, "*Best unilateral approximation for the partition function of the two-dimensional Ising model*", in "Trans. of the 12-th Prague Conference", 1994.

[**97**] D. K. Pickard, "*Asymptotic inference for an Ising lattice*", J. Appl. Prob. **13**, pp. 486–497, 1976.

[**98**] D. K. Pickard, "*Asymptotic inference for an Ising lattice II*", Adv. Appl. Prob. **9**, pp. 476–501, 1977.

[99] D. K. Pickard, "*Asymptotic inference for an Ising lattice III. Non-zero field and ferromagnetic states*", J. Appl. Prob. **16**, pp. 12–24, 1979.

[100] D. K. Pickard, "*Asymptotic inference for an Ising lattice IV: Besag's coding method*", J. Appl. Prob. **16**, pp. 12–24, 1981.

[101] D. K. Pickard, "*Inference for general Ising models*", J. Appl. Prob. **19A**, pp. 345–357, 1982.

[102] D. K. Pickard, "*Inference for discrete Markov fields: the simplest nontrivial case*", J. Amer. Statist. Assoc. **82**, pp. 90–96, 1987.

[103] B. D. Ripley, "*Modelling spatial patterns (with discussions)*", J. Roy. Stat. Soc. **39B**, pp. 172–212, 1977.

[104] B. D. Ripley, "*Statistics, images and pattern recognition*", Canad. J. of Stat. **14**, pp. 83–111, 1986.

[105] A. Särkka, "*Pseudo-likelihood approach for Gibbs point processes in connection with field observation*", Statistics **26**, no. 1, pp. 89–97, 1995.

[106] D. J. Strauss, "*Analyzing binary lattice data with the nearest-neighbour property*", J. Appl. Prob. **12**, pp. 702–712, 1975.

[107] R. Takacs, "*Estimator for the pair potential of a Gibbsian point process*", Statistics **17**, pp. 429–433, 1986.

[108] L. Younes, "*Couplage de l'estimation et du recuit pour des champs de Gibbs*", CRAS de Paris **303I**, no. 13, 1986.

[109] L. Younes, "*Estimation and annealing for Gibbsian fields*", Annales de l'Institute Henri Poincare **24**, pp. 269–294, 1988.

[110] L. Younes, "*Problèmes d'estimation parametrique pour des champs de Gibbs Markoviens. Application en traitement d'image*", Thesis at Université d'Orsay, 1988.

[111] L. Younes, "*Parametric inference for imperfectly observed Gibbsian fields*", Prob. Theory and Rel. Fields **82**, pp. 625–645, 1989.

[112] L. Younes, "*Maximum likelihood estimation for Gibbsian fields*", in "Spat. Stat. and Imaging", IMS, L. N.-M. S., ed.: A. Possolo, 1991.