Combinatorial series expansions

for point processes

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
an der Ludwig-Maximilians-Universität München

eingereicht von
Leonid Kolesnikov

München, den 23.01.2023

1. Gutachter: Prof. Dr. Sabine Jansen
2. Gutachter: Prof. Dr. Daniel Ueltschi
3. Gutachter: Prof. Dr. Franz Merkl

Tag der mündlichen Prüfung: 17.05.2023

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.2011, § 8, Abs. 2, Pkt. 5.)

Hiermit erkläre ich an Eides statt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

München, den 23.01.2023

Leonid Kolesnikov

Zusammenfassung

Zufällige Konfigurationen abzählbar vieler Teilchen in messbaren Räumen wie $\mathbb{Z}^d$ und $\mathbb{R}^d$ werden üblicherweise durch Zufallsvariablen, die Werte in geeigneten Mengen von Zählmaßen annehmen, modelliert. Solche Zufallsvariablen werden als Punktprozesse bezeichnet. Falls die Teilchen, informell gesprochen, „gleichmäßig und voneinander unabhängig“ im Raum verteilt sein sollen, spricht man von einem Poisson-Punktprozess (PPP). Wichtige zugehörige Größen wie Teilchendichten können für den PPP direkt und explizit berechnet werden.

Betrachten wir allerdings Gibbs-Punktprozesse (GPP) — Verallgemeinerungen vom PPP, bei welchen die Annahme der Unabhängigkeit der Teilchen fallen gelassen wird, d.h., Wechselwirkungen zwischen den Teilchen erlaubt sind — so lassen sich Größen, die Informationen über diese Systeme von interagierenden Teilchen kodieren, im Allgemeinen nicht explizit berechnen. Eine Möglichkeit, dennoch Informationen über das System aus diesen schwer zu handhabenden Funktionen zu gewinnen, liefert ein perturbativer Ansatz: Man kann die GPP als Modifikationen eines zugrundeliegenden PPP bezüglich eines Wechselwirkungspotenzials betrachten und die fraglichen Funktionen als Potenzreihen in der Dichte dieses PPP, eines die Aktivität genannten Parameters, um die Null entwickeln. Unter der Voraussetzung, dass die Wechselwirkungen hinreichend schwach sind, erwartet man, dass diese Aktivitätsentwicklungen in einem Regime geringer Dichte sinnvoll sind. Der Ansatz ist als Clusterentwicklung bekannt (manchmal verwendet man die Bezeichnung auch restriktiver für eine partikuläre Entwicklung des Logarithmus der Zustandssumme — der freien Energie des Systems). Man interessiert sich natürlicherweise für die Konvergenzradien dieser Aktivitätsentwicklungen.

In „Cluster expansions: Necessary and sufficient conditions“ folgen wir der Tradition eines analytischen Zugangs zur Clusterentwicklung, welcher die sogenannten Kirkwood-Salsburg-Gleichungen hinzuzieht. Das sind Integral-Fixpunktgleichungen über einem geeigneten Banachraum von Funktionen, welche von den zum gegebenen Paar-Wechselwirkungspotential gehörenden Korrelationsfunktionen gelöst werden. Wir leiten eine abstrakte hinreichende Bedingung für die Konvergenz der entsprechenden Aktivitätsentwicklungen — die im Fall repulsiver Paarwechselwirkungen auch notwendig ist — in einem recht allgemeinen Setup her. Unser Kriterium formulieren wir hinsichtlich der Existenz von Funktionen, die Ungleichungen vom Kirkwood-Salsburg-Typ erfüllen. Wir zeigen nicht nur, wie man klassische hinreichende Bedingungen im Rahmen dieses Ansatzes vereinen kann, sondern auch, wie neue Konvergenzkriterien unter dessen Anwendung gefunden werden können, sowohl in diskreten als auch in stetigen Setups.

Die Aktivitätsentwicklungen der Korrelationsfunktionen können als exponentielle erzeugende Funktionen bestimmter Mengen von Graphen ausgedrückt werden. Grundsätzlich besteht die Hoffnung, Informationen über die graphische Entwicklung der Korrelationsfunktionen zu erlangen, indem man die zugrundeliegenden kombinatorischen Strukturen von einem konstruktiven Standpunkt aus versteht: Elementare konstruktive Beziehungen zwischen solchen „kombinatorischen Spezies“ lassen sich unkompliziert in Beziehungen zwischen den zugehörigen erzeugenden Funktionen übersetzen. In „Logarithms of Catalan generating functions: A combinatorial approach“ demonstrieren wir diesen Ansatz und präsentieren ein kombinatorisches Resultat, welches den Logarithmus der erzeugenden Funktion der (verallgemeinerten) Catalan-Zahlen hinsichtlich verschiedener baum- oder pfadähnlicher Strukturen interpretiert. Auf dem Niveau formaler Reihen liefert das Resultat ein Analogon zu der

klassischen Clusterentwicklung der freien Energie, wo die graphische Entwicklung der Zustandssumme (i.e., die exponentielle erzeugende Funktion der Graphen) durch die erzeugende Funktion der Catalan-Zahlen (i.e., die exponentielle erzeugende Funktion regulärer geordneter Bäume) als Ausgangspunkt ersetzt wird. Das erlaubt uns auch, bekannte Ausdrücke für die Koeffizienten des Logarithmus der erzeugenden Funktion der Catalan-Zahlen auf ein simples Aufzählungsproblem zurückzuführen.

Einige Fragen, die in der Theorie der GPP aufkommen — wie die Eindeutigkeit zu gegebenen Aktivitäten und Paarpotentialen gehörender Gibbsmaße — stehen im starken Zusammenhang mit Fragen aus verwandten Modellen für Perkolation. Ein konkretes Beispiel für ein auf dem PPP in $\mathbb{R}^d$ basierendes Modell für Perkolation ist das random connection model (RCM), in dem, informell gesprochen, ein Zufallsgraph dadurch gebildet wird, dass man Kanten zwischen zwei Punkten des PPP unabhängig voneinander mit einer Wahrscheinlichkeit, die von einer gegebenen connection function diktiert wird, zieht. Die Wahrscheinlichkeit dafür, dass zwei vorab in $\mathbb{R}^d$ fixierte Punkte durch (Pfade von) Kanten in diesem Graphen verbunden sind, wird durch die pair connectedness function beschrieben. Man kann die pair connectedness function als formale Potenzreihe in der Dichte z des zugrundeliegenden PPP entwickeln. Zwischen der Paarkorrelationsfunktion eines GPP und der pair connectedness function im RCM gibt es eine Entsprechung, die durch ein System von Integralgleichungen folmalisiert werden kann, allem voran unter Einbeziehung der sogenannten Ornstein-Zernike-Gleichung (OZE). Diese Gleichung setzt die Potenzreihenentwicklung der pair connectedness function, welche bezüglich bestimmter zusammenhängender Graphen ausgedrückt werden kann, in Beziehung zur sogenannten direct-connectedness function, von der man erwartet, dass eine graphische Entwicklung bezüglich bestimmter „doppelt zusammenhängender“ Graphen möglich ist. In „The direct-connectedness function in the random connection model“ leiten wir eine solche Entwicklung für die direct-connectedness function rigoros her, zeigen, dass sie tatsächlich die OZE in einem gewissen nicht-trivialen Konvergenzbereich der Intensität erfüllt, und setzen sie zu anderen bekannten Entwicklungen, wie der lace expansion, in Beziehung.

Abstract

Random configurations of countably many particles in measurable spaces such as $\mathbb{Z}^d$ and $\mathbb{R}^d$ are commonly modeled by random variables with values in a suitable set of counting measures. These random variables are called point processes. If the particles are supposed to be distributed, informally speaking, "uniformly and independently of each other" in space, the random variable is referred to as the Poisson point process (PPP). Important associated quantities such as particle densities can be computed directly and explicitly for the PPP.

However, if we consider Gibbs Point Processes (GPP) — generalizations of the PPP, where the assumption of independence between the particles is dropped, i.e., interactions between the particles are allowed — the quantities encoding the information about those systems of interacting particles become intractable in general. A way to still extract information about the system from those intractable functions is to use a perturbative approach: One can view the GPP as a modification of an underlying PPP in terms of an interaction potential and express the functions in question as power series in the density of that PPP, a parameter called the activity, around zero. Under the assumption that the interactions are sufficiently weak, one expects those activity expansions to be reasonable in a low-density regime. This approach is known as cluster expansion (sometimes the terminology is used more restrictive to describe a particular type of expansion for the logarithm of the partition function — the free energy of the system). Naturally, one is interested in the convergence radii of those activity expansions.

In "Cluster expansions: Necessary and sufficient conditions" we follow a tradition of analytic arguments involving the so-called Kirkwood-Salsburg equations. Those are integral fixed-point equations over suitable Banach spaces of functions that are solved by the correlation functions associated to the given pair-interaction potential. We are able to derive an abstract sufficient condition for the convergence of the corresponding activity expansions — that is also necessary in the case of repulsive interactions — in a quite general setup. Our criterion is formulated in terms of existence of functions satisfying Kirkwood-Salsburg-type inequalities. We show not only how the classical sufficient conditions can be unified under this approach, but also how it can be employed to find novel criteria for convergence, both in discrete and in continuous setups.

The activity expansions for the correlation functions can be expressed as exponential generating functions for certain sets of graphs. Generally speaking, one can hope to gain knowledge about such graphical expansions by understanding the underlying combinatorial structures from a constructive point of view: Basic constructive relations between these "combinatorial species" are easily translated to relations between the associated generating functions. In "Logarithms of Catalan generating functions: A combinatorial approach" we demonstrate this approach and present a new combinatorial result, interpreting the logarithm of the generating function for (generalized) Catalan numbers in terms of different tree-like or path-like structures. The result provides an analogue to the classical cluster expansion of the free energy on the level of formal power series — where the graphical expansion of the partition function (i.e., the exponential generating function for graphs) is replaced by the Catalan generating function (i.e., the exponential generating function for regular plane trees) as a starting point. It also allows us to recover known expressions for the coefficients of the logarithm of the Catalan generating functions via rather simple exact enumeration of those

combinatorial structures.

Some questions arising in the theory of GPP — such as the uniqueness of Gibbs measures corresponding to given activities and pair-interaction potentials — are strongly connected to questions arising in associated percolation models. A particular model driven by the PPP in $\mathbb{R}^d$ is the random connection model (RCM) where, informally speaking, a random graph is obtained by drawing edges between points of the PPP independently of each other with probabilities dictated by a given connection function. The probability for two points fixed in $\mathbb{R}^d$ to be connected through (paths of) edges of that graph is described by the pair connectedness function. One can expand the pair connectedness function as a power series in the density z of the underlying PPP. There is a correspondence between the pair correlation function of a GPP and the pair connectedness function in the RCM which can be formalized by systems of integral equations, most notably involving the so-called Ornstein-Zernike equations (OZE). This equation relates the power series expansion of the pair connectedness function given in terms of certain connected graphs to the so-called direct connectedness function which is expected to possess a graphical expansion in terms of certain "doubly connected" graphs. In "The direct-connectedness function in the random connection model" we rigorously derive such an expansion for the direct-connectedness function showing that it indeed does satisfy the OZE in a certain non-trivial domain of convergence and relate it to other known expansions like the lace expansion.

Acknowledgements

First and foremost, I would like to thank my supervisor, Sabine Jansen, for her truly inexhaustible patience with me. Her insight and her guidance have proven to be invaluable and it has always been a pleasure to work with her. I could not have wished for a better supervisor.

I also would like to express my gratitude towards Daniel Ueltschi and Franz Merkl for accepting to review this thesis.

A special thanks and shoutout to my co-author Kilian Matzke. Our collaboration was a great learning experience for me. I thoroughly enjoyed it.

For all the fun we had at the Institute and outside of it, I thank the people from the probability group and other colleagues, including but not limited to: Alejandro, Anna, Felix, Jakob, Leon, Matija, Matthew, Nannan, Philipp, Simon, Stefan, Tamás, Thomas and Umberto. I also thank Markus Heydenreich for his kindness and his (major) contribution to the great atmosphere at the Institute. Looking back, I am humbled and thankful for the opportunity to work in such an enjoyable, fruitful environment.

Finally, I would like to thank my family and friends for all the support I received over the years.

Contents

1 Introduction

1.1 Motivation and brief outline

Suppose our goal is to model a gas as a system of interacting particles in some physical space. The statistical mechanics approach to this task involves two main assumptions: The positions of the particles (gas molecules) are random and very large systems can be represented by infinite systems. That is, instead of considering microscopic dynamics on deterministic configurations of particles in a very large volume, one encodes the microscopic dynamics into probability distributions (ensembles) and studies a random configuration on an infinite volume — since the latter approach is more accessible. We consider the grand-canonical ensemble where neither the number of particles nor the energy of the system are deterministic. If there are no interactions between the particles, one refers to the system as an ideal gas.

Mathematically, the random configurations of particles are modeled by random variables with values in a suitable space of counting measures. Those random variables are referred to as point processes. An ideal gas is modeled via the Poisson point process (PPP), the particle density is tuned via the intensity parameter z. In the case of present interactions, the so-called Gibbs point processes (GPP) or (grand-canonical) Gibbs measures are used. GPP can be viewed as modifications of the PPP corresponding to a fixed interaction potential and an activity z (the latter is the intensity of the underlying PPP). Besides applications in statistical mechanics [9, 17], GPP appear in various models from stochastic geometry [5] and from spatial statistics [36]. We rigorously introduce the setup of GPP with pair-interactions and the necessary notation in Subsection 1.2.

Typically, in statistical mechanics, one is interested in the analyticity of certain functions (such as free energy or correlation functions) encoding information about the system of interacting particles. The functions of interest are, in general, intractable in the Gibbs setup (i.e., in the presence of interactions between particles), however, one would like to extract information from them by differentiating in some parameter, e.g., in the activity parameter z. Therefore, one expresses those quantities as power series in z around zero. This perturbative approach is referred to as cluster expansion. The study of those formal power series in regard to their structure and domains of convergence is our main goal; we focus primarily on cluster expansion results in this introduction.

The original approach to the classical cluster expansion of the free energy involves expressing it as an exponential generating function for connected graphs and trying to see what the structure of those graphs tells about the convergence of the generating function (typically, via tree-graph bounds [4] or, more recently, via tree-graph identities [13]). We discuss this combinatorial trick that initiated the cluster expansion approach (known as the Mayer trick [34]) and introduce the activity expansions ρ_n for the correlation functions (given as exponential generating functions for so-called multi-rooted graphs [25]) in Subsection 1.3.

In Subsection 1.4, we briefly discuss different approaches to cluster expansion known from the literature and introduce our results from “Cluster expansions: Necessary and sufficient conditions” (contribution [a]). Our main result (see [a, Theorem 2.1] and Theorem 1.5) is a novel sufficient condition for the convergence of the cluster expansion for the correlation functions in a general pair-interaction setup that is also necessary in the case of repulsive interactions. To obtain the result, we combine combinatorial considerations with analytical

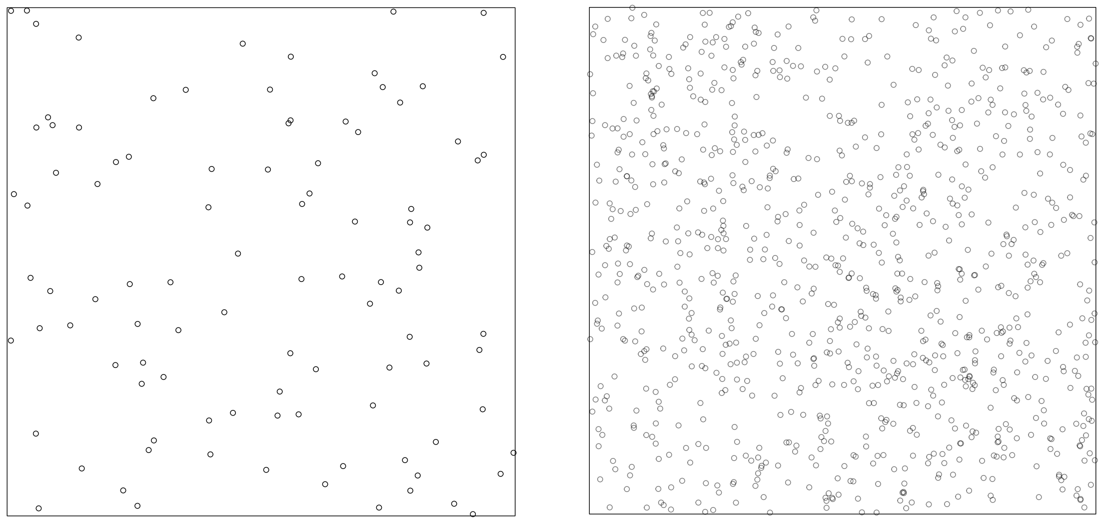

Figure 1: On the left, a sample of a PPP with the homogeneous intensity $z \equiv 1$ is depicted; on the right a sample of a PPP with $z \equiv 10$.

arguments à la Gruber and Kunz [18], involving Kirkwood-Salsburg equations — certain integral fixed-point equations which the sequence of correlation functions solves exactly and which can be formally expressed as

$$K_z \rho(z) + e_z = \rho(z),$$

where $\rho(z)$ denotes the activity expansions for the sequence of correlation functions and K_z denotes the Kirkwood-Salsburg operator (see [a,25]). Our characterization of the convergence domain of $\rho(z)$ is formulated in terms of existence of sequences of non-negative measurable solutions ξ of the corresponding sign-flipped Kirkwood-Salsburg inequalities

$$\tilde{K}_z \xi + e_z \leq \xi, \tag{1.1}$$

where $\tilde{K}_z$ denotes the sign-flipped Kirkwood-Salsburg operator (see (1.4) and compare to the Kirkwood-Salsburg operator from [a, 25]).

This approach reduces finding sufficient conditions for the convergence of the activity expansions $\rho(z)$ to finding functions ξ satisfying Kirkwood-Salsburg-type inequalities (1.1). In particular, the classical conditions like Kotecký-Preiss [28], Gruber-Kunz [18] or Fernández-Procacci [13] are easily recovered by employing it. Moreover, we are able to use this method to derive various novel sufficient conditions for different hard-core setups, both discrete and continuous, in [a].

In Subsection 1.5, we return to the sequence of formal power series $\rho(z)$, given by the exponential generating functions for so-called multi-rooted graphs. We justify the role of $\rho(z)$ as the starting point and main object of our studies in [a] by relating it to other expansions known from the literature and to finite-volume Gibbs correlation functions. To do so, we formulate and prove rigorously several results including Proposition 1.7, Proposition 1.9 and Corollary 1.10.

Notice that to show analyticity of the free energy or of the correlation functions other graphical expansions can be helpful (like the virial expansion [27, 34], where power series in the density of the GPP are obtained) and might, in general, yield larger domains of analyticity. An exemplary result that points out the limitations of cluster expansion in a classical setup is given in Subsection 1.6.

From a purely combinatorial point of view, the classical cluster expansion of the free energy involves taking the logarithm of the exponential generating function for graphs — which results in the exponential generating function for connected graphs. What if we consider the logarithm of the exponential generating function for regular plane trees instead? For which combinatorial species, in particular what kind of tree-like structures, is the resulting formal power series the generating function? These combinatorial questions arise in the context of statistical mechanics when studying discrete one-dimensional systems of non-overlapping rods (see [24, Chapter 5.2]). In Subsection 1.6 we summarize our results from "Logarithms of Catalan generating functions: A combinatorial approach" (contribution [b]) providing answers to these questions: The generating function for generalized Catalan numbers can be viewed as the exponential generating function for regular plane trees and the logarithm of this formal power series can be represented by the exponential generating function for, e.g., tree-like structures called cycle-rooted trees. We give several alternative combinatorial interpretations for the logarithm of these Catalan generating functions that allow us to determine the coefficients by solving rather simple counting problems. The results are purely combinatorial, we view the occurring generating functions as formal power series and do not concern ourselves with questions of convergence.

Finally, in Subsection 1.7, we summarize our results from "The direct-connectedness function in the random connection model" (contribution [c]). They are motivated by the following question: In the random connection model (RCM), a particular PPP-driven model of continuum percolation [19, 35], how can the study of the percolation phase transition provide answers to important question concerning the phase transition in terms of uniqueness of an associated Gibbs measure? We expand certain connectivity functions in the RCM in a fashion which is partially inspired by cluster expansion: Consider the pair connectedness function given by the probability that two points (fixed a priori in the space) are connected via paths of a random graph with the vertex set given by a PPP. Like the pair correlation function of a GPP, the pair connectedness function admits a graphical expansion in terms of connected graphs. In statistical mechanics, the Ornstein-Zernike equation (OZE) provides relations between correlation functions and direct correlation functions [29]. What is the analogue for the two-point direct correlation function in the RCM setup? We rigorously derive an expansion for the function satisfying the OZE relations with the pair-connectedness function. While the existence of such an analytic function in a low-density regime is known by [31], we provide an explicit graphical expansion and quantify its domain of convergence.

1.2 General setup and notation

Let $(\mathbb{X}, \mathcal{X}, \lambda)$ be a measure space, where $\mathcal{X}$ is the Borel-σ-algebra on the metric space $\mathbb{X}$ and λ is a σ-finite Borel measure on $\mathcal{X}$. Let $\mathcal{X}_b \subset \mathcal{X}$ be the set of bounded Borel sets.

An activity function z is defined as a measurable map $z : \mathbb{X} \to \mathbb{R}_0^+$. Here we assume additionally that $\int_B z(x)\lambda(\mathrm{d}x) < \infty$ for all $B \in \mathcal{X}_b$. To an activity function, we assign the measure λ_z, given by

$$\lambda_z(B) := \int_B z(x)\lambda(\mathrm{d}x), \quad B \in \mathcal{X}.$$

Although one could introduce complex activities here and work in the usual complex analysis setup (in which case λ_z would be a complex measure in general), we only consider physically relevant positive activities in this introduction.

We call a measure η on $(\mathbb{X}, \mathcal{X})$ a locally finite counting measure if $\eta(B) \in \mathbb{N}_0$ for all $B \in \mathcal{X}_b$. We denote the set of locally finite counting measures on $(\mathbb{X}, \mathcal{X})$ by $\mathcal{N}_f$. Each non-zero $\eta \in \mathcal{N}_f$ can be written as $\eta = \sum_{i=1}^{\kappa} \delta_{x_i}$ for $\kappa \in \mathbb{N} \cup \{\infty\}$ and $x_1, x_2, \ldots \in \mathbb{X}$; by S_η we denote the support $\{x_1, x_2, \ldots\}$ of η. Consider the family $\{N_B\}_{B \in \mathcal{X}_b}$ of maps N_B given by $N_B : \mathcal{N}_f \to \mathbb{N}_0$, $\eta \mapsto N_B(\eta) := \eta(B)$ for every $B \in \mathcal{X}_b$. We equip $\mathcal{N}_f$ with the σ-algebra $\mathfrak{N}$ generated by the family $\{N_B\}_{B \in \mathcal{X}_b}$, i.e., $\mathfrak{N} = \sigma(\{N_B | B \in \mathcal{X}_b\})$.

Definition 1.1. *Let ν be a σ-finite Borel measure on $\mathcal{X}$. The Poisson point process with intensity ν is the unique probability distribution on $(\mathcal{N}_f, \mathfrak{N})$ that satisfies:*

- *The number of points in any bounded Borel set $B \in \mathcal{X}$ is Poisson distributed with intensity given by $\nu(B)$, i.e., for all $B \in \mathcal{X}_b$ and $k \in \mathbb{N}_0$*
$$\mathbb{P}(\{\eta \in \mathcal{N}_f |\ N_B(\eta) = k\}) = \mathrm{e}^{-\nu(B)} \frac{\nu(B)^k}{k!}.$$
- *The numbers of points in disjoint regions are independent of each other, i.e., for all $m \in \mathbb{N}$ and all pairwise disjoint sets $B_1, \ldots, B_m \in \mathcal{X}_b$, the associated random variables $N_{B_1}, \ldots, N_{B_m}$ are independent.*

For the proof of the existence and the uniqueness of the PPP, see [30].

In infinite volume we want to consider the PPP with intensity measure λ_z. In a finite volume $\Lambda \in \mathcal{X}_b$, we can just consider the PPP with intensity measure $\mathbb{1}_\Lambda z \mathrm{d}\lambda$. In the latter case, we want to introduce modifications of the PPP involving interactions between particles. We assume that the particles are interacting pairwise (and not in larger cliques) and introduce a pair-interaction potential v, defined as a measurable and symmetric map $v : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \cup \{\infty\}$. For $x \in \mathbb{X}$ and $\eta \in \mathcal{N}_f$, we define the total interaction $W(x; \eta)$ of x with the configuration $\eta = \sum_{i=1}^{\kappa} \delta_{y_i}$ by
$$W(x; \eta) := \sum_{i=1}^{\kappa} v(x; y_i).$$
Furthermore, we assign a family of measurable functions $(H_n)_{n \in \mathbb{N}}$, $H_n : \mathbb{X}^n \to \mathbb{R} \cup \{\infty\}$, to v, given by
$$H_n(x_1, \ldots, x_n) = \sum_{1 \le i < j \le n} v(x_i, x_j), \quad H(x_1) = 0,$$
for any $n \in \mathbb{N}$ and $(x_1, \ldots, x_n) \in \mathbb{X}^n$. We write $H(\sum_{i=1}^{n} \delta_{x_i}) := H_n(x_1, \ldots, x_n)$. For $\Lambda \in \mathcal{X}_b$, the finite-volume Hamiltonian H_Λ is given by $H_\Lambda(\eta) = H(\eta_\Lambda)$, where η_Λ denotes the configuration restricted to Λ, i.e., $\eta_\Lambda(B) := \eta(B \cap \Lambda)$ for all $B \in \mathcal{X}$. H_Λ encodes the energy of the configuration on the finite volume Λ — the total interaction between particles positioned inside Λ.

A Gibbs point process on $\Lambda \in \mathcal{X}_b$ can be defined via an explicit Radon-Nykodým derivative with respect to the PPP on Λ. The derivative involves the so-called Boltzmann factor e^{-H_Λ}:

Definition 1.2. *The probability distribution on* $(\mathcal{N}_f, \mathfrak{N})$ *which has the Radon-Nikodým derivative*

$$\eta \mapsto \frac{\mathrm{e}^{\lambda_z(\Lambda)}}{\Xi_\Lambda(z)} \mathrm{e}^{-H_\Lambda(\eta)}$$

with respect to the the finite volume PPP with intensity measure $\mathbb{1}_\Lambda z \mathrm{d}\lambda$ *is called the Gibbs point process or (grand-canonical) Gibbs measure on* Λ *with empty boundary conditions (associated to potential* v *and activity* z*). The normalization constant* Ξ_Λ *is called the (grand-canonical) partition function with empty boundary conditions and is given by*

$$\Xi_\Lambda(z) := 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \mathrm{e}^{-\sum_{1 \le i < j \le k} v(y_i, y_j)} \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

Remark 1.1. Different boundary conditions are possible — via modification of the Hamiltonian H_Λ to include interactions with particles on the “boundary” of Λ.

For the infinite volume case, such an explicit definition is, in general, not possible. We introduce infinite-volume Gibbs measures via the GNZ equations, named after Georgii, Nguyen and Zessin. Those reflect a structural property of finite-volume Gibbs measures expressed by a system of integral equations:

Definition 1.3. *A probability distribution* $\mathbb{P}$ *on* $(\mathcal{N}_f, \mathfrak{N})$ *is called a Gibbs measure with interaction potential* v *and activity* z *if*

$$\mathbb{E}\left[\sum_{x \in S_\eta} N_{\{x\}}(\eta) F(x, \eta)\right] = \int_{\mathbb{X}} \mathbb{E}[F(x, \eta + \delta_x) \mathrm{e}^{-W(x;\eta)}] \lambda_z(\mathrm{d}x)$$

for every measurable map $F : \mathbb{X} \times \mathcal{N}_f \to [0, \infty)$*. We denote the set of Gibbs measures with interaction potential* v *and activity* z *by* $\mathscr{G}(v, z)$*.*

Remark 1.2. One can find the proof that finite-volume Gibbs measures satisfy this property in [9]. Moreover, by [9], the GNZ equations are equivalent to the DLR equations, named after Dobrushin, Lanford and Ruelle, that provide the classical definition of Gibbs measures in infinite volume.

Remark 1.3. The questions of uniqueness and existence of Gibbs measures are both highly non-trivial. The phase transition associated to the uniqueness of Gibbs measures is a central topic of ongoing research in the field of statistical mechanics (see, e.g., [7]).

In particular, for the trivial interaction potential $v \equiv 0$, the GNZ equations simplify to the Mecke formula [31] and we recover the PPP with intensity measure λ_z as a unique solution.

Finally, we would like to consider a sequence of functions uniquely associated to a Gibbs measure and encoding all the relevant properties of the system: the so-called n-point correlation functions. In discrete setups, the n-point correlation functions are just given by the probabilities “to encounter particles at n fixed positions” in the random configuration; in continuous setups they are given by certain measure densities. The n-point correlation functions can be rigorously defined as densities of the factorial measures [25]. An alternative definition is the following:

Definition 1.4. *Let $\mathbb{P} \in \mathscr{G}(v,z)$ and $n \in \mathbb{N}$. We define the n-point correlation function $\rho_n(z)$ by*

$$\rho_n(x_1,\dots,x_n;z) := \prod_{j=1}^{n} z(x_j)\mathrm{e}^{-H_n(x_1,\dots,x_n)}\mathbb{E}\left[\prod_{j=1}^{n} \mathrm{e}^{-W(x_j;\eta)}\right] \tag{1.2}$$

for all $(x_1,\dots,x_n) \in \mathbb{X}^n$.

Remark 1.4. Notice that there is a one-to-one correspondence between Gibbs measures and sequences of correlation functions. This follows from the observation that the functions defined in (1.2) coincide with the densities of the factorial measures associated to $\mathbb{P}$, see [25, Lemma 2.2].

For the special case of the PPP with intensity measure λ_z, one recovers the correlation functions $(\rho_n(z))_{n\in\mathbb{N}}$ given by

$$\rho_n(x_1,\dots,x_n;z) = z(x_1)\dots z(x_n).$$

However, in the case of non-trivial interactions, such functions are, in general, non-tractable. To work around that fact, one can employ a perturbative approach and approximate quantities of interest, such as the correlation function, as formal power series in the intensity z of the underlying PPP. I.e., starting from a simplification given by the correlation function of the PPP, one adds terms of higher order encoding the particle interactions via the Boltzman factor and expects that the resulting series approximate the Gibbs correlation functions — if the interactions are weak and the particle density is low enough. The so-called cluster expansion is such a perturbative approach. It does not only ensure that the partial sums of the resulting power series approximate the correlation functions, but also that the series — which are the Taylor expansions of the correlation functions around $z = 0$ — converge in a non-trivial region. The study of those power series expansions and their domains of convergence is the topic of [a] and constitutes the initial research goal behind this thesis. We now want to introduce the approach in more details — to discuss our main result from [a] and to present some supplementary results on cluster expansion not included therein.

1.3 Mayer trick and combinatorics

We follow the so-called Mayer trick introduced in the 1940's [34] that is at the heart of all combinatorial approaches to cluster expansion. The trick is based on the idea to encode interactions between pairs of+ particles by weights of edges in an associated weighted graph, leading to a graphical expansion of the partition function. To elaborate the idea, we introduce several sets of graphs and a way to assign weights (associated to the pair-interaction potential) to those graphs.

For every $n \in \mathbb{N}$, let us denote by $\mathcal{G}_n$ the set of all graphs on the vertex set $[n] := \{1,\dots,n\}$ and let us denote by $\mathcal{C}_n \subset \mathcal{G}_n$ the set of all connected graphs on $[n]$.

To introduce graph weight, we define Mayer's f function associated to the potential v by

$$f(x,y) = \mathrm{e}^{-v(x,y)} - 1, \quad x,y \in \mathbb{X}.$$

Let $G \in \mathcal{G}_n$ be a graph with vertex set $[n]$ and edge set $E(G)$. For $(x_1, \dots, x_n) \in \mathbb{X}^n$, the graph weight $w(G; x_1, \dots, x_n)$ is given as the product of the edge weights dictated by Mayer's f function:

$$w(G; x_1, \dots, x_n) := \prod_{i,j \in E(G)} f(x_i, x_j).$$

We want to expand the logarithm of the partition function $\Xi_\Lambda(z)$, that is the normalization constant for a finite-volume Gibbs measure with empty boundary conditions, around the activity $z = 0$. To do so, we first notice that $\Xi_\Lambda(z)$ is given by the exponential generating function for graphs: For all $k \in \mathbb{N}$ and $(y_1, \dots, y_k) \in \mathbb{X}^k$ we have

$$\prod_{1 \le i < j \le k} (1 + f(y_i, y_j)) = \sum_{G \in \mathcal{G}_k} \prod_{\{i,j\} \in E(G)} f(y_i, y_j)$$

and therefore

$$\Xi_\Lambda(z) = 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \prod_{1 \le i < j \le k} (1 + f(y_i, y_j)) \lambda_z^k(\mathrm{d}\boldsymbol{y}) = \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \sum_{G \in \mathcal{G}_k} \prod_{\{i,j\} \in E(G)} f(y_i, y_j) \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

By a well-known combinatorial argument (see, e.g., [12, 15] and compare to the proof of Proposition 1.9) based on the fact that the weight of a graph is given by the product of the weights of its connected components, the exponential generating function for graphs is given by the exponential of the exponential generating function for connected graphs. That means, that taking the logarithm of $\Xi_\Lambda(z)$ corresponds to discarding the graphs that are not connected, i.e., $\log \Xi_\Lambda(z)$ can be expressed as the exponential generating function for connected graphs: For $n \in \mathbb{N}$, we define the n-th Ursell function by

$$\varphi_n^{\mathsf{T}}(x_1, \dots x_n) := \sum_{G \in \mathcal{C}_n} w(G; x_1, \dots, x_n), \quad (x_1, \dots, x_n) \in \mathbb{X}^n,$$

then we can capture the following identity between formal power series:

$$\log \Xi_\Lambda(z) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \varphi_k^{\mathsf{T}}(y_1, \dots y_k) \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

In the context of statistical mechanics, $\log \Xi_\Lambda$ is the free energy of the system and encodes the relevant thermodynamic properties of the system, which can be extracted from it by taking derivatives; naturally, one is interested in the convergence radius of this activity expansion.

Alternatively, instead of $\log \Xi_\Lambda$, one can (and we will in the following) consider the activity expansions of the correlation functions. Again, those are given by exponential generating functions for a certain type of graphs that we call multi-rooted graphs: Let $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$. We call a graph $G \in \mathcal{G}_{n+k}$ a multi-rooted graph with root vertices $\{1, \dots, n\}$ if every vertex $j \in \{n+1, \dots, n+k\}$ connects to some vertex $i \in \{1, \dots, n\}$. We denote the collection of multi-rooted graphs with root vertices $\{1, \dots, n\}$ by $\mathcal{D}_{n,n+k}$. Furthermore, we define

$$\psi_{n,n+k}(x_1, \dots, x_{n+k}) := \sum_{G \in \mathcal{D}_{n,n+k}} w(G; x_1, \dots, x_{n+k}), \quad (x_1, \dots, x_{n+k}) \in \mathbb{X}^{n+k}.$$

Notice that for $n = 1$, we recover the standard Ursell functions, i.e., $\psi_{1,1+k} = \varphi^{\mathsf{T}}_{1+k}$.

Finally, we can define our main object of study: For an activity function z and $n \in \mathbb{N}$, the activity expansion $\rho_n(z)$ for the n-point correlation function is given by

$$\rho_n(x_1, \dots, x_n; z) = \sum_{k \geq 0} \frac{1}{k!} \int_{\mathbb{X}^k} \psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k) z(x_1) \dots z(x_n) \lambda_z^k(\mathrm{d}\boldsymbol{y}) \tag{1.3}$$

for $(x_1, \dots, x_n) \in \mathbb{X}^n$.

1.4 Our main cluster expansion result: Characterizing the domain of convergence

Later, in Subsection 1.5, we will justify the fact that the activity expansions ρ_n are our main object of interest — by proving that they provide power series expansions for the correlation functions corresponding to Gibbs measures with empty boundary conditions and by connecting them to expansions known from literature. But what can be said about the radii of convergence of those power series? In the literature, there are essentially three classical approaches to cluster expansions for discrete systems with hard-core interactions:

- The combinatorial approach: This method is based on the free-energy expansion in terms of connected graphs due to Mayer's trick and was clearly formulated by Battle [1] first, though there is preceding work by other authors. The convergence of the expansion is established by tree-graph bounds [1,4], i.e., by bounding sums over weighted connected graphs by sums over weighted trees; later Fernández and Procacci [13] improved on the method by rediscovering an old argument by O. Penrose [38] and employing so-called tree partition schemes to establish tree-graph identities. Another novel ingredient contributing to their improvement was an iterative argument asserting that it is enough to bound the expression involving single-generation trees to obtain convergence (and analyticity) of the free-energy expansion.
- The inductive approach: This method was introduced by Kotecký and Preiss [28]; later it was refined by Dobrushin [10] into its "final" form, abandoning any reference to graphical expansion or power series altogether (Dobrushin himself referred to it as a "no-cluster expansion"). While the proofs of given sufficient conditions are elegant and accessible, they do not provide a clear path for further improvement of the conditions since there is no recipe to come up with new inductive hypotheses. Recently, an inductive proof à la Dobrushin of the Fernández-Procacci condition emerged [14].
- The analytical approach: This method was introduced by Gruber and Kunz [18]. It involves a system of integral fixed-point equations over a suitable Banach space of functions — the Kirkwood-Salsburg equations — that is shown to be solved by the activity expansions for the correlation functions. The convergence of the expansion is essentially shown via the Banach fixed-point theorem and the argument involves showing that an operator associated to the Kirkwood-Salsburg equations is a contraction on the Banach space. Notice that the uniqueness of the solution ensures uniqueness of the Gibbs measure. Using a slight modification of the approach — the extended Gruber-Kunz approach — Bissacot, Fernández and Procacci [3] were able to reproduce the Fernández-Procacci condition.

Notice that, until recently, in the discrete setup of polymer models, the three methods produced at best the same bound on the convergence radius of the cluster expansion — the one provided by the Ferández-Procacci condition (see [3, 13] for a comparison of classical conditions in discrete setups). However, we are able to improve on the Fernández-Procacci condition using the analytical approach (see [a, Proposition 2.4]).

A lot was done to generalize those classic results to hold in continuous setups and for larger classes of pair interactions. Ueltschi [42] generalized the classic Kotecký-Preiss condition to hold, in particular, in continuous setups and for soft-core interaction; Faris [11] and Jansen [25] followed by generalizing the Fernández-Procacci criterion.

Let us briefly summarize our results from our contribution [a], "Cluster expansions: Necessary and sufficient conditions": We follow the spirit of Bissacot, Fernández and Procacci [3], employing a similar (but slightly modified) analytical approach á la Gruber and Kunz in a much more general setup (and combining it with some combinatorial considerations). The activity expansions ρ_n of Gibbs correlation functions are given by exponential generating functions for multi-rooted graphs. Multi-rooted graphs have a certain structural property: Taking a multi-rooted graph and removing an arbitrary root with the incident edges produces a multi-rooted graph, where every neighbor of the removed root becomes a root vertex itself. The weight of the original multi-rooted graph is equal to the weight of the resulting graph times the weight of the edges removed. This structural property, on the level of generating functions, is expressed by the Kirkwood-Salsburg equations that can be exploited to find sufficient conditions for convergence.

Consider an activity z and a non-negative potential $v \geq 0$. Let us formulate a particular sign-flipped version of the classical Kirkwood-Salsburg equations by defining the sign-flipped Kirkwood-Salsburg operator $\tilde{K}_z$ — that acts on families $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$ of measurable symmetric functions $\xi_n : \mathbb{X}^n \to \mathbb{R}^+$ — as

$$\begin{aligned}(\tilde{K}_z\xi)_n(x_1,\dots,x_n) := z(x_1) \prod_{i=2}^{n}(1+f(x_1,x_i))\Big(&\mathbb{1}_{\{n\geq 2\}}\xi_{n-1}(x_2,\dots,x_n) \\ &+\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\prod_{j=1}^{k}\big|f(x_1,y_j)\big|\,\xi_{n-1+k}(x_2,\dots,x_n,y_1,\dots,y_k)\lambda^k(\mathrm{d}\boldsymbol{y})\Big), \quad (1.4)\end{aligned}$$

for all $n \in \mathbb{N}$ and $(x_1,\dots,x_n) \in \mathbb{X}^n$. One can show that the non-negative versions $\tilde{\rho}_n(z)$ of the activity expansions $\rho_n(z)$, given by

$$\tilde{\rho}_n(x_1,\dots,x_n;z) := \sum_{k=0}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\big|\psi_{n,n+k}(x_1,\dots,x_n,y_1,\dots,y_k)\big|z(x_1)\cdots z(x_n)\lambda_z^k(\mathrm{d}\boldsymbol{y})$$

for all $(x_1,\dots,x_n)$, satisfy the equations

$$\tilde{K}_z\tilde{\rho}(z) + e_z = \tilde{\rho}(z).$$

Thus, on level of formal series, the activity expansions $\tilde{\rho}_n(z)$ can be written as Neumann-type series $\sum_{n=0}^{\infty}\tilde{K}_z^n e_z$ and the partial sums $\sum_{n=0}^{N}\tilde{K}_z^n e_z$ of those series are given by truncated versions of $\tilde{\rho}_n(z)$ involving graphs with at most $N \in \mathbb{N}$ nodes. While the classic Kotecký-Preiss and Gruber-Kunz conditions ensure the convergence of the Neumann series $\sum_{n=0}^{\infty}\tilde{K}_z^n$

with respect to an operator norm by ensuring that the operator $\tilde{K}_z$ is a contraction, this is not necessary for the convergence of the Neumann-type series $\sum_{n=0}^{\infty} \tilde{K}_z^n e_z$ in general. We want to provide a condition which is both sufficient and necessary for the convergence of the Neumann-type series $\sum_{n=0}^{\infty} \tilde{K}_z^n e_z$ in the case of repulsive interaction, which is the idea behind the extended Gruber-Kunz approach [3].

In the following, we characterize the convergence of the activity expansions $\tilde{\rho}_n$ in terms of existence of non-negative, measurable solutions to the associated Kirkwood-Salsburg-type inequalities; it is a somewhat simplified version of [a, Theorem 2.1], for the special case of repulsive interactions (i.e., for $v \geq 0$):

Theorem 1.5. *In the case of repulsive interactions, the following two conditions are equivalent:*

(i) There is a family $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$ of measurable symmetric functions $\xi_n : \mathbb{X}^n \to \mathbb{R}^+$ such that

$$z(x_1)\delta_{n,1} + (\tilde{K}_z \boldsymbol{\xi})_n\,(x_1,\ldots,x_n) \leq \xi_n(x_1,\ldots,x_n) \tag{1.5}$$

for all $n \in \mathbb{N}$ and $(x_1,\ldots,x_n) \in \mathbb{X}^n$.

(ii) The series $\rho_n(x_1,\ldots,x_n;z)$ converges absolutely, for all $n \in \mathbb{N}$ and $(x_1,\ldots,x_n) \in \mathbb{X}^n$.

Moreover, if (i) is satisfied, then

$$\sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \bigl|\psi_{n,n+k}(x_1,\ldots,x_n,y_1,\ldots,y_k)\bigr| z(x_1)\cdots z(x_n)\lambda_z^k(\mathrm{d}\boldsymbol{y}) \leq \xi_n(x_1,\ldots,x_n)$$

on $\mathbb{X}^n$, for all $n \in \mathbb{N}$.

This approach allows us to unify existing convergence criteria in a joint context (e.g., the classic Kotecký-Preiss and Fernández-Procacci sufficient conditions) but it also allows us to find new conditions by designing suitable ansatz functions ξ (e.g., as approximations of the the activity expansions $\tilde{\rho}_n$) to satisfy the Kirkwood-Salsburg-type inequalities (1.5). We demonstrate the potential of this approach to find new sufficient conditions in various discrete and continuous setups with hard-core repulsive interactions and are able to use it to improve on the Fernández-Procacci condition in the setup of abstract polymer models (see [a, Proposition 2.4]).

1.5 Supplementary cluster expansion results

1.5.1 Activity expansions of the correlation functions

The results from this section — except for Lemma 1.9 — do not appear in [a] and are supposed to be read as supplementary results providing context for the choice of the power series ρ_n as the candidate activity expansions for the correlation functions. We would like to connect different known representations for the activity expansions to each other and to the correlation functions themselves.

First, we want to consider some combinatorial results to understand the graphical expansions (1.3) on the level of the underlying graphs: In the literature there are two types of expansions that (can be interpreted to) stem from two different possibilities to construct multi-rooted graphs on a given vertex set. While the following expansions (see Proposition 1.7 and Proposition 1.9) themselves can not be considered novel and appear in different setups under various assumptions, we provide a combinatorial interpretation of the different representations while rigorously proving them to hold in our quite general setup without assuming concrete sufficient conditions for convergence.

First we introduce a simple procedure to construct multi-rooted graphs, which allows us to explain how the series ρ_n relates to the series ρ_n^{T} given by the exponential generating function for connected graphs: For all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$, let

$$\rho_n^{\mathsf{T}}(x_1, \dots, x_n) := \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \varphi_{n,n+k}^{\mathsf{T}}(x_1, \dots, x_n, y_1, \dots, y_k) z(x_1) \cdots z(x_n) \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

Lemma 1.6. *For all $n \in \mathbb{N}$, all $k \in \mathbb{N}_0$ and all $(x_1, \dots, x_{n+k}) \in \mathbb{X}^{n+k}$, we have*

$$\psi_{n,n+k}(x_1, \dots, x_{n+k}) = \sum_{\{V_1, \dots, V_m\}} \prod_{\ell=1}^{m} \varphi_{|V_\ell|}^{\mathsf{T}}\big((x_j)_{j \in V_\ell}\big)$$

where the sum runs over all set partitions $\{V_1, \dots, V_m\}$ of $\{1, \dots, n+k\}$ such that every block V_ℓ contains at least one root vertex $i \in \{1, \dots, n\}$.

Proof. To construct a multi-rooted graph $G \in \mathcal{D}_{n,n+k}$, consider a partition $\{V_1, ..., V_m\}$, $m \leq n$, of the vertex set $\{1, ..., n+k\}$, such that every block V_l of the partition contains at least one vertex from the set of roots $\{1, ..., n\}$. Then simply pick connected graphs $G_1, ..., G_m$ on the vertex sets $V_1, ..., V_m$. The weight of the resulting multi-rooted graph G is given by

$$w(G; x_1, \dots, x_{n+k}) = \prod_{\ell=1}^{m} w\big(G_\ell; (x_j)_{j \in V_\ell}\big). \tag{1.6}$$

Since every multi-rooted graph can be obtained in that fashion and since the sets of the graphs corresponding to different partitions $V_1, ..., V_m$ are disjoint, the claim of the lemma follows immediately from (1.6) by summation over all connected graphs $G_1, ..., G_m$ on $V_1, .., V_m$ and summation over all partitions $V_1, ..., V_m$, $m \leq n$. $\square$

The relations (between the coefficients $\psi_{n,n+k}$ of the activity expansions ρ_n and the Ursell functions) established in Lemma 1.6 lead to the following representation of ρ_n:

Proposition 1.7. *Suppose that all series $\rho_n(x_1, \dots, x_n; z)$ are absolutely convergent for some activity function z. Then also all series $\rho_n^{\mathsf{T}}(x_1, \dots, x_n; z)$ are absolutely convergent, and we have for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$*

$$\rho_n(x_1, \dots, x_n; z) = \sum_{\{R_1, \dots, R_m\}} \prod_{i=1}^{m} \rho_{|R_i|}^{\mathsf{T}}\big((x_j)_{j \in R_i}; z\big), \tag{1.7}$$

where the sum runs over all set partitions $\{R_1, \dots, R_m\}$ of $\{1, \dots, n\}$.

Remark 1.5. The right-hand side of (1.7) is the representation for the correlation functions in finite volume from [42, Theorem 2].

Proposition 1.7 follows from Lemma 1.6, it is proven at the end of this subsection (see 1.5.2).

Unlike in the construction underlying our proof of Lemma 1.6, one might obtain multi-rooted graphs by considering partitions of the root vertices and partitions of the non-root vertices separately (see also Figure 2). That is precisely the construction underlying the proof of the following lemma (Lemma 1.8). The lemma leads to a different representation of ρ_n in terms of the Ursell functions (established in Proposition 1.9). While that second representation is of its own interest, it additionally allows us to show that the series ρ_n do indeed correspond to the correlation functions (we use it in the proof of Corollary 1.10).

Lemma 1.8. *For all $n \in \mathbb{N}$, all $k \in \mathbb{N}_0$ and all $(x_1, \dots, x_{n+k}) \in \mathbb{X}^{n+k}$, we have*

$$\psi_{n,n+k}(x_1, \dots, x_{n+k}) = \prod_{1 \le i < j \le n} \big(1 + f(x_i, x_j)\big) \times \sum_{\{V_1,\dots,V_r\}} \prod_{\ell=1}^{r} \Bigg(\prod_{\substack{1 \le i \le n, \\ j \in V_\ell}} \big(1 + f(x_i, x_j)\big) - 1 \Bigg) \varphi^{\mathsf{T}}_{|V_\ell|}\big((x_j)_{j \in V_\ell}\big), \quad (1.8)$$

where the sum runs over all set partitions $\{V_1, \dots, V_r\}$ of non-root vertices $\{n+1, \dots, n+k\}$.

Proof. Every multi-rooted graph $G \in \mathcal{D}_{n,n+k}$ can be constructed in the following way. On the root set $\{1, \dots, n\}$ pick an arbitrary graph G_0. On the complement of the root set do the following construction: Partition the set of the non-root vertices into r sets $V_1, ..., V_r$, $r \le k$. For every block V_ℓ, pick a connected graph G_ℓ with vertex set V_ℓ, and in addition a non-empty set of edges $E_\ell \subset \big\{\{i,j\} \mid i \in \{1, \dots, n\},\, j \in V_\ell\big\}$. Then the graph G on $\{1, \dots, n+k\}$ with the edge set given by the union of $E_1, \dots, E_\ell$ and the edge sets of $G_0, G_1, \dots, G_r$ is contained in $\mathcal{D}_{n,n+k}$, its graph weight is

$$w(G; x_1, \dots, x_{n+k}) = w(G_0; x_1, \dots, x_n) \prod_{\ell=1}^{r} \Bigg(\prod_{\{i,j\} \in E_\ell} f(x_i, x_j) \Bigg) w\big(G_\ell; (x_j)_{j \in V_\ell}\big).$$

Summation over G_0 yields $\prod_{1 \le i < j \le n}(1 + f(x_i, x_j))$. Summation over connected graphs G_ℓ yields $\varphi^{\mathsf{T}}_{|V_\ell|}(\boldsymbol{x}_{V_\ell})$. Finally, we see that summation over the edge sets E_ℓ yields the factor $\prod_{1 \le i \le n, j \in V_\ell}(1 + f(x_i, x_j)) - 1$. □

The relations (between the coefficients $\psi_{n,n+k}$ of the activity expansions ρ_n and the Ursell functions) established in Lemma 1.8 lead to the following a representation of the series ρ_n:

Proposition 1.9. *Suppose that all series $\rho_n(x_1, \dots, x_n; z)$ are absolutely convergent for some activity function z. Then*

$$\rho_n(x_1, \dots, x_n; z) = z(x_1) \cdots z(x_n) \prod_{1 \le i < j \le n} \big(1 + f(x_i, x_j)\big) \times \exp\Bigg(\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \Bigg[\prod_{\substack{1 \le i \le n \\ 1 \le j \le k}} \big(1 + f(x_i, y_j)\big) - 1 \Bigg] \varphi^{\mathsf{T}}_k(y_1, \dots, y_k) \lambda_z^k(\mathrm{d}\boldsymbol{y}) \Bigg),$$

for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$.

Remark 1.6. This representation is well-established in discrete setups in the literature: E.g., in the setup of abstract polymers in finite volume, the exponential on the right-hand side for $n = 1$ coincides with the expansion $\Theta^{\Lambda}_{x_1}$ of the reduced 1-function in x_1 from [3]. For general $n \geq 1$ the exponential appears in [3] in the form of a reconstruction formula involving the 1-expansions $\Theta^{\Lambda \setminus x_{i+1},\dots,x_n}_{x_i}$, $i \in \{1, ..., n\}$.

The proposition follows from Lemma 1.8, it is proven in Section 1.5.2.

We see how the two different constructions of multi-rooted graphs lead to two different representations of the activity expansions ρ_n given by (1.3) appearing throughout the literature in various setups.

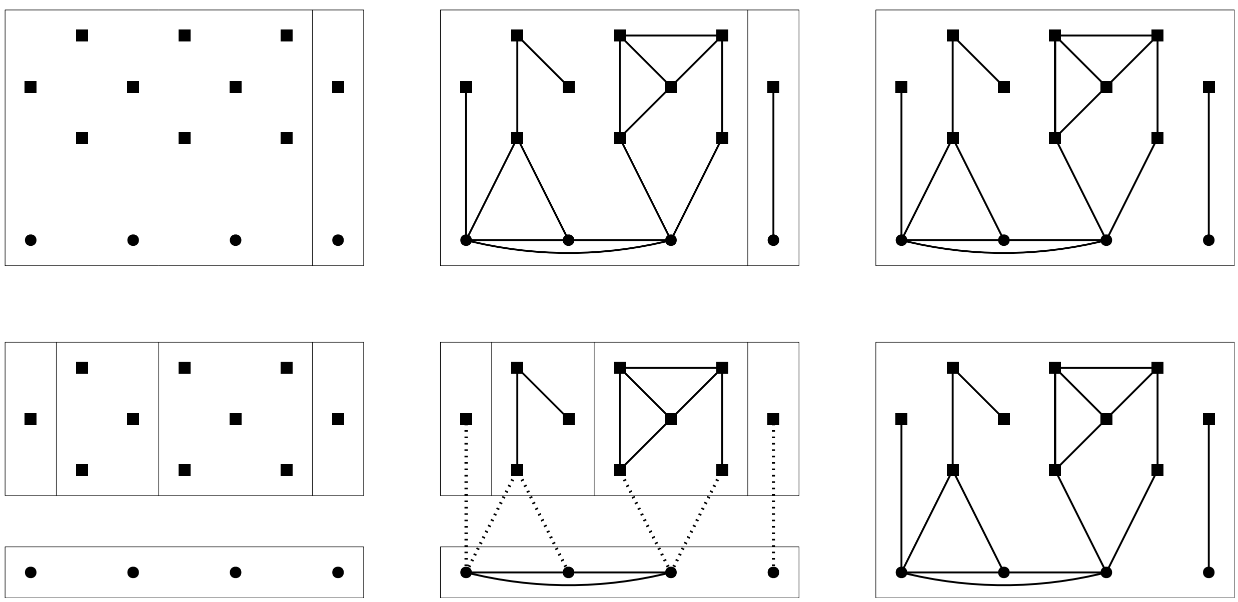

Figure 2: Here, we see a multi-rooted graph $G \in \mathcal{D}_{4,14}$ — the round nodes depict roots, the square nodes depict non-root vertices. In the first line, using the example of G, we illustrate the construction of multi-rooted graphs underlying Lemma 1.6, where in the first step the vertex set is partitioned in blocks containing at least one root. In the second line, we illustrate the construction of multi-rooted graphs underlying Lemma 1.8, where in the first step the roots and the non-root vertices are partitioned separately.

Now, we would like to convince the reader even further that the power series $(\rho_n)_{n \in \mathbb{N}}$ are indeed the right object to study — by providing a result indicating that, for every $n \in \mathbb{N}$, ρ_n corresponds to the n-point correlation function of the grand-canonical Gibbs measure with empty boundary conditions. To do so, we want to introduce some auxiliary notation first: Let us fix an interaction potential $v \geq 0$ and an activity z. For $\Lambda \in \mathcal{X}_b$, let P_Λ denote the associated finite-volume grand-canonical Gibbs measure with empty boundary conditions (non-negativity of the potential v is sufficient for the existence of the finite-volume Gibbs measure). For $n \in \mathbb{N}$, define the candidate expansions for the finite-volume n-point correlation functions, denoted by $\rho_{n,\Lambda}$, just as ρ_n in (1.3) but with integrals over Λ^n instead of $\mathbb{X}^n$.

Corollary 1.10. *Let z be an activity function and suppose that $\rho_n(x_1,\dots,x_n;z)$ converges absolutely for all $n \in \mathbb{N}$ and all $(x_1,\dots,x_n) \in \mathbb{X}^n$. Then for every Borel set $\Lambda \subset \mathbb{X}$ with $\int_\Lambda \tilde{\rho}_1(x;z)\lambda(\mathrm{d}x) < \infty$, the functions $\rho_{n,\Lambda}(z)$ are the correlation functions of the finite-volume Gibbs measure P_Λ.*

Remark 1.7. Suppose there exists a sequence $(\Lambda_m)_{m\in\mathbb{N}}$ in $\mathcal{X}$ such that $\int_{\Lambda_m} \tilde{\rho}_1(x)\lambda_z(\mathrm{d}x) < \infty$ for all $m \in \mathbb{N}$. Then as $m \to \infty$, the finite-volume correlation functions $\rho_{n,\Lambda_m}(z)$ converge pointwise to the functions $\rho_n(z)$. Usually this is accompanied by a convergence of Gibbs measures $\mathsf{P}_{\Lambda_m} \to \mathsf{P}$ and the functions $\rho_n(z)$ are the correlation functions of the infinite volume Gibbs measure P. A rigorous statement on convergence is beyond the purpose of this work; some relevant considerations and references are given in [25, Section 2.3 and Appendix B].

Proof. The corollary essentially follows from Proposition 1.9 proven at the end of this subsection. We show that

$$\rho_{n,\Lambda_m}(x_1,\dots,x_n;z) = z(x_1)\cdots z(x_n) \prod_{1\le i<j\le n} \bigl(1+f(x_i,x_j)\bigr) \int_{\mathcal{N}_f} \mathrm{e}^{-\sum_{i=1}^n W(x_i,\eta)} \mathsf{P}_{\Lambda_m}(\mathrm{d}\eta). \tag{1.9}$$

To prove this identity, we observe that the expected value on the right side of (1.9) can be rewritten as a ratio of partition functions. To do so, notice that the interactions of random points in η with points $x_1,\dots,x_n$ on the right side of (1.9) can be absorbed into a modified activity as follows: For fixed $(x_1,\dots,x_n) \in \mathbb{X}^n$ define the activity $\hat{z}$ by

$$\widehat{z}(y) := z(y)\mathrm{e}^{-\sum_{i=1}^n v(x_i;y)} = z(y)\prod_{i=1}^n \bigl(1+f(x_i,y)\bigr), \quad y \in \mathbb{X}.$$

For a Borel set $\Lambda \subset \mathbb{X}$ with $\lambda_z(\Lambda) < \infty$, consider the grand-canonical partition function at activity z given by

$$\Xi_\Lambda(z) := 1 + \sum_{k=1}^\infty \frac{1}{k!} \int_{\Lambda^k} \mathrm{e}^{-\sum_{1\le i<j\le k} v(y_i,y_j)} \lambda_z^n(\mathrm{d}\boldsymbol{y}),$$

using the notation introduce above, we can write

$$\int_{\mathcal{N}_f} \mathrm{e}^{-\sum_{i=1}^n W(x_i,\eta)} \mathsf{P}_\Lambda(\mathrm{d}\eta) = \frac{\Xi_\Lambda(\widehat{z})}{\Xi_\Lambda(z)}.$$

Now, assume for a moment that

$$\sum_{k=1}^\infty \frac{1}{k!} \int_{\Lambda^k} \bigl|\varphi_k^\mathsf{T}(y_1,\dots,y_k)\bigr| \lambda_z^k(\mathrm{d}\boldsymbol{y}) < \infty, \tag{1.10}$$

then it is well-known from the theory of cluster expansions (as mentioned in Subsection 1.3) that

$$\log \Xi_\Lambda(z) = \sum_{k=1}^\infty \frac{1}{k!} \int_{\Lambda^k} \varphi_k^\mathsf{T}(y_1,\dots,y_k)\lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

If the convergence condition (1.10) holds true for z, then because of $v \geq 0$ the modified activity $\widehat{z}$ satisfies a similar condition and $\log \Xi_\Lambda(\widehat{z})$ has a similar expansion. Therefore

$$\log \frac{\Xi_\Lambda(\widehat{z})}{\Xi_\Lambda(z)} = \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \Bigg(\prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq k}} \big(1 + f(x_i, y_j)\big) - 1 \Bigg) \varphi_k^{\mathsf{T}}(\boldsymbol{y}) \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

and we recognize the exponent from Proposition 1.9. The identities from Proposition 1.9 hold true in finite volume as well, so we conclude that Eq. (1.9) holds true.

Thus it remains to check that a sequence $(\Lambda_m)_{m\in\mathbb{N}}$ with $\lambda_z(\Lambda_m) < \infty$ for all $m \in \mathbb{N}$ and $\Lambda_m \nearrow \mathbb{X}$, which satisfies the convergence condition (1.10), exists. Since λ is σ-finite and $\tilde{\rho}_1$ is pointwise finite by assumption, the measure given by $\tilde{\rho}_1(x) d\lambda(x)$ is again σ-finite. That ensures the existence of a sequence $(\Lambda_m)_{m\in\mathbb{N}} \subset \mathbb{X}$ with $\lambda_z(\Lambda_m) < \infty$ and $\Lambda_m \nearrow \mathbb{X}$ such that $\int_{\Lambda_m} \tilde{\rho}_1(x) d\lambda(x) < \infty$. To see that the sets Λ_m thus satisfy (1.10) notice that

$$\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\Lambda^k} \big|\varphi_k^{\mathsf{T}}(y_1, \dots, y_k)\big| \, \lambda_z^k(\mathrm{d}\boldsymbol{y}) \leq \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \int_{\Lambda^k} \big|\varphi_k^{\mathsf{T}}(y_1, \dots, y_k)\big| \, \lambda_z^k(\mathrm{d}\boldsymbol{y})$$

holds for every Borel set Λ. If, moreover, $\int_\Lambda \tilde{\rho}_1(x) d\lambda(x) < \infty$ holds, one can interchange the order of summation and integration in the expression on the right-hand side of the last inequality to obtain

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} \int_{\Lambda^k} \big|\varphi_k^{\mathsf{T}}(y_1, \dots, y_k)\big| \, \lambda_z^k(\mathrm{d}\boldsymbol{y}) = \int_\Lambda \tilde{\rho}_1(x) d\lambda(x) < \infty,$$

which concludes the proof of the corollary. □

1.5.2 Proofs of Propositions 1.7 and 1.9

Proof of Proposition 1.7. The proposition follows from Lemma 1.6. Clearly there is a one-to-one correspondence between on the one hand set partitions $\{V_1, \dots, V_m\}$ of $\{1, \dots, n+k\}$ such that every block contains at least one root vertex $i \in \{1, \dots, n\}$, and on the other hand pairs consisting of (i) a set partition $P = \{R_1, \dots, R_m\}$ of the roots $\{1, \dots, n\}$ and (ii) a collection $(V'_R)_{R \in P}$ of sets, indexed by the blocks of the partition P, such that the sets V'_{R_j} are pairwise disjoint and their union is $\{n+1, \dots, n+k\}$ (we do not exclude $V'_R = \varnothing$).

As a consequence,

$$\psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k) = \sum_{P = \{R_1, \dots, R_m\}} \sum_{(J_R)_{R\in P}} \prod_{\ell=1}^{m} \varphi^{\mathsf{T}}_{|R_\ell| + |J_\ell|}\big(\boldsymbol{x}_{R_\ell}, \boldsymbol{y}_{J_\ell}\big) \tag{1.11}$$

where the sum is over set partitions P of $\{1, \dots, n\}$ and tuples $(J_R)_{R\in P}$ of pairwise distinct sets J_R (possibly empty) such that $\bigcup_{R\in P} J = \{1, \dots, k\}$. It is a general fact that if $A_k^{(\ell)}(y_1, \dots, y_k)$, $k \in \mathbb{N}_0$, $\ell = 1, \dots, m$, is a family of symmetric functions with

$$\sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \big|A_k^{(\ell)}(y_1, \dots, y_k)\big| \lambda_z^k(\mathrm{d}\boldsymbol{y}) < \infty,$$

then

$$\sum_{\{R_1,\dots,R_m\}} \prod_{\ell=1}^m \left(\sum_{k=0}^\infty \frac{1}{k!} \int_{\mathbb{X}^k} A_k^{(\ell)}(\boldsymbol{y}) \lambda_z^k(\mathrm{d}\boldsymbol{y}) \right) = \sum_{k=0}^\infty \frac{1}{k!} \int_{\mathbb{X}^k} \sum_{(J_1,\dots,J_m)}^{(k)} \prod_{\ell=1}^m A_{|J_\ell|}^{(\ell)}(\boldsymbol{y}_{J_\ell}) \lambda_z^k(\mathrm{d}\boldsymbol{y}), \quad (1.12)$$

where the sum over $(J_1,\dots,J_m)$ is over tuples of pairwise disjoint sets (with $J_\ell = \varnothing$ allowed) with union $\bigcup J_\ell = \{1,\dots,k\}$. Assume for a moment that the truncated functions ρ_n^T converge absolutely, i.e., assume that

$$\prod_{i=1}^n z(x_i) \sum_{k=0}^\infty \frac{1}{k!} \int_{\mathbb{X}^k} \left| \varphi_{n+k}^\mathsf{T}(x_1,\dots,x_n,y_1,\dots,y_k) \right| \lambda_z^k(\mathrm{d}\boldsymbol{y}) < \infty \quad (1.13)$$

for all $n \in \mathbb{N}$ and $(x_1,\dots,x_n) \in \mathbb{X}^n$. Then, for every fixed $\boldsymbol{x}$ and set partition $\{R_1,\dots,R_m\}$ we can apply Eq. (1.12) to $A_k^{(\ell)}(y_1,\dots,y_k) = \varphi_{|R_\ell|+k}^\mathsf{T}(\boldsymbol{x}_{R_\ell}, y_1,\dots,y_k) \prod_{i\in R_\ell} z(x_i)$. Together with (1.11), this yields the formula from the proposition. Thus it remains to check the absolute convergence (1.13).

We prove (1.13) by induction over $n \geq 1$. First, notice that (1.13) holds for $n = 1$ and for all $x_1 \in \mathbb{X}$, since for every $k \in \mathbb{N}$ the set of 1-rooted graphs $\mathcal{D}_{1,1+k}$ is exactly the set of connected graphs $\mathcal{C}_{1+k}$:

$$\rho_1^\mathsf{T}(x_1) = \sum_{k=0}^\infty \frac{z(x_1)}{k!} \int_{\mathbb{X}^k} \varphi_{1+k}^\mathsf{T}(x_1,y_1,\dots,y_k) \lambda_z^k(\mathrm{d}\boldsymbol{y}) = \rho_1(x_1)$$

and the 1-point functions $\rho_1(z)$ converge absolutely by assumption.

Now assume that the truncated functions $\rho_1^\mathsf{T},\dots,\rho_n^\mathsf{T}$ converge absolutely for some $n \geq 1$. To obtain absolute convergence (1.13) for $n+1$, we realize that for our choice of the functions $A_k^{(\ell)}$ the identity (1.12) can be rewritten as

$$\rho_{n+1}(x_1,\dots,x_{n+1};z) = \rho_{n+1}^\mathsf{T}(x_1,..,x_{n+1};z) + \sum_{\substack{P=\{R_1,\dots,R_m\} \\ P\neq\{\{1,\dots,n+1\}\}}} \prod_{i=1}^m \rho_{|R_i|}^\mathsf{T}\big((x_j)_{j\in R_i};z\big), \quad (1.14)$$

where the finite sum on the right side runs over all partitions P of the roots $\{1,\dots,n+1\}$ except for the trivial partition $P = \{\{1,\dots,n+1\}\}$. Consider

$$\sum_{\substack{P=\{R_1,\dots,R_m\} \\ P\neq\{\{1,\dots,n+1\}\}}} \prod_{i=1}^m \rho_{|R_i|}^\mathsf{T}\big((x_j)_{j\in R_i};z\big)$$

and notice that it converges absolutely by the inductive hypothesis (as a finite sum of Cauchy products of absolutely convergent series), while $\rho_{n+1}(x_1,\dots,x_{n+1};z)$ converges absolutely by assumption. Therefore,

$$\tilde{\rho}_{n+1}^\mathsf{T}(x_1,\dots,x_{n+1};z) \leq \tilde{\rho}_{n+1}(x_1,\dots,x_{n+1};z) + \sum_{\substack{P=\{R_1,\dots,R_m\} \\ P\neq\{\{1,\dots,n+1\}\}}} \prod_{i=1}^m \tilde{\rho}_{|R_i|}^\mathsf{T}\big((x_j)_{j\in R_i};z\big) < \infty.$$

and we get the absolute convergence of $\rho_{n+1}^\mathsf{T}(x_1,..,x_{n+1};z)$, which concludes the induction and the proof of the proposition. $\square$

Proof of Proposition 1.9. The proposition follows from Lemma 1.8. By Lemma 1.8,

$$\psi_{n,n+k}(x_1,\dots,x_n,y_1,\dots,y_k) = \prod_{1\le i<j\le n}\big(1+f(x_i,x_j)\big) \times \sum_{\{V_1,\dots,V_r\}}\prod_{\ell=1}^{r} A_{n,|V_\ell|}\big(x_1,\dots,x_n;(y_j)_{j\in V_\ell}\big), \quad (1.15)$$

where the sum runs over all partitions $\{V_1,\dots,V_r\}$ of $\{n+1,\dots,n+k\}$ and

$$A_{n,k}\big(x_1,\dots,x_n;y_1,\dots,y_k\big) := \Bigg(\prod_{\substack{1\le i\le n,\\ 1\le j\le k}}\big(1+f(x_i,y_j)\big)-1\Bigg)\varphi_k^{\mathsf{T}}\big(y_1,\dots,y_k\big).$$

Assume for a moment that

$$\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\big|A_{n,k}(\boldsymbol{x};\boldsymbol{y})z(y_1)\cdots z(y_k)\big|\lambda^k(\mathrm{d}\boldsymbol{y})<\infty. \quad (1.16)$$

Then it follows by the exponential formula from combinatorics (see, e.g., [12]) that

$$1+\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\sum_{\{V_1,\dots,V_r\}}\prod_{\ell=1}^{r}A_{n,|V_\ell|}\big(\boldsymbol{x};\boldsymbol{y}_{V_\ell}\big)\lambda_z^k(\mathrm{d}\boldsymbol{y}) = \exp\Bigg(\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}A_{n,k}(\boldsymbol{x};\boldsymbol{y})\lambda_z^k(\mathrm{d}\boldsymbol{y})\Bigg)$$

and the equality from the proposition follows from (1.15). Thus it remains to check the absolute convergence (1.16).

Therefore, we first notice that by the sign-flipped Kirkwood-Salsburg equations — satisfied by [a, Proposition 3.6] — the following holds for every $x_0\in\mathbb{X}$:

$$\begin{aligned}\tilde\rho_1(x_0;z) &= z(x_0)\left(1+\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\prod_{j=1}^{k}|f(x_0,y_j)|\,\tilde\rho_k(y_1,...,y_k;z)d\lambda^k(\mathrm{d}\boldsymbol{y})\right)\\ &\ge z(x_0)\int_{\mathbb{X}}|f(x_0,y)|\,\tilde\rho_1(y;z)d\lambda(y),\end{aligned}$$

in particular, that last integral is absolutely convergent:

$$\forall x_0\in\mathbb{X}:\quad \int_{\mathbb{X}}|f(x_0,y)|\,\tilde\rho_1(y;z)d\lambda(y)<\infty. \quad (1.17)$$

Standard arguments (see, e.g., [42, Eq. (10)]) yield the bound

$$\Big|\prod_{i=1}^{n}\prod_{j=1}^{k}\big(1+f(x_i,y_j)\big)-1\Big| \le \sum_{i=1}^{n}\sum_{j=1}^{k}|f(x_i,y_j)|,$$

which yields

$$\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\big|A_{n,k}(\boldsymbol{x};\boldsymbol{y})z(y_1)\cdots z(y_k)\big|\lambda^k(\mathrm{d}\boldsymbol{y}) \le \sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\sum_{i=1}^{n}\sum_{j=1}^{k}|f(x_i,y_j)||\varphi_k^{\mathsf{T}}(\boldsymbol{y})|\lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

Finally, interchanging the order of summation and integration on the right-hand side of the last inequality and using absolute convergence (1.17), we obtain (1.16):

$$\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\sum_{i=1}^{n}k|f(x_i,y_1)|\big|\varphi_k^{\mathsf{T}}(\boldsymbol{y})\big|\lambda_z^k(\mathrm{d}\boldsymbol{y})=\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\int_{\mathbb{X}^k}\sum_{i=1}^{n}|f(x_i,y_1)|\big|\varphi_k^{\mathsf{T}}(\boldsymbol{y})\big|\lambda_z^k(\mathrm{d}\boldsymbol{y})$$
$$=\sum_{i=1}^{n}\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\int_{\mathbb{X}^k}|f(x_i,y_1)|\big|\varphi_k^{\mathsf{T}}(\boldsymbol{y})\big|\lambda_z^k(\mathrm{d}\boldsymbol{y})=\sum_{i=1}^{n}\int_{\mathbb{X}}|f(x_i,y_1)|\tilde{\rho}_1(y_1;z)d\lambda(y_1)<\infty.$$

This concludes the proof of the proposition. □

1.5.3 Limitations of cluster expansion

We conclude our discussion of cluster expansion with an observation about the general limitations of the approach. Notice that the cluster expansion is a low density expansion and the domain in which the correlation functions are analytic can be in general much larger than the domain of convergence of the activity expansions — see, e.g., [24] where we have a one-dimensional system of non-overlapping rods such that the convergence of the activity expansion for the pressure implies that the activities decay exponentially in the length of the rods, but the pressure is analytic even for exponentially growing activities.

Historically, the convergence of the cluster expansion is often used to argue uniqueness of Gibbs measures, however, the convergence radius of the activity expansion is typically much lower than the critical activity corresponding to the uniqueness phase transition and other methods (e.g., based on disagreement percolation [2,22] or Dobrushin uniqueness [23]) provide far superior bounds for that critical activity.

Moreover, graphical expansions of the free energy or correlation functions in other parameters are possible, e.g., virial expansions in the density of the Gibbs processes [27]. The coefficients are given in terms of irreducible (doubly connected) graphs. Although, classically, the convergence of the virial expansion is proven via inversion formulas starting from an activity expansion (so that the limitations are inherited from cluster expansion methods) it was conjectured that the virial expansions have a larger radius of convergence than activity expansions (at least in the case of repulsive interactions). In particular models, this conjecture was confirmed (see, e.g., [24, 26, 33]).

Notice that, in the case of repulsive interactions, our main result does not only provide sufficient conditions, but also necessary ones. We show an exemplary result that demonstrates the limitations of the cluster expansion approach in a very classical setup. The result is based on a necessary condition that stems from [a, Theorem 2.7] — which is a minor modification of our main result Theorem 2.1 formulated for subset polymer systems. First, we would like to introduce this classical setup:

Let $\mathbb{X}$ consist of all finite, non-empty subsets of $\mathbb{Z}^d$, let $\mathcal{X}$ be the power set of $\mathbb{X}$ and let the reference measure λ be given by the counting measure on $\mathcal{X}$. We define the hard-core compatibility interactions by setting

$$f(X,Y):=-\mathbb{1}_{\{X\cap Y\neq\varnothing\}},\quad X,Y\in\mathbb{X}.$$

For simplicity, we want to consider a single-type polymer system here, i.e., we assume that there exists a finite non-empty set $S \subset \mathbb{Z}^d$ and a scalar $z > 0$ such that

$$z(X) = \begin{cases} z, & X \text{ is a translate of } S, \\ 0, & \text{otherwise.} \end{cases} \tag{1.18}$$

The elements of $\mathbb{X}$ are called polymers. With respect to the Gibbs measure, configurations with overlapping polymers are prohibited and the polymers that appear in the configuration have the same shape and size. To denote the size of those polymers, we set $V := |S|$.

Lemma 1.11 (Limitations of cluster expansions; discrete case)**.** *In the homogeneous case of (one-type) subset polymers, the convergence radius of the activity expansions $\tilde{\rho}_n(z)$ is bounded by the inverse of the polymer size V.*

Proof. A necessary condition for the absolute convergence of $\tilde{\rho}_n(z)$ (see [a, Theorem 2.7] and choose $D' := \varnothing$) is that for every $x \in \mathbb{Z}^d$ there exists a measurable function $a : \mathbb{X} \to [0, \infty)$ with

$$z \sum_{Y \in \mathbb{X}, Y \ni x} \mathrm{e}^{a(Y)} \leq \mathrm{e}^{a(\{x\})} - 1.$$

By the proof of [a, Theorem 2.7], we can choose a to satisfy the property $a(D_1) \leq a(D_2)$ for $D_1 \subset D_2$. Therefore, we get

$$\begin{aligned} z \leq & \frac{\mathrm{e}^{a(\{x\})}}{\sum_{Y \in \mathbb{X}, Y \ni x} \mathrm{e}^{a(Y)}} - \frac{1}{\sum_{Y \in \mathbb{X}, Y \ni x} \mathrm{e}^{a(Y)}} \\ \leq & \frac{1}{\sum_{Y \in \mathbb{X}, Y \ni x} \mathrm{e}^{a(Y) - a(\{x\})}} \leq \frac{1}{|\{Y \in \mathbb{X} | x \in Y\}|} = \frac{1}{V}. \end{aligned}$$

□

Remark 1.8. Analogous results can be formulated for non-overlapping objects of a single type in continuous setups, consider, e.g., the hard sphere model. That is, let $\mathbb{X}$ be given by $\mathbb{R}^d$, let λ be the Lebesgue measure on $\mathbb{R}^d$ and, for a positive radius $R > 0$, define the pair potential v on $R^d \times \mathbb{R}^d$ by

$$v(x, y) = \begin{cases} 0, & \text{if } ||x - y|| > 2R, \\ \infty, & \text{if } ||x - y|| \leq 2R. \end{cases}$$

In this setup, a similar argument yields $1/|B_R(0)|$ as an upper bound for the convergence radius of the activity expansions $\tilde{\rho}_n(z)$, where $|B_R(0)|$ denotes the (Lebesgue) volume of the open ball of radius R around 0.

Remark 1.9. Notice that the classic Kotecký-Preiss condition yields the bound $\frac{1}{V\mathrm{e}}$ for the convergence radius of the cluster expansion, so that the theoretical improvement possible past the classic result is limited to a mere multiplicative factor e.

1.6 Generating functions for generalized Catalan numbers

Now, let us briefly discuss the setup and the results from our second contribution [b]. In [24], a one-dimensional system of non-overlapping rods on $\mathbb{Z}$ is considered and the activity expansion for the associated pressure is studied. Under the assumption that all the rods have the same length $k \geq 2$ and activity z, the convergence of the activity expansion follows from the convergence of the power series $F_k(z)$ satisfying the fixed-point equation

$$\mathrm{e}^{F_k(z)} = 1 + z\mathrm{e}^{kF_k(z)} \tag{1.19}$$

on the level of formal power series. Naturally, the following question arises: What are the coefficients of $F_k(z)$ and how can they be interpreted combinatorially?

It is not hard to see that the formal power series $G_k(z) = \mathrm{e}^{F_k(z)}$ is given by

$$G_k(z) := 1 + \sum_{n\geq 1} \frac{z^n}{n}\binom{kn}{n-1},$$

see, e.g., [32]. One way to interpret the fixed-point equation (1.19) is to consider so-called k-ary rooted plane trees: Informally speaking, those are tree-graphs such that every vertex has exactly 0 children (i.e., it is a leaf) or k children ordered from left to right (see Figure 4). One can relate (1.19) to the following structural property of k-ary rooted plane trees: Such a tree is either empty or it decomposes into the root and a sequence of k k-ary rooted plane trees (this recursive relation can be taken as a formal definition [16]). Therefore, by standard combinatorial arguments (see, e.g., [12]), the generating function for k-ary rooted plane trees satisfies (1.19). And, for every $n \in \mathbb{N}$, the n-th coefficient of G_k enumerates k-ary rooted plane trees on n vertices: Indeed, the coefficients

$$c_k^{(0)} := 1, \quad c_k^{(n)} := \frac{1}{n}\binom{kn}{n-1}, \quad n \in \mathbb{N}$$

of G_k are called k^{th} generalized Catalan numbers (or Fuss-Catalan numbers, see [32]). For

$\frac{1}{n}\binom{kn}{n-1}$	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8	n=9	...
k=2	1	2	5	14	42	132	429	1430	4862	...
k=3	1	3	12	55	273	1428	7752	43263	246675	...
k=5	1	5	35	285	2530	23751	231880	2330445	23950355	...

Table 1: Here we provide the first few (generalized) Catalan numbers for $k = 2, 3, 5$.

$k = 2$, we recover the prominent Catalan numbers

$$c_2^{(0)} = 1, \quad c_2^{(n)} = \frac{2n!}{(n+1)!n!}, \quad n \in \mathbb{N},$$

that constitute one of the most studied sequences of natural numbers in combinatorics, admitting countless significant interpretations relevant in enumeration problems. Among the combinatorial structures enumerated by Catalan numbers (over 200 are listed in [40]) are binary plane trees. This interpretation generalizes to k^{th} generalized Catalan numbers, in the sense that they were shown to enumerate k-ary rooted plane trees [21]. This can be proven by using the recursive relation

$$c_k^{(0)} = 1, \quad c_k^{(n+1)} = \sum_{n=\ell_1+\ldots+\ell_j} \prod_{i=1}^{j} c_k^{(\ell_i)}, \quad n \in \mathbb{N},$$

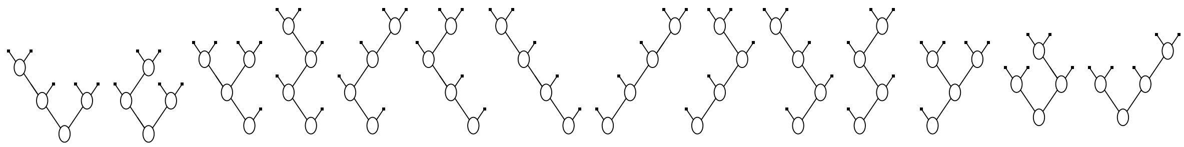

Figure 3: There are $c_2^{(4)} = 14$ unlabeled binary rooted plane trees with 4 (non-leaf) vertices.

which corresponds to the recursive structure of the k-ary rooted plane trees described above.

The question posed in [24] that we answer in [b] can now be reformulated as follows: For which labeled tree-like structures is $F_k(z)$ the exponential generating function and what are the underlying constructions (on the level of graphs) connecting those structures to k-ary rooted plane trees? Notice that we need to consider labeled structures here, $F_k(z) = \log G_k(z)$ can not be interpreted as an ordinary generating function for a combinatorial species, but only as an exponential generating function, since the coefficients of $F_k(z)$ do not need to be natural numbers.

We provide such combinatorial interpretations in terms of tree-like structures in [b, Theorem 3.5] and [b, Theorem 3.7]. The structures underlying the interpretation of $F_k(z)$ from the latter result are called cycle-rooted trees and are inherently of cyclic nature, which can be understood as follows: Informally speaking, a cycle-rooted tree is obtained from a k-ary rooted plane tree by bending the right-most branch of the tree into a cycle — identifying the root and the right-most leaf of the tree. Our bijective result [b, Lemma 3.4] that identifies the trees with sets of cycle-rooted trees can be viewed as a generalization of the well-known decomposition of permutations into cycles.

An alternative interpretation, using the fact that the Catalan generating function G_k can also be viewed as the generating functions for certain monotone lattice paths, the so-called k-good path [21], is stated in [b, Theorem 2.7] — it is based on our bijective result [b, Lemma 2.4] stating that a k-good path can be decomposed into a set of cyclic path-like structures we call lattice k-ornaments. Informally speaking, those structures are obtained by bending k-good paths into cycles — identifying the starting point and the end-point of the path. This combinatorial interpretation allows for a particularly simple enumerative proof to determine explicit expressions for the coefficients of $F_k(z)$ (as well as for the coefficients of the higher powers $(F_k(z))^a$, $a \geq 2$). Notice that closed expressions for the coefficients of $F_k(z)$ are

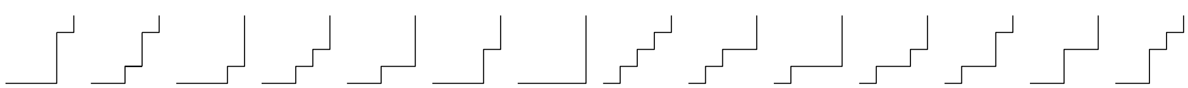

Figure 4: There are $c_2^{(4)} = 14$ 2-good paths with 4 vertical steps.

known in the literature — the formula for $k = 2$ (that we prove in [b, Theorem 2.8]) was presented by Donald Knuth in his 2014 Christmas lecture. Soon formulas for the coefficients of $F_k(z)$ were proven for $k \geq 2$ (see [24]), and also for the coefficients of higher powers $(F_k(z))^a$ with $a \geq 2$ (see [6, 39]). All those proofs involve general algebraic inversion formulas like Lagrange inversion. Thus our results identifying those coefficients (see [b, Theorem 2.8 and Theorem 2.9]) can not be considered novel, but demonstrate the usefulness of our combinatorial interpretation of $F_k(z)$ that allows us to reduce the problem of identifying the coefficients of its powers to a simple enumeration problem.

In [b], all appearing power series are treated as formal power series and we do not provide statements about the convergence of the series. However, the domain of convergence of the power series $F_k(z)$ was characterized in [24].

1.7 The random connection model and the Ornstein-Zernike equation

Consider a GPP on $\mathbb{R}^d$ with repulsive pair interactions (i.e., $v \geq 0$ and $0 \leq f = \mathrm{e}^v - 1 \leq 1$). One can hope to gain knowledge about the uniqueness of the Gibbs measure by choosing a suitable notion of connectivity on the point configurations and studying the associated percolation phase transition. A recipe to obtain connectedness functions is given in [8] by physicists: On the level of graphical expansions, e.g., in the density of the GPP, one obtains connectedness functions from the correlation functions of the GPP by discarding certain graphs that contribute to the graphical expansions of the latter. To do so, one rewrites Mayer's f function as a certain sum of two functions, $f = f^+ + f^\star$, splitting every edge into two possible edges — one weighted with f^+, the other with $f^\star$. Naturally, this leads to expansions of the correlation functions in terms of weighted graphs with the two types of edges. To obtain the expansions for connectedness functions, graphs that satisfy certain connectivity assumptions with respect to f^+-edges are kept, the rest is discarded. While [8] provides elegant power series expansions of the connectedness functions, the convergence of those expansions is not treated in the physics literature and mathematically rigorous analysis is exceedingly rare (see [31]).

In [b], we consider a simplification, where the starting point is the PPP of intensity $\lambda \geq 0$ in $\mathbb{R}^d$. Consider the random connection model (RCM), informally described as follows: The vertex set is given by the points of the PPP (and possibly a finite number of deterministic points) and edges between the vertices are drawn independently with probabilities given by a radially symmetric connection function $\varphi : \mathbb{R}^d \to [0,1]$. The pair connectedness function $\tau_\lambda : \mathbb{R}^d \times \mathbb{R}^d \to [0,1]$ is given by the probabilities $\tau_\lambda(x_1, x_2)$ that two deterministic points $x_1, x_2 \in \mathbb{R}^d$ (a priori fixed as vertices) are connected in the resulting random graph. For a rigorous introduction of the RCM, see [19, 31].

The pair connectedness function admits a graphical expansion in terms of graphs with two types of edges, $(+)$-edges weighted by φ and $(-)$-edges weighted by $-\varphi$; we refer to those as $(\pm)$-graphs. This corresponds to the recipe from [8], where one decomposes Mayer's f function — notice that $f \equiv 0$ for the PPP — as $f = \varphi + (-\varphi)$. The resulting notion of connectivity then coincides with the one in the RCM. Notice that the negative weights $-\varphi$ can be interpreted in terms of Mayer's f function of an associated GPP with repulsive pair interactions (in particular, if we draw edges deterministically between sufficiently close points, $-\phi$ can be interpreted as Mayer's f function for a hard spheres model).

The expansion of $\tau_\lambda(x_1, x_2)$ is then given by

$$\tau_\lambda(x_1, x_2) = \sum_{n \geq 0} \frac{\lambda^n}{n!} \int \sum_{\substack{G \in \mathcal{C}^\pm(\vec{x}_{[n+2]}): \\ x_1 \overset{+}{\longleftrightarrow} x_2}} \mathbf{w}^\pm(G) \mathrm{d}\vec{x}_{[3,n+2]}, \quad x_1, x_2 \in \mathbb{R}^d, \tag{1.20}$$

where the graphs we sum over satisfy the following assumptions: Those are $(\pm)$-graphs on

$\{x_1, \ldots, x_{n+2}\}$ that are connected (in terms of paths consisting of possibly both types of edges) and, additionally, x_1, x_2 are connected by a path consisting of only (+)-edges.

The Ornstein-Zernike equation (OZE) is an integral equation first derived in [37]. In statistical mechanics, it initially related — essentially via a convolution formula — the correlation functions to the direct correlation functions. It was pointed out by Hill in [20] that it was possible to use the same relations for connectedness functions in percolation as well. The goal is to find a solution g_λ, called the direct-connectedness function, to the Ornstein-Zernike equation

$$\tau_\lambda(x, y) = g_\lambda(x, y) + \lambda \int_{\mathbb{R}^d} g_\lambda(x, z)\tau_\lambda(z, y)\mathrm{d}z, \quad x, y \in \mathbb{R}^d. \tag{1.21}$$

The existence of a unique integrable and essentially bounded function $g_\lambda : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ solving (1.21) was proven in [31].

In (1.21), both the functions τ_λ and g_λ are unknown, moreover, unlike the pair connectedness function τ_λ, the direct-connectedness function g_λ cannot be interpreted probabilistically. Usually, a complementary equation — a so-called closure relation — is provided to obtain an approximation of g_λ as a solution of a modified integral equation (a classic example is the Percus-Yevick approximation, see [41]). In [b], we do not introduce a closure relation, but consider a graphical expansion as a definition — showing that it converges and solves the OZE in a certain low-intensity regime. From [8], we know the physicists expansion of g_λ (stated there without proof) given by

$$g_\lambda(x_1, x_2) = \sum_{n\geq 0} \frac{\lambda^n}{n!} \int \sum_{\substack{G \in \mathcal{D}^{\pm}_{x_1,x_2}(\vec{x}_{[n+2]}): \\ x_1 \overset{+}{\longleftrightarrow} x_2}} \mathbf{w}^{\pm}(G)\mathrm{d}\vec{x}_{[3,n+2]}. \tag{1.22}$$

The difference to the expansion (1.20) of $\tau_\lambda(x_1, x_2)$ is that on the graphs in (1.22) one imposes an additional assumption — the existence of two vertex-disjoint ($\pm$)-paths between x_1 and x_2 (see Figure 5). In other words, the expansion (1.22) of $g_\lambda(x_1, x_2)$ is obtained from the expansion (1.20) of τ_λ by discarding those ($\pm$)-graphs that possess vertices pivotal for the x_1-x_2-connection.

In [c, Section 6.2], we show that the convergence of the expansion (1.22) follows from [31], but we do not provide quantitative bounds for the domain of convergence of this expansions. Instead, the idea behind the main result [c, Theorem 1.1] is to perform a resummation in (1.22) and show that the resulting series converges and solves the OZE in a certain domain conjectured to be bigger than the domain of convergence of (1.22). The graphical expansion obtained by the resummation is not a power series — the direct-connectedness function can be viewed to be partially expanded in the intensity of the underlying PPP. The rough idea behind the resummation is to split every graph contributing to the expansion into a core graph and a shell graph — so that the direct-connectedness function decomposes into a core function and a shell function (see [c, Definition 4.2]). We then provide probabilistic bounds for both the core (in terms of certain “connection probabilities”, see [c, Proposition 4.1]) and the shell (in terms of certain “disconnection probabilities”, see [c, Eq. (4.25)]).

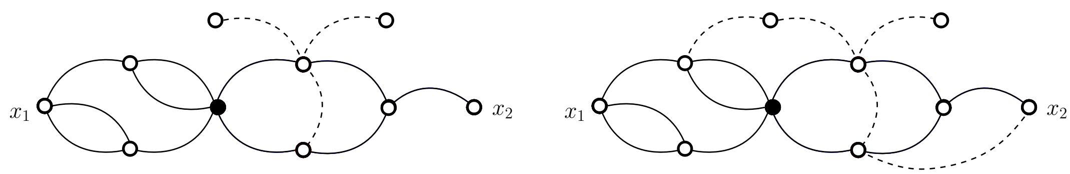

Figure 5: We see two graph contributing to the expansion (1.20) of $\tau_\lambda(x_1, x_2)$. The $(+)$-edges are depicted as solid lines and the $(-)$-edges as dashed lines. Notice that both graphs are connected (with respect to arbitrary edges) and in both graphs there is a path consisting of solely $(+)$-edges connecting the vertices x_1 and x_2. The right graph does not possess a vertex pivotal to the x_1-x_2-connection (with respect to arbitrary edges) — thus it also appears in the expansion (1.22) of $g_\lambda(x_1, x_2)$. But in the left graph the solid black vertex is pivotal for the x_1-x_2 connection — thus it does not appear in (1.22).

Bibliography: Own Contributions

[a] S. Jansen and L. Kolesnikov. Cluster expansions: Necessary and sufficient convergence conditions. *J. Stat. Phys.*, 189(33):1–47, 2022.

[b] S. Jansen and L. Kolesnikov. Logarithms of Catalan generating functions: A combinatorial approach. arXiv e-prints, 2022.

[c] S. Jansen, L. Kolesnikov, and K. Matzke. The direct-connectedness function in the random connection model. *Adv. in Appl. Probab.*, 1–44, 2022.

Bibliography: Additional References

[1] G. A. Battle. A new combinatoric estimate for cluster expansions. *Commun. Math. Phys.*, 94:133–139, 1984.

[2] S. Betsch and G. Last. On the uniqueness of Gibbs distributions with a non-negative and subcritical pair potential. 2021.

[3] R. Bissacot, R. Fernández, and A. Procacci. On the convergence of cluster expansions for polymer gases. *J. Stat. Phys.*, 139(4):598–617, 2010.

[4] D. C. Brydges. A short course on cluster expansions. In *Phénomènes critiques, systèmes aléatoires, théories de jauge, Part I, II (Les Houches, 1984)*, pages 129–183. North-Holland, Amsterdam, 1986.

[5] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, third edition, 2013.

[6] W. Chu. Logarithms of a binomial series: Extension of a series of Knuth. *Math. Commun.*, (24):83–90, 2019.

[7] D. Conache, A. Daletskii, Y. Kondratiev, and T. Pasurek. Gibbs states of continuum particle systems with unbounded spins: Existence and uniqueness. *J. Math. Phys.*, 59(1):013507, 2018.

[8] A. Coniglio, U. DeAngelis, and A. Forlani. Pair connectedness and cluster size. *J. Phys. A*, 10(7):1123–1139, Jul 1977.

[9] D. Dereudre. Introduction to the theory of gibbs point processes. In *Stochastic geometry*, pages 181–229. Springer, 2019.

[10] R. L. Dobrushin. Estimates of semiinvariants for the Ising model at low temperatures. *Topics in Statistics and Theoretical Physics, Amer. Math. Soc. Transl.*, 177(2):59–81, 1996.

[11] W. G. Faris. A connected graph identity and convergence of cluster expansions. *J. Math. Phys.*, 49(11):113302, 2008.

[12] W. G. Faris. Combinatorics and cluster expansions. *Prob. Surveys*, 7:157–206, 2010.

[13] R. Fernández and A. Procacci. Cluster expansion for abstract polymer models. New bounds from an old approach. *Comm. Math. Phys.*, 274(1):123–140, 2007.

[14] P. M. S. Fialho. Abstract polymer gas: A simple inductive proof of the Fernández-Procacci criterion. *J. Stat. Phys.*, 178:1354–1361, 2020.

[15] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems.* Cambridge University Press, Cambridge, 2018.

[16] R. Garnier and J. Taylor. *Discrete Mathematics: Proofs, Structures and Applications, Third Edition.* Taylor & Francis, 2009.

[17] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[18] C. Gruber and H. Kunz. General properties of polymer systems. *Comm. Math. Phys.*, 22(2):133–161, 1971.

[19] M. Heydenreich, R. van der Hofstad, G. Last, and K. Matzke. Lace Expansion and Mean-Field Behavior for the Random Connection Model. page arXiv:1908.11356, 2019.

[20] T. L. Hill. Molecular Clusters in Imperfect Gases. *J. Chem. Phys.*, 23(4):617–622, 1955.

[21] P. Hilton and J. Pedersen. Catalan Numbers, Their Generalization, and Their Uses. *Math. Intell.*, 13:64–75, 1991.

[22] C. Hofer-Temmel and P. Houdebert. Disagreement percolation for Gibbs ball models. *Stoch. Process. Appl.*, 129(10):3922–3940, 2019.

[23] P. Houdebert and A. Zass. An explicit continuum Dobrushin uniqueness criterion for Gibbs point processes with non-negative pair potentials. *J. Appl. Probab.*, 59(2):541–555, 2022.

[24] S. Jansen. Cluster and virial expansions for the multi-species Tonks gas. *J. Stat. Phys.*, 161(5):1299–1323, 2015.

[25] S. Jansen. Cluster expansions for Gibbs point processes. *Adv. Appl. Probab.*, 51:1129–1178, 2019.

[26] S. Jansen. Thermodynamics of a hierarchical mixture of cubes. *J. Stat. Phys.*, 179:309–340, 2020.

[27] S. Jansen, T. Kuna, and D. Tsagkarogiannis. Virial inversion and density functionals. *J. Funct. Anal.*, 284, 2023.

[28] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Comm. Math. Phys.*, 103(3):491–498, 1986.

[29] T. Kuna and D. Tsagkarogiannis. Convergence of density expansions of correlation functions and the Ornstein-Zernike equation. *Ann. Henri Poincaré*, 19(4):1115–1150, 2018.

[30] G. Last and M. Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[31] G. Last and S. Ziesche. On the Ornstein–Zernike equation for stationary cluster processes and the random connection model. *Adv. Appl. Probab.*, 49(4):1260–1287, 2017.

[32] C.-H. Lin. Some combinatorial interpretations and applications of Fuss-Catalan numbers. *ISRN Discrete Math.*, 2011, 01 2011.

[33] D. H. U. Marchetti. The virial series for a gas of particles with uniformly repulsive pairwise interaction and its relation with the approach to the mean field. *Braz. J. Probab. Stat. 29*, 29(2):502–539, 2015.

[34] J. E. Mayer and M. G. Mayer. *Statistical Mechanics.* John Wiley and Sons, Inc. London: Chapman and Hall, limited edition, 1940.

[35] R. Meester and R. Roy. *Continuum percolation*, volume 119 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1996.

[36] J. Møller and R. P. Waagepetersen. *Statistical inference and simulation for spatial point processes*, volume 100 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[37] L. S. Ornstein and F. Zernike. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci. (Amsterdam)*, 17:793–806, 1914.

[38] O. Penrose. Convergence of fugacity expansions for classical systems. In *Statistical Mechanics, Foundations and Applications.* Benjamin, New York, 1967.

[39] H. Prodinger. Logarithms of a binomial series: A Stirling number approach. *Ars Math. Contemp.*, 17:271–275, 10 2019.

[40] R. P. Stanley. *Catalan Numbers.* Cambridge University Press, 2015.

[41] S. Torquato. *Random heterogeneous materials. Microstructure and macroscopic properties*, volume 16 of *Interdisciplinary Applied Mathematics.* Springer-Verlag, New York, 2002.

[42] D. Ueltschi. Cluster expansions and correlation functions. *Mosc. Math. J.*, 4(2):511–522, 2004.

2 Cluster expansions: Necessary and sufficient convergence conditions

This section contains the published version of:
S. Jansen and L. Kolesnikov. Cluster expansions: Necessary and sufficient convergence conditions. *J. Stat. Phys.*, 189(33):1–47, 2022.

Personal contribution:
This article is a collaborative effort based on discussions with my supervisor, Sabine Jansen. I developed and implemented many ideas incorporated in our main results, with occasional input from Sabine. I also took a leading role in drafting and writing the article.

Journal of Statistical Physics (2022) 189:33
https://doi.org/10.1007/s10955-022-02992-6

Cluster Expansions: Necessary and Sufficient Convergence Conditions

Sabine Jansen[1,2] · Leonid Kolesnikov[1]

Received: 16 January 2022 / Accepted: 12 September 2022 / Published online: 24 September 2022

Abstract

We prove a new convergence condition for the activity expansion of correlation functions in equilibrium statistical mechanics with possibly negative pair potentials. For non-negative pair potentials, the criterion is an if and only if condition. The condition is formulated with a sign-flipped Kirkwood–Salsburg operator and known conditions such as Kotecký–Preiss and Fernández–Procacci are easily recovered. In addition, we deduce new sufficient convergence conditions for hard-core systems in $\mathbb{R}^d$ and $\mathbb{Z}^d$ as well as for abstract polymer systems. The latter improves on the Fernández–Procacci criterion.

Keywords Cluster expansions · Correlation functions · Kirkwood–Salsburg equations · Combinatorics of connected graphs · Abstract polymer models · Hard-core germ–grain models · Subset polymers · Hard spheres

Mathematics Subject Classification 82B05 · 82B21

1 Introduction

Since its introduction by Mayer in the early 40s, the method of cluster expansions was—and remains—a very important tool in equilibrium statistical mechanics. A classical application yields the analyticity of the logarithm of the partition function for a physical system at equilibrium by deriving a Taylor expansion in the activity or density parameter around zero. Such results can be quite useful, for example, in the study of phase transitions or the decay of correlations, for a vast class of models.

In 1971, Gruber and Kunz introduced in their seminal paper [12] systems of non-overlapping geometric objects—referred to as *polymers*—given by subsets of a lattice. They

Communicated by Yvan Velenik.

Leonid Kolesnikov
kolesni@math.lmu.de

1 Mathematisches Institut, Ludwig-Maximilians-Universität, Theresienstr. 39, 80333 Munich, Germany

2 Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, 80799 Munich, Germany

presented a rigorous mathematical formalism in order to provide convergent cluster expansions for this model. Instead of the logarithm of the partition function, they considered the correlation functions of the system and derived convergent activity expansions by using a system of integral equations, the so-called Kirkwood–Salsburg equations, and solving the corresponding fixed point equation on a suitable Banach space. However, in the following years less analytical appoaches were favoured by researchers: Combinatorial proofs such as in [3], relying on tree-graph identities [23], and inductive proofs following the idea by Kotecký and Preiss [18] and its development in [4] by Dobrushin. The inductive method was presented in the more general setup of abstract polymers (where the underlying space is not necessarily a lattice, nor are the polymers necessarily given by geometric objects). Notice that abstract polymer models are universal in the sense that a large class of classical models can be represented as polymer models due to the combinatorial structure of the corresponding partition functions (see, e.g., [11] for an application to the Ising model). Moreover, an interesting connection with probability theory was pointed out by Scott and Sokal in [29]: Convergence of cluster expansions in abstract polymer models is related to the Lovász Local Lemma—better sufficient conditions can provide refinements of the latter (see, e.g., [2]).

In 2008, Fernández and Procacci proved a new sufficient criterion in the setup of abstract polymers improving on the result by Kotecký and Preiss. The initial proof [8] relies on combinatorial arguments, an alternative proof via an induction à la Dobrushin [10] appeared recently (finally, in this paper we provide an analytical proof in the spirit of Gruber–Kunz).

Overall, in the last two decades, a notable effort was made to generalize classical sufficient conditions in the abstract polymer setup (including the condition by Fernández and Procacci) to hold in continuous spaces and for systems with soft-core (or even more general) interactions, see [5, 14, 22, 24, 32].

We want to go further by employing a Kirkwood–Salsburg approach in the rather general setup of Gibbs point processes (or, in terms of statistical mechanics, grand-canonical Gibbs measures) defined via pairwise interactions. It is well-known that—under mild additional moment conditions, which are automatically satisfied for non-negative pair potentials—there is a one-to-one correspondence between the set of those Gibbs measures and the associated families of correlation functions (also known as *factorial moment densities*). In the special case of a discrete space and hard-core interactions, the value of the n-point correlation function is given simply by the probability to see n particles at the prescribed positions in the random configuration of particles. The correlation functions can be expanded as power series in the activity parameter z, i.e., in the intensity of the underlying Poisson process. We denote the Taylor expansion for the n-point correlation function in z around zero by ρ_n and write $\boldsymbol{\rho}$ for the family of those expansions. In general, the series $\boldsymbol{\rho}$ need not to be convergent at all; we are, however, interested in conditions which ensure pointwise convergence (towards the correlation functions). Furthermore, we want to consider the more general case where the underlying Poisson point process is inhomogeneous, i.e., where a different intensity value may be assigned to every point in the space, the activity z is a function and the expansions $\boldsymbol{\rho}$ are multivariate power series in z. For a rigorous introduction of Gibbs point processes and corresponding correlation functions, see [14], but notice that here we do not assume the interaction potential to be non-negative (unless explicitly stated).

The starting point of the paper and the central quantity to investigate are the activity expansions $\boldsymbol{\rho}$ which we consider independently of their interpretation in term of the correlation functions. Let us outline the main ideas present in the paper. The coefficients of the multivariate power series $\boldsymbol{\rho}$ are defined in terms of a certain family of rooted graphs to which we refer as *multi-rooted graphs* (see [14, 30]). Using the terminology from [6], the activity expansions $\boldsymbol{\rho}$ are given by the exponential generating functions of the coloured weighted combinatorial species of multi-rooted graphs with a fixed set of roots. The set of all multi-rooted graphs has an essential structural property—it is invariant under the operation of removal of a root. Taking a multi-rooted graph and removing an arbitrary root (as well as all edges incident to it), one gets again a multi-rooted graph on a smaller vertex set, where every neighbour of the removed root becomes a root vertex itself. The weight of the original graph is equal to the weight of the resulting graph times the weight of the edges removed. The corresponding property of the generating functions is expressed by the Kirkwood–Salsburg equations. Every possible rule for the choice of the root to remove induces a different combinatorial operation and therefore a different system of Kirkwood–Salsburg equations for the generating functions.

In this work we provide a condition for absolute convergence of the activity expansions $\boldsymbol{\rho}$ in terms of the existence of a measurable function solving a system of Kirkwood–Salsburg type inequalities (in the case of repulsive interactions, that condition is also a necessary one). Our main result, Theorem 2.1, is inspired by [1]; it is a slightly modified, strongly generalized version of Claim 1 therein. The goal, however, is not only to obtain abstract conditions which are both necessary and sufficient for convergence of the cluster expansions—but also to demonstrate how these characterizations provide a universal approach to prove model-specific sufficient conditions on different levels of generality, both in discrete and continuous setups with repulsive interactions. A two-lane mechanism arises: On the one hand, for a candidate family of ansatz functions $\boldsymbol{\xi}$ (given, for example, as approximations of $\boldsymbol{\rho}$) one can search for conditions that ensure that these functions $\boldsymbol{\xi}$ satisfy the Kirkwood–Salsburg inequalities; on the other hand, given candidate sufficient conditions, one can construct a suitable family of ansatz functions $\boldsymbol{\xi}$ tailored to satisfy the Kirkwood–Salsburg inequalities under these conditions.

This approach provides a unifying framework for the known conditions, but it also allows to prove stronger results. To emphasize this possibility, we derive a new sufficient condition for absolute convergence of the activity expansions $\boldsymbol{\rho}$ in the setup of abstract polymers. In that general setup, our condition improves on any known condition that we are aware of.

A more detailed outline of the main ideas intoduced above can be found in [16] (for the case of non-negative pairwise interactions and without rigorous proofs).

In the further course of the paper, we investigate two particular hard-core setups as examples—the subset polymers in $\mathbb{Z}^d$ and hard objects in $\mathbb{R}^d$. There, the sets of roots of the multi-rooted graphs correspond to configurations of geometric objects. By breaking the geometric objects into smaller "pieces" to which we refer as *snippets*, we can identify these configurations with configurations of snippets (e.g., in the case of subset polymers we can identify a configuration of polymers with the disjoint union of monomers covering this configuration). Picking a root of a multi-rooted graph—the combinatorial operation underlying the Kirkwood–Salsburg equations—corresponds to picking a snippet. Different rules to pick a snippet in general give rise to characterizations of absolute convergence in terms of different Kirkwood–Salsburg inequalities. This way the latter can be tailored to a candidate sufficient condition. Thus different sufficient conditions can be derived by playing both with the choice of different systems of Kirkwood–Salsburg inequalities and the choice of different ansatz functions satisfying these inequalities. We illustrate this mechanism by deriving some sufficient conditions for a class of hard-core interaction models, in particular for multi-type systems of hard spheres in $\mathbb{R}^d$.

The paper is organized as follows: In Sect. 2.1 we introduce the basic notation and present the general framework. Furthermore, in Theorem 2.1 we state our main result, a characterization of the domain of absolute convergence for the activity expansions $\boldsymbol{\rho}$, and use it to recreate the classical sufficient conditions by Kotecký and Preiss as well as the sufficient conditions by Fernández and Procacci in a rather general setup (see Corollaries 2.7 and 2.9, respectively). In Sect. 2.2, the same approach is used to prove a new, improved sufficient condition in the setup of abstract polymers (Proposition 2.4). The proof of the proposition relies on an auxiliary result (Lemma 2.5) which is proved in Appendix A. In the Sects. 2.3 and 2.4 we consider the special case of hard-core interactions. Both in the continuum (Sect. 2.3) and in the discrete setup (Sect. 2.4), we provide model-specific characterizations of the convergence domain, stated in Theorems 2.6 and 2.7, respectively. As an immediate consequence of Theorem 2.7 we obtain an elementary proof of the well known Gruber–Kunz condition (Corollary 2.8). In Sect. 3, we present a forest-graph equality and other combinatorial results in order to prove Theorems 2.1, 2.6 and 2.7. Finally, in Sect. 4, Theorems 2.6 and 2.7 are used to obtain practitioner-type sufficient conditions for a class of hard-core interaction models, including new sufficient conditions for subset polymers in $\mathbb{Z}^d$ (Theorem 4.1) and hard objects in $\mathbb{R}^d$ (Theorems 4.2 and 4.4).

The reader interested primarily in the discrete setup of subset polymers is encouraged to jump directly to Subsect. 2.4, its main result being the characterization of the domain of convergence for the activity expansions $\boldsymbol{\rho}$ given by Theorem 2.7 (compare to Theorem 3.13). The main ideas behind the proof of Theorem 2.7 in Subsect. 3.4 and behind the application of Theorem 2.7 in Subsect. 4.1 can be transferred to the continuous setup as well.

2 Main Results

2.1 (Locally) Stable Pair Potentials

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space, λ a σ-finite reference measure, and v a pair potential, i.e., $v : \mathbb{X} \times \mathbb{X} \to \mathbb{R} \cup \{\infty\}$ is measurable and symmetric—in the sense that $v(x, y) = v(y, x)$ for any $x, y \in \mathbb{X}$. Corresponding to the potential v, *Mayer's f function* is given by

$$f(x, y) = \mathrm{e}^{-v(x,y)} - 1.$$

We call the pair potential v *stable* if there exists a measurable map $B : \mathbb{X} \to \mathbb{R}_+$ such that for any $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{X}$

$$\prod_{1 \le i < j \le n} (1 + f(x_i, x_j)) \le \mathrm{e}^{\sum_{k=1}^{n} B(x_k)} \tag{2.1}$$

holds; we call v *locally stable* or *Penrose stable* (due to O. Penrose, see [23]) if there exists a measurable map $C : \mathbb{X} \to \mathbb{R}_+$ such that for any $x_0 \in \mathbb{X}$, $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{X}$ satisfying $\prod_{1 \le i < j \le n} (1 + f(x_i, x_j)) \neq 0$

$$\prod_{i=1}^{n} (1 + f(x_0, x_i)) \le \mathrm{e}^{C(x_0)} \tag{2.2}$$

holds. Notice that every locally stable potential is stable and that every non-negative potential v is locally stable (with the choice $C \equiv 0$).

An *activity function* is a measurable map $z : \mathbb{X} \to \mathbb{R}$. Physically relevant activities are non-negative but for the purpose of studying the convergence of expansions it can be helpful to admit negative (or complex) activities as well. We define the (signed) measure λ_z on $\mathcal{X}$ by

$$\lambda_z(B) := \int_B z(x) \lambda(\mathrm{d}x), \quad B \in \mathcal{X}. \tag{2.3}$$

The weight of a graph G with vertex set $[n] = \{1, \dots, n\}$ and edge set $E(G)$ is

$$w(G; x_1, \dots, x_n) := \prod_{\{i,j\} \in E(G)} f(x_i, x_j).$$

Let $\mathcal{G}_n$ be the set of all graphs with vertex set $[n]$, $\mathcal{C}_n \subset \mathcal{G}_n$ the set of connected graphs and

$$\varphi_n^{\mathsf{T}}(x_1, \dots x_n) := \sum_{G \in \mathcal{C}_n} w(G; x_1, \dots, x_n)$$

the n-th Ursell function. For $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$, let $\mathcal{D}_{n,n+k} \subset \mathcal{G}_{n+k}$ be the collection of all graphs G such that every vertex $j \in \{n+1, \dots, n+k\}$ connects to at least one of the vertices $i \in \{1, \dots, n\}$. We may view the vertices $\{1, \dots, n\}$ as roots and call the graphs $G \in \mathcal{D}_{n,n+k}$ *multi-rooted graphs* or, following the footnote 53 in [30], *root-connected* graphs. Consider the functions

$$\psi_{n,n+k}(x_1, \dots, x_{n+k}) := \sum_{G \in \mathcal{D}_{n,n+k}} w(G; x_1, \dots, x_{n+k}).$$

For $n = 1$, the functions coincide with the standard Ursell functions, i.e., $\psi_{1,1+k} = \varphi_{1+k}^{\mathsf{T}}$. We are interested in the associated series

$$\rho_n(x_1, \dots, x_n; z) := \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k) z(x_1) \cdots z(x_n) \lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

The summand for $k = 0$ is to be read as $\psi_{n,n}(x_1, \dots, x_n) z(x_1) \cdots z(x_n)$. The series ρ_n corresponds to the n-point correlation function of a grand-canonical Gibbs measure [30, Eqs. (4–7)], see also [14]—it is the expansion of the correlation function in the activity z around 0.

We will say that the activity expansions $\boldsymbol{\rho}$ converge absolutely for a non-negative activity function z if

$$\sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} |\psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k)| z(x_1) \cdots z(x_n) \lambda_z^k(\mathrm{d}\boldsymbol{y}) < \infty$$

for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$.

Our main concern is to derive necessary and sufficient convergence conditions, but sometimes it is useful to view the series as purely formal; relevant background on formal power series whose variable is a measure (here $\lambda_z(\mathrm{d}x)$) is given in [17, Appendix A].

Next we introduce sign-flipped Kirkwood–Salsburg operators. A *selection rule* $s(\cdot)$ is a map from $P(\mathbb{X}) := \sqcup_{n=1}^{\infty} \mathbb{X}^n$ to $\mathbb{N}$ such that $s(x_1, \dots, x_n) \in \{1, \dots, n\}$ for all $(x_1, \dots, x_n) \in P(\mathbb{X})$. To lighten notation we write x_s rather than $x_{s(x_1,\dots,x_n)}$. Further let $(x_2', \dots, x_n')$ be the vector obtained from $(x_1, \dots, x_n)$ by deleting the entry x_s, leaving the order otherwise unchanged. For the simplest selection rule that picks the first entry $s = 1$, we have $x_i' = x_i$. The *sign-flipped Kirkwood–Salsburg operator* $\tilde{K}_z^s$ with selection rule $s(\cdot)$ acts on families $\boldsymbol{\xi} = (\xi_n)_{n \in \mathbb{N}}$ of measurable symmetric functions $\xi_n : \mathbb{X}^n \to \mathbb{R}_+$ as

$$\begin{aligned}(\tilde{K}_z^s \xi)_n(x_1, \dots, x_n) := z(x_s) \prod_{i=2}^{n} (1 + f(x_s, x_i')) \Big(&\mathbb{1}_{\{n \geq 2\}} \xi_{n-1}(x_2', \dots, x_n') \\ &+ \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \prod_{j=1}^{k} \big| f(x_s, y_j) \big| \, \xi_{n-1+k}(x_2', \dots, x_n', y_1, \dots, y_k) \lambda^k(\mathrm{d}\boldsymbol{y}) \Big),\end{aligned} \tag{2.4}$$

for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$. Here we allow the functions $(\tilde{K}_z^s \xi)_n$ to assume the value "∞". For non-negative potentials and on a suitably reduced domain, $\tilde{K}_z^s$ differs from the standard Kirkwood–Salsburg operator [26, Chapter 4.2] by a mere sign-flip: it has $|f(x_s, y_i)|$ instead of $f(x_s, y_i)$.

Theorem 2.1 *Let $z(\cdot)$ be a non-negative activity and $s(\cdot)$ any selection rule. Consider the following two conditions:*

(i) *There is a family $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$ of measurable symmetric functions $\xi_n : \mathbb{X}^n \to \mathbb{R}_+$ such that*

$$z(x_1)\delta_{n,1} + (\tilde{K}_z^s \boldsymbol{\xi})_n\,(x_1, \dots, x_n) \le \xi_n(x_1, \dots, x_n) \tag{2.5}$$

for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$.

(ii) *The series $\rho_n(x_1, \dots, x_n; z)$ converges absolutely, for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$.*

Condition (i) is sufficient for (ii) to hold; moreover, if (i) is satisfied, then

$$\sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \big|\psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k)\big| z(x_1)\cdots z(x_n)\lambda_z^k(\mathrm{d}\boldsymbol{y}) \le \xi_n(x_1, \dots, x_n) \tag{2.6}$$

on $\mathbb{X}^n$, for all $n \in \mathbb{N}$.

In addition, if we assume the pair potential to be non-negative, then (ii) implies (i) as well, so that the two conditions are equivalent in this case.

Remark 2.1 We formulate this theorem—as well as the following results—for non-negative activities, mainly for the purpose of notational convenience. Naturally, such conditions for absolute convergence can be formulated in the usual framework of complex analysis by exchanging complex activities z with $|z|$ in the convergence criteria.

We prove the theorem in Subsect. 3.2. The known sufficient convergence conditions of Kotecký–Preiss and Fernández–Procacci types are easily recovered from Theorem 2.1. We start with the Kotecký–Preiss type criterion [18], as extended to soft-core and continuum systems by Ueltschi in [32] (and to stable interactions by Ueltschi and Poghosyan in [24]).

Corollary 2.2 *Let z be a non-negative activity function and assume stable interactions in the sense of (2.1) for some $B \ge 0$. If there exists a measurable function $a : \mathbb{X} \to \mathbb{R}_+$ such that for all $x \in \mathbb{X}$*

$$\int_{\mathbb{X}} \big|f(x, y)\big| \mathrm{e}^{a(y)} \lambda_z(\mathrm{d}y) + 2B(x) \le a(x), \tag{2.7}$$

then the activity expansions $\rho_n(x_1, \dots, x_n; z)$ converge absolutely and the bounds

$$\rho_n(x_1, \dots, x_n; z) \le z(x_1)\cdots z(x_n)\mathrm{e}^{a(x_1)+\cdots+a(x_n)}$$

hold for all $n \in \mathbb{N}$ and $(x_1, \dots, x_n) \in \mathbb{X}^n$. Notice that for non-negative pair interactions, we can choose $B \equiv 0$ in condition (2.7).

Remark 2.2 Notice that via the substitution $\hat{a} = a - 2B$ the above criterion is equivalent to the existence of a measurable function $\hat{a} : \mathbb{X} \to \mathbb{R}_+$ such that for all $x \in \mathbb{X}$

$$\int_{\mathbb{X}} \big|f(x, y)\big| \mathrm{e}^{\hat{a}(y)+2B(y)} \lambda_z(\mathrm{d}y) \le \hat{a}(x).$$

Proof Assume that (2.7) holds and define $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}},\ \xi_n : \mathbb{X}^n \to [0,\infty)$, by

$$\xi_n(x_1, ..., x_n) := z(x_1)\cdots z(x_n)\mathrm{e}^{a(x_1)+\cdots+a(x_n)}$$

for some $a(\cdot)$ satisfying (2.7). The interactions fulfill the stability condition (2.1), therefore for every $n \in \mathbb{N}$ and $x_1, ..., x_n \in \mathbb{X}$ there exists an index $j \in \{1, \ldots, n\}$ such that the bound

$$\prod_{1\le i\le n, i\ne j} (1 + f(x_j, x_i)) \le \mathrm{e}^{2B(x_j)} \tag{2.8}$$

holds. Choose the selection rule s that always picks an element x_j satisfying (2.8) from $(x_1, \ldots, x_n)$. Plugging our choice of $\boldsymbol{\xi}$ into the left-hand side of Eq. (2.5) and bounding the interaction term as $\prod_{i=2}^n (1 + f(x_s, x_i')) \le \mathrm{e}^{2B(x_s)}$, we recognize an exponential series, and find altogether that the left-hand side of (2.5) is bounded by

$$z(x_s)z(x_2')\cdots z(x_n')\,\mathrm{e}^{a(x_2')+\cdots+a(x_n')}\exp\Big(\int_{\mathbb{X}} |f(x_s, y)|\mathrm{e}^{a(y)}\lambda_z(\mathrm{d}y) + 2B(x_s)\Big).$$

By condition (2.7), this is in turn bounded by $\xi_n(x_1, \ldots, x_n)$. It follows that condition (i) of Theorem 2.1 is satisfied. □

Analogously, one shows that the criterion by Fernández and Procacci [8], extended to soft-core and continuum systems by Faris in [5] and by Jansen in [14], is sufficient for absolute convergence of the activity expansions $\boldsymbol{\rho}$. We prove the result in the slightly more general setup of locally stable interactions.

Corollary 2.3 *Let z be a non-negative activity function and assume locally stable interactions in the sense of (2.2) for some $C \ge 0$. If there exists a measurable function $\mu : \mathbb{X} \to [0,\infty)$ such that for all $x \in \mathbb{X}$*

$$z(x)\left(1 + \sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k} \mathrm{e}^{\sum_{j=1}^k C(y_j)}\prod_{j=1}^k \big|f(x, y_j)\big| \prod_{1\le i<j\le k}(1 + f(y_i, y_j))\lambda_\mu^k(\mathrm{d}\boldsymbol{y})\right) \le \mu(x), \tag{2.9}$$

then the activity expansions $\rho_n(x_n, \ldots, x_n; z)$ converge absolutely and the bounds

$$\rho_n(x_n, \ldots, x_n; z) \le \prod_{1\le i<j\le n}\big(1 + f(x_i, x_j)\big)\prod_{i=1}^n \mu(x_i)$$

hold for all $n \in \mathbb{N}$ and $(x_1, \ldots, x_n) \in \mathbb{X}^n$. Notice that for non-negative pair interactions, we can choose $C \equiv 0$ in condition (2.9).

Remark 2.3 The Fernández–Procacci condition improves on the the Kotecký–Preiss condition—in the sense that the assumptions of Corollary 2.2 yield the assumptions of Corollary 2.3. In other words, Corollary 2.3 in general guarantees convergence of $\boldsymbol{\rho}$ on a larger domain of activities .

Proof Assume that (2.9) holds and define $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}},\ \xi_n : \mathbb{X}^n \to [0,\infty)$, by

$$\xi_n(x_1, \ldots, x_n) := \prod_{1\le i<j\le n}\big(1 + f(x_i, x_j)\big)\prod_{i=1}^n \mu(x_i)$$

for some μ satisfying (2.9). Let s be the selection rule that always selects the first entry—so that $x_s = x_1$ and $x_i' = x_i$ for $i \geq 2$. For locally stable pair potentials, we have

$$\begin{aligned}
&\prod_{i=2}^{n}(1+f(x_1,x_i))\xi_{n+k-1}(x_2,\dots,x_n,y_1,\dots,y_k)\\
&= \prod_{1\le i<j\le n}(1+f(x_i,x_j))\prod_{1\le i<j\le k}(1+f(y_i,y_j))\prod_{i=2}^{n}\prod_{j=1}^{k}(1+f(x_i,y_j))\prod_{i=2}^{n}\mu(x_i)\prod_{j=1}^{k}\mu(y_j)\\
&\le \Big(\prod_{1\le i<j\le n}(1+f(x_i,x_j))\prod_{i=2}^{n}\mu(x_i)\Big)\Big(\prod_{1\le i<j\le k}(1+f(y_i,y_j))\prod_{j=1}^{k}\mu(y_j)\Big)\mathrm{e}^{\sum_{j=1}^{k}C(y_j)}, \qquad (2.10)
\end{aligned}$$

where we used the local stability to estimate

$$\prod_{i=2}^{n}\prod_{j=1}^{k}(1+f(x_i,y_j)) = \prod_{j=1}^{k}\prod_{i=2}^{n}(1+f(x_i,y_j)) \le \prod_{j=1}^{k}\mathrm{e}^{C(y_j)} = \mathrm{e}^{\sum_{j=1}^{k}C(y_j)}.$$

We plug our choice of $\boldsymbol{\xi}$ into the left-hand side of (2.5) and use the estimate (2.10) together with the assumption (2.9) to find that condition (i) of Theorem 2.1 is satisfied. □

Remark 2.4 We see that Theorem 2.1 provides a mechanism to prove sufficient conditions for absolute convergence—by constructing a sequence of ansatz functions $\boldsymbol{\xi}$ tailored to satisfy the Kirkwood–Salsburg inequalities under the given condition. Conversely, given an appropriate sequence of ansatz functions $\boldsymbol{\xi}$, obtained, for example, as an approximation of $\boldsymbol{\rho}$, one can try to determine the corresponding sufficient condition for convergence.

We now proceed to demonstrate the usefulness of that approach by deriving a sufficient condition that improves on the classical examples above.

2.2 Abstract Polymer Models

In the following we want to consider the setup of abstract polymers [1, 8], in which the two classical conditions above—Kotecký–Preiss and Fernández–Procacci—were first introduced.

Let $\mathbb{X}$ be a countable set (the set of polymers), let $\mathcal{X}$ be the powerset of $\mathbb{X}$ and let λ simply be given by the counting measure. Moreover, let $R \subset \mathbb{X}\times\mathbb{X}$ be a symmetric and reflexive relation. We write $x \nsim y$ for $(x,y)\in R$ (and say that x and y are *incompatible*) and $x \sim y$ for $(x,y)\notin R$ (and say that x and y are *compatible*). Moreover, we call a subset $X \subset \mathbb{X}$ compatible if $x \sim y$ for all $x \neq y \in X$ and write $X \sim z$ for $z \in \mathbb{X}$ if $z \sim x$ for all $x \in X$. We set $\Gamma(x) := \{y \in \mathbb{X} |\ y \nsim x\}$ for any $x \in \mathbb{X}$ and extend this notation to $\Gamma(X) := \cup_{x\in X}\{y \in \mathbb{X} |\ y \nsim x\}$ for any $X \subset \mathbb{X}$. Notice that we do not require $\Gamma(x)$ to be finite sets and that $x \in \Gamma(x)$ for every $x \in \mathbb{X}$. Finally, we consider hard-core interactions corresponding to Mayer's f function given by $f(x,y) := -\mathbb{1}_{\{x\nsim y\}}$.

In this setting we prove a new, improved sufficient condition for absolute convergence of the activity expansions $\boldsymbol{\rho}$.

Proposition 2.4 *Let z be a non-negative activity function and assume that there exists $\mu : \mathbb{X} \to [0, \infty)$ such that for all $x \in \mathbb{X}$*

$$z(x)\left(1+\sum_{k\geq 1}\sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x,\ y_i\sim y_j}}\prod \mu(y_i)\prod_{w\in\Gamma(Y)}\mathrm{e}^{\mu(w)}\right)\leq \mu(x)\prod_{w\in\Gamma(x)}\mathrm{e}^{\mu(w)}, \tag{2.11}$$

where the inner sum on the left-hand side runs over compatible subsets $Y = \{y_1, \dots, y_k\} \subset \Gamma(x)$. Then the activity expansions $\rho_n(x_1, \dots, x_n; z)$ converge absolutely and the bounds

$$\rho_n(x_1,\dots,x_n;z)\leq \prod_{1\leq i<j\leq n}\mathbb{1}_{\{x_i\sim x_j\}}\prod_{i=1}^{n}\mu(x_i)\prod_{w\in\Gamma(\{x_1,\dots,x_n\})}\mathrm{e}^{\mu(w)}$$

hold for all $n \in \mathbb{N}$ and all $(x_1, \dots, x_n) \in \mathbb{X}^n$.

The proof of the proposition essentially exploits the following auxiliary result:

Lemma 2.5 *Let $\mu : \mathbb{X} \to [0, \infty)$. Then the following holds for every $x_1 \in \mathbb{X}$, $n \in \mathbb{N}$ and $X = \{x_2, \dots, x_n\} \subset \mathbb{X}$ such that $x_1 \sim x_i$ for all $i \in \{2, \dots, n\}$:*

$$\frac{\mu(x_1)\prod_{w\in\Gamma(x_1)}\mathrm{e}^{\mu(w)}}{1+\sum_{k\geq 1}\sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}}\prod_{i=1}^{k}\mu(y_i)\prod_{w\in\Gamma(Y)}\mathrm{e}^{\mu(w)}}$$
$$\leq\frac{\mu(x_1)\prod_{w\in\Gamma(x_1)\cap\Gamma(X)^C}\mathrm{e}^{\mu(w)}}{1+\sum_{k\geq 1}\sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j\\ y_i\sim X}}\prod_{i=1}^{k}\mu(y_i)\prod_{w\in\Gamma(Y)\cap\Gamma(X)^C}\mathrm{e}^{\mu(w)}}, \tag{2.12}$$

where $\Gamma(W)$ is given by $\cup_{i=1}^{n}\Gamma(w_i)$ for any $n \in \mathbb{N}$ and $W = \{w_1, \dots, w_n\} \subset \mathbb{X}$. The inner sum in the denominator on the left-hand side runs over compatible subsets $Y = \{y_1, \dots, y_k\} \subset \Gamma(x_1)$; the inner sum in the denominator on the right-hand side runs over all such subsets Y which additionally satisfy the constraint $Y \cap \Gamma(X) = \varnothing$, i.e., $y_i \sim X$ for all $i \in \{1, \dots, k\}$.

The lemma is of rather technical nature; for the interested reader, the proof is to be found in Appendix A.

Remark 2.5 The general idea behind the proof of Proposition 2.4 is to argue as in the proofs of the classical conditions presented in the previous section (Corollaries 2.2 and 2.3)—but to choose a sequence of ansatz functions $\boldsymbol{\xi}$ which, heuristically speaking, encode more of the structure of the exact solution to the Kirkwood–Salsburg equations (i.e., of the activity expansions $\boldsymbol{\rho}$) than the ansatz functions chosen in the proof of those corollaries. The intuition thereby is that "less multiplicative" ansatz functions $\boldsymbol{\xi}$ provide better convergence criteria.

Proof of Proposition 2.4 Assume that (2.11) holds and define $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$, $\xi_n : \mathbb{X}^n \to [0, \infty)$, by setting

$$\xi_n(x_1,\dots,x_n):=\prod_{1\leq i<j\leq n}\mathbb{1}_{\{x_i\sim x_j\}}\prod_{i=1}^{n}\mu(x_i)\prod_{w\in\Gamma(\{x_1,\dots,x_n\})}\mathrm{e}^{\mu(w)} \tag{2.13}$$

for some μ satisfying (2.11), for any $n \in \mathbb{N}$ and every $(x_1, .., x_n) \in \mathbb{X}^n$. Thereby we again use the convention $\Gamma(\{w_1, \ldots, w_n\}) = \cup_{i=1}^n \Gamma(w_i)$ for $\{w_1, ..., w_n\} \subset \mathbb{X}$. As in the preceeding proofs of the classical sufficient conditions, we show that our choice of $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$ satisfies the system of Kirkwood–Salsburg inequalities (2.5) from Theorem 2.1. To lighten the notation, we choose the same selection rule s as in the proof of Lemma 2.3 and denote by X the set $\{x_2, ..., x_n\}$. Notice that the left-hand side of (2.5) is equal to

$$z(x_1) \prod_{1\le i<j\le n} \mathbb{1}_{\{x_i \sim x_j\}} \prod_{i=2}^n \mu(x_i) \prod_{w\in\Gamma(X)} \mathrm{e}^{\mu(w)}$$
$$\times \left(1 + \sum_{k\ge1} \sum_{Y=\{y_1,...,y_k\}} \prod_{j=1}^k \mathbb{1}_{\{y_j \nsim x_1\}} \prod_{\substack{2\le i\le n\\ 1\le j\le k}} \mathbb{1}_{\{x_i\sim y_j\}} \prod_{1\le i<j\le k} \mathbb{1}_{\{y_i\sim y_j\}} \right.$$
$$\left. \prod_{j=1}^k \mu(y_j) \prod_{w\in\Gamma(Y)\cap\Gamma(X)^C} \mathrm{e}^{\mu(w)} \right).$$

By Lemma 2.5, the assumption that z satisfies the condition (2.11) implies that z also satisfies the inequality

$$z(x_1) \left(1 + \sum_{k\ge1} \sum_{Y=\{y_1,...,y_k\}} \prod_{j=1}^k \mathbb{1}_{\{y_j \nsim x_1\}} \prod_{\substack{2\le i\le n\\ 1\le j\le k}} \mathbb{1}_{\{x_i\sim y_j\}} \prod_{1\le i<j\le k} \mathbb{1}_{\{y_i\sim y_j\}} \right.$$
$$\left. \prod_{j=1}^k \mu(y_j) \prod_{w\in\Gamma(Y)\cap\Gamma(X)^C} \mathrm{e}^{\mu(w)} \right)$$
$$\le \mu(x_1) \prod_{w\in\Gamma(x_1)\cap\Gamma(X)^C} \mathrm{e}^{\mu(w)}$$

and thus, for our choice of $\boldsymbol{\xi}$, the left-hand side of (2.5) is bounded from above by

$$\prod_{1\le i<j\le n} \mathbb{1}_{\{x_i \sim x_j\}} \prod_{i=1}^n \mu(x_i) \prod_{w\in\Gamma(X)} \mathrm{e}^{\mu(w)} \prod_{w\in\Gamma(x_1)\cap\Gamma(X)^C} \mathrm{e}^{\mu(w)}$$
$$= \prod_{1\le i<j\le n} \mathbb{1}_{\{x_i \sim x_j\}} \prod_{i=1}^n \mu(x_i) \prod_{w\in\Gamma(X\cup\{x_1\})} \mathrm{e}^{\mu(w)} = \xi_n(x_1, ..., x_n),$$

which—by Theorem 2.1—yields the claim of the proposition. □

Example 2.1 Consider non-overlapping (hard-core interactions) cubes on $\mathbb{Z}^2$ of side-length 2 with translationally invariant activity z. The sufficient condition on z for the absolute convergence of $\rho(z)$ given by the Fernández–Procacci criterion provides the bound

$$z \le \max_{\mu\ge0} \frac{\mu}{1 + 9\mu + 16\mu^2 + 8\mu^3 + \mu^4} \approx 0.057271,$$

while our condition from Proposition 2.4 provides

$$z \le \max_{\mu\ge0} \frac{\mu e^{9\mu}}{1 + 9e^{9\mu}\mu + (6e^{15\mu} + 8e^{16\mu} + 2e^{17\mu})\mu^2 + 8e^{21\mu}\mu^3 + e^{25\mu}\mu^4} \approx 0.060833.$$

This corresponds to an improvement of approximately 6%.

2.3 Hard-Core Systems in the Continuum

Let $\mathscr{K}'$ be the collection of non-empty compact subsets of $\mathbb{R}^d$, equipped with the Hausdorff distance and Borel σ-algebra [19, Chapter I-4], and $\mathbb{X} \subset \mathscr{K}'$ a non-empty measurable subset. Here we want to additionally assume that $\mathbb{X}$ consists of bounded convex sets that are non-empty and regular closed, i.e., that are equal to the closure of its non-empty interior. Notice that such sets are compact and have finite positive Lebesgue measure that is equal to the Lebesgue measure of their interior. In practice $\mathbb{X}$ will consist of easily described subsets. For example, when dealing with closed balls $B_r(x) \subset \mathbb{R}^d$ we may identify $\mathbb{X}$ with $\mathbb{R}^d \times \mathbb{R}_+$. Consider the hard-core interactions given by the potential $v(X, Y) := \infty \mathbb{1}_{\{X \cap Y \neq \varnothing\}}$, Mayer's f function then is

$$f(X, Y) = -\mathbb{1}_{\{X \cap Y \neq \varnothing\}}.$$

Clearly the function is well-defined for general subsets $X, Y \subset \mathbb{R}^d$ that are not necessarily in $\mathbb{X}$, the domains of definition of the functions φ_n^{T} and $\psi_{n,n+k}$ extend accordingly.

For $D \subset \mathbb{R}^d$ and a measure λ_z on $\mathcal{X}$ defined as in (2.3), consider the formal series

$$T(D; z) := 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \varphi_{1+k}^{\mathsf{T}}(D, Y_1, \ldots, Y_k) \lambda_z^k(\mathrm{d}\boldsymbol{Y}). \tag{2.14}$$

As is well-known [6, Eq. (3.12)]

$$T(D; z) = \exp\Big(-\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{\exists i:\, Y_i \cap D \neq \varnothing\}} \varphi_k^{\mathsf{T}}(Y_1, \ldots, Y_k) \lambda_z^k(\mathrm{d}\boldsymbol{Y})\Big) \tag{2.15}$$

on the level of formal power series.

Moreover, if the domain D can be written as a finite union of disjoint objects $X_i \in \mathbb{X}$, say $D = X_1 \cup \ldots \cup X_n$ for $n \in \mathbb{N}$, then the identity

$$\begin{aligned} 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \varphi_{1+k}^{\mathsf{T}}(D, Y_1, \ldots, Y_k) \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \\ = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \psi_{n,n+k}(X_1, \ldots, X_n, Y_1, \ldots, Y_k) \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \end{aligned}$$

holds by Lemma 3.8 below and we recognize that the series $T(D, z)$ provide expansions for the reduced correlation functions in the sense that

$$\rho_n(X_1, \ldots, X_n; z) = z(X_1) \cdots z(X_n) \mathbb{1}_{\{X_1, \ldots, X_n \text{ disjoint}\}} T(X_1 \cup \cdots \cup X_n; z).$$

The absolute convergence of the expansions $\rho(z)$ for the correlation functions is implied by the absolute convergence of $T(D; z)$, i.e., by the pointwise convergence

$$1 + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} |\varphi_{1+k}^{\mathsf{T}}(D, Y_1, \ldots, Y_k)| \lambda_z^k(\mathrm{d}\boldsymbol{Y}) < \infty,$$

for all domains D that are unions of finitely many objects $X_i \in \mathbb{X}$.

Assume we are given a systematic way to chop up the objects $X \in \mathbb{X}$ into smaller bits and pieces, called *snippets* (think: analogous to representing a polymer as a collection of monomers in the discrete setup of subset polymers). That is, choose a positive number $\varepsilon > 0$ and assume that there is a designated collection $\mathbb{E}_\varepsilon$ of bounded Borel sets in $\mathbb{R}^d$, each of which is contained in some open ball of radius ε, and a *chopping map*

$$C : \mathbb{X} \to \mathcal{P}(\mathbb{E}_\varepsilon), \quad X \mapsto C(X)$$

such that for every $X \in \mathbb{X}$, $C(X) = \{E_1, \dots, E_m\}$ with $m \in \mathbb{N}$ and $E_1, \dots, E_m$ a set partition of X. We additionally want to assume that the topological boundary of every snippet is a λ-null set, i.e., $\lambda(\overline{E} \backslash E^\circ) = 0$ for all $E \in \mathbb{E}_\varepsilon$ (where $\overline{E}$ denotes the topological closure and E° the interior of E).

Let $\mathbb{D}_\varepsilon$ be the set of bounded domains $D \subset \mathbb{R}^d$ that can be written as the union of finitely many disjoint snippets. The empty set $D = \varnothing$ is an element of $\mathbb{D}_\varepsilon$. For two disjoint subsets $D_0, D_1 \subset \mathbb{R}^d$ with $D_0 \neq \varnothing$ and for finitely many objects $Y_1, \dots, Y_k \in \mathbb{X}$, $k \in \mathbb{N}$, set

$$I(D_0; D_1; Y_1, \dots, Y_k) := \Big(\prod_{i=1}^{k} \mathbb{1}_{\{D_0 \cap Y_i \neq \varnothing,\, D_1 \cap Y_i = \varnothing\}}\Big)\Big(\prod_{1 \le i < j \le k} \mathbb{1}_{\{Y_i \cap Y_j = \varnothing\}}\Big). \quad (2.16)$$

Theorem 2.6 *Let $z(\cdot)$ be a non-negative activity function. The following two conditions are equivalent:*

(i) *There exists a non-negative map $a : \mathbb{D}_\varepsilon \to \mathbb{R}_+$ such that for all $D \in \mathbb{D}_\varepsilon$, the map $\mathscr{K}' \ni F \mapsto a(D \cup F)$ is measurable and the following system of inequalities is satisfied: For all non-empty $D \in \mathbb{D}_\varepsilon$ with $C(D) = \{E_1, \dots, E_m\} \subset \mathbb{E}_\varepsilon$ for some $m \in \mathbb{N}$, there exists an $s \in \{1, \dots, m\}$ such that—setting $D' := D \backslash E_s$—we have*

$$\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} I(E_s; D'; Y_1, \dots, Y_k) \mathrm{e}^{a(D' \cup Y_1 \cup \dots \cup Y_k) - a(D')} \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \le \mathrm{e}^{a(E_s \cup D') - a(D')} - 1.$$

(ii) *$T(D; z)$ is absolutely convergent for all $D \in \mathbb{D}_\varepsilon$.*

Moreover, if one of the equivalent conditions (hence, both) holds true, then, for all $D \in \mathbb{D}_\varepsilon$, we have

$$\big|\log T(D; z)\big| \le \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{\exists i:\, Y_i \cap D \neq \varnothing\}} \big|\varphi_k^T(Y_1, \dots, Y_k)\big| \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \le a(D). \quad (2.17)$$

2.4 Subset Polymers

Let $\mathbb{X}$ consist of the finite non-empty subsets of $\mathbb{Z}^d$ (or any other countable set), and let $\mathcal{X} = \mathcal{P}(\mathbb{X})$ be the σ-algebra containing all subsets of $\mathbb{X}$. The reference measure λ is simply the counting measure. The interaction is a pure hard-core interaction as in Sect. 2.3. Notice that this setup is a special case of the abstract polymer setup introduced in Sect. 2.2. For a finite set $D \subset \mathbb{Z}^d$, define $T(D; z)$ as in (2.14). In statistical physics $T(D; z)$ corresponds to the probability that no polymer intersects D. If D is a polymer or a union of disjoint polymers, it corresponds to a *reduced correlation function* in the sense of [12].

Notice how in the case of subset polymers every polymer always can be "chopped" in a canonical way—into a disjoint collection of monomers. Those play the role of snippets from the previous section—that simplifies the formulation of a criterion for absolute convergence of the activity expansions $\boldsymbol{\rho}$ (compare next result with Theorem 2.6).

Theorem 2.7 *Let $(z(X))_{X \in \mathbb{X}}$ be a non-negative activity. The following two conditions are equivalent:*

(i) *There exists a function* $a(\cdot)$ *from the finite subsets of* $\mathbb{Z}^d$ *to* $[0,\infty)$ *such that* $a(\varnothing) = 0$ *and the following system of inequalities is satisfied: For all finite, non-empty subsets* $D \subset \mathbb{Z}^d$ *there exists an* $x \in D$ *such that—setting* $D' := D\backslash\{x\}$*—we have*

$$\sum_{\substack{Y\in\mathbb{X}:\\ Y\ni x,\, Y\cap D'=\varnothing}} z(Y)\mathrm{e}^{a(D'\cup Y)-a(D')} \le \mathrm{e}^{a(D'\cup\{x\})-a(D')} - 1. \tag{2.18}$$

(ii) $T(D;z)$ *is absolutely convergent for all finite subsets* $D \subset \mathbb{Z}^d$*.*

Moreover, if one of the equivalent conditions (hence, both) holds true, then, for all finite subsets $D \subset \mathbb{Z}^d$*, we have*

$$\left|\log T(D;z)\right| \le \sum_{k=1}^{\infty}\frac{1}{k!}\sum_{(Y_1,\dots,Y_k)\in\mathbb{X}^k} 1\!\!1_{\{\exists i:\, Y_i\cap D\neq\varnothing\}}\left|\varphi_k^T(Y_1,\dots,Y_k)\right| z(Y_1)\cdots z(Y_k) \le a(D). \tag{2.19}$$

The theorem is similar to Claim 1 in [1, Sect. 4.2]. As noted in [1], Theorem 2.7 allows for an easy recovery of the *extended Gruber–Kunz criterion*. The criterion is named after Gruber and Kunz [12], who proved a similar condition but with a strict inequality. See [8] for a comparison of the Gruber–Kunz criterion to other classical conditions.

Corollary 2.8 *Let* $(z(X))_{X\in\mathbb{X}}$ *be a non-negative activity. Suppose there exists some* $\alpha \ge 0$ *such that for all* $x \in \mathbb{Z}^d$*,*

$$\sum_{Y\ni x} z(Y)\,\mathrm{e}^{\alpha|Y|} \le \mathrm{e}^{\alpha} - 1. \tag{2.20}$$

Then $T(D;z)$ *is absolutely convergent, for all finite subsets* $D \subset \mathbb{Z}^d$*.*

Proof Set $a(D) := \alpha|D|$, where $\alpha > 0$ satisfies the inequality from (2.20). Because of the additivity of $a(\cdot)$, we have $a(D\cup Y) = a(D)+a(Y)$ for all finite, disjoint subsets $D, Y \subset \mathbb{Z}^d$. Therefore condition (2.18) becomes

$$\sum_{\substack{Y\ni x:\\ Y\cap D=\varnothing}} z(Y)\,\mathrm{e}^{a(Y)} \le \mathrm{e}^{a(\{x\})} - 1,$$

which depends on D only through the constraint $Y \cap D \neq \varnothing$ on the left-hand side. By the non-negativity of the activity z, it is clearly sufficient that

$$\sum_{Y\ni x} z(Y)\,\mathrm{e}^{a(Y)} \le \mathrm{e}^{a(\{x\})} - 1,$$

which holds true for all $x \in \mathbb{Z}^d$ because of (2.20). □

Another immediate consequence of Theorem 2.7 is that convergence of cluster expansions implies exponential decay of the activities in the object size. Precisely, set

$$V(D) := \sum_{\substack{Y\in\mathbb{X}:\\ Y\cap D\neq\varnothing}} z(Y).$$

Notice that if the activity is translationally invariant and not identically zero, one can choose an arbitrary polymer $X \in \mathbb{X}$ with positive activity, say $z_0 > 0$, and obtain the bound

$$V(D) \ge z_0|D|. \tag{2.21}$$

Theorem 2.9 *If $z(\cdot)$ is a non-negative activity and $T(D; z)$ is absolutely convergent for all finite subsets $D \subset \mathbb{Z}^d$, then necessarily*

$$\sum_{Y \ni x} z(Y)\mathrm{e}^{V(Y)} < \infty$$

for all $x \in \mathbb{Z}^d$.

Proof By condition (i) in Theorem 2.7, evaluated at $D' = \varnothing$, there exists a non-negative function $a(\cdot)$ such that

$$\sum_{Y \ni x} z(Y)\mathrm{e}^{a(Y)} \le \mathrm{e}^{a(\{x\})} - 1 < \infty.$$

For any polymer $Y \in \mathbb{X}$, the value $a(Y)$ is necessarily larger than $V(Y)$ by (2.19) and the claim follows. □

For translationally invariant systems, Theorem 2.9 says that if the activity expansions are absolutely convergent, then necessarily the activities are exponentially small in the size of the object—by (2.21) we can observe that $z(X) = O(\exp(-z_0|X|))$ when $|X| \to \infty$. Let us emphasize that the necessary exponential decay is an intrinsic limitation of the activity expansion, which cannot be eliminated by tinkering with different sufficient convergence conditions. Rigorous results for one-dimensional and hierarchical models [13, 15] suggest that the exponential decay is not needed for the convergence of the multi-species virial expansion, however for general systems this is so far an unproven conjecture.

3 Combinatorial Lemmas: Proof of Theorems 2.1, 2.6, and 2.7

3.1 Forest Partition Schemes: Alternating Sign Property

To obtain a better understanding of the series ρ_n given by the generating functions of multi-rooted graphs, we now consider a particular way to construct the latter—by taking a designated spanning forest and successively adding edges to it. This perspective onto multi-rooted graphs leads to a forest-graph equality analogous to the familiar tree-graph identity for connected graphs [8, Proposition 5] and allows for a direct proof of an alternating sign property for the coefficients $\psi_{n,n+k}$ of ρ_n in the case of repulsive interactions (i.e., for non-negative potentials).

The forest-graph equality builds on the notion of *forest partition schemes*—maps that assign spanning forests to multi-rooted graphs in $\mathcal{D}_{n,n+k}$ and thereby in a specific manner provide partitions of $\mathcal{D}_{n,n+k}$.

In the following, we let $\mathcal{F}_{n,n+k}$ denote the set of forest graphs on the vertex set $[n+k]$ consisting of n rooted trees, where the vertices $\{1, \dots, n\}$ are the roots of the trees (recall that a forest is an acyclic graph and a tree is a connected acyclic graph).

Definition 3.1 *(Forest partition scheme)* A *forest partition scheme* is a family of maps $\pi_{n,k}$: $\mathcal{D}_{n,n+k} \to \mathcal{F}_{n,n+k}$ such that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$, and all $F \in \mathcal{F}_{n,n+k}$, there exists a graph $R_{n,k}(F) \in \mathcal{D}_{n,n+k}$ with

$$\pi_{n,k}^{-1}(\{F\}) = \{G \in \mathcal{D}_{n,n+k} \mid E(F) \subset E(G) \subset E\big(R_{n,k}(F)\big)\} =: [F, R_{n,k}(F)].$$

To lighten the notation, we introduced partition schemes as families of maps on uncoloured structures. Notice, however, that partition schemes may be defined on coloured structures and may be allowed to depend on the colouring of the vertex set. Therefore, one could introduce families of maps $\pi_{k,n}(\boldsymbol{x}_{[n]})$, indexed additionally by colourings $\boldsymbol{x}_{[n]} \in \mathbb{X}^n$ of $[n] = \{1, ..., n\}$. Same graphs on the vertex set $[n]$ with different colourings $\boldsymbol{x}_{[n]}$ of the vertices can be mapped onto different forests under such partition schemes.

The existence of forest partition schemes is ensured by the existence of a large class of tree partition schemes, e.g, the Penrose tree partition scheme (see [8, 33]; for coulouring-dependent schemes see also [25, 31]).

Example 3.1 A particular forest partition scheme can be defined as follows: For a given multi-rooted graph, construct a connected graph from it by adding a ghost-vertex and connecting it to every root directly by an edge. Then apply the Penrose partition scheme to the resulting connected graph to obtain a spanning tree of this connected graph. Finally, by removing the ghost vertex as well as every edge incident to it, one gets a spanning forest of the initial multi-rooted graph. For the map given by this construction, the characterizing properties of a forest partition scheme follow from the corresponding properties of the Penrose tree partition scheme.

Naturally, the choice of the Penrose tree partition scheme in the example above is somewhat arbitrary; any tree partition scheme which does not "delete" any edge incident to the ghost vertex in the above construction yields a forest partition scheme via the same procedure.

Proposition 3.2 (Forest-graph equality) *Let $(\pi_{n,k})_{n\in\mathbb{N},\, k\in\mathbb{N}_0}$ be a forest partition scheme and let $(R_{n,k})_{n\in\mathbb{N},\, k\in\mathbb{N}_0}$ provide the corresponding family of multi-rooted graphs as in Definition 3.1. Then*

$$\psi_{n,n+k}(x_1, \dots, x_{n+k}) = \sum_{F\in\mathcal{F}_{n,n+k}} \prod_{\{i,j\}\in E(F)} f(x_i, x_j) \prod_{\{i,j\}\in E(R_{n,k}(F))\setminus E(F)} \big(1 + f(x_i, x_j)\big)$$

for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$, and $(x_1, \dots, x_{n+k}) \in \mathbb{X}^{n+k}$.

Proof The proof is similar to the standard proof of the tree-graph equality [8, Proposition 5]. We have

$$\begin{aligned}\psi_{n,n+k}(x_1, \dots, x_{n+k}) &= \sum_{G\in\mathcal{D}_{n,n+k}} w\big(G, (x_1, \dots, x_{n+k})\big)\\ &= \sum_{F\in\mathcal{F}_{n,n+k}} \sum_{\substack{G\in\mathcal{D}_{n,n+k}:\\ \pi_{n,k}(G)=F}} w\big(G, (x_1, \dots, x_{n+k})\big)\\ &= \sum_{F\in\mathcal{F}_{n,n+k}} \prod_{\{i,j\}\in E(F)} f(x_i, x_j) \prod_{\{i,j\}\in E(R_{n,k}(F))\setminus E(F)} \big(1 + f(x_i, x_j)\big).\end{aligned}$$

□

In the case of repulsive interactions, the forest-graph equality allows for a direct proof of the alternating sign property for the graph weights $\psi_{n,n+k}$. For $n = 1$, it reduces to the well-known alternating sign property [8, Eq. (2.8)]

$$\varphi_n^{\mathsf{T}}(x_1, \dots, x_n) = (-1)^n \big|\varphi_n^{\mathsf{T}}(x_1, \dots, x_n)\big| \tag{3.1}$$

of the Ursell functions.

Corollary 3.3 *For non-negative potentials, we have*

$$\psi_{n,n+k}(x_1,\ldots,x_{n+k}) = (-1)^k \big|\psi_{n,n+k}(x_1,\ldots,x_{n+k})\big|$$

for all $n \in \mathbb{N}$, all $k \in \mathbb{N}_0$, and all $(x_1,\ldots,x_{n+k}) \in \mathbb{X}^{n+k}$.

Proof Each forest $F \in \mathcal{F}_{n,n+k}$ has exactly k edges. Indeed, the forest F consists of trees $T_1,\ldots,T_n$. Let m_i be the number of vertices of the tree T_i; thus $m_1 + \cdots + m_n = n + k$. Each tree T_i has exactly $m_i - 1$ edges, therefore the number of edges of the forest is given by $\sum_{i=1}^n (m_i - 1) = k$. Since $f \le 0$ and $1 + f \ge 0$ for non-negative potentials, it follows that

$$\psi_{n,n+k}(x_1,\ldots,x_{n+k}) = (-1)^k \sum_{F\in\mathcal{F}_{n,n+k}} \prod_{\{i,j\}\in E(F)} |f(x_i,x_j)| \prod_{\{i,j\}\in E(R_{n,k}(F))\setminus E(F)} \big(1 + f(x_i,x_j)\big),$$

hence $(-1)^k \psi_{n,n+k}(x_1,\ldots,x_{n+k}) \ge 0$. □

We will use the alternating sign property to establish that—in the case of non-negative potentials—condition (i) in Theorem 2.1 is not only sufficient but also necessary for absolute convergence of ρ.

We conclude this section with a lemma that is not needed for the proof of Theorem 2.1 but enters the analysis of hard-core models, see the proof of Lemma 3.8 below.

Lemma 3.4 *For all $n \in \mathbb{N}$, all $k \in \mathbb{N}_0$, and all $(x_1,\ldots,x_{n+k}) \in \mathbb{X}^{n+k}$, we have*

$$\begin{aligned}\psi_{n,n+k}(x_1,\ldots,x_{n+k}) = &\prod_{1\le i<j\le n} \big(1 + f(x_i,x_j)\big)\\ &\times \sum_{\{V_1,\ldots,V_r\}} \prod_{\ell=1}^{r} \Bigg(\prod_{\substack{1\le i\le n,\\ j\in V_\ell}} \big(1 + f(x_i,x_j)\big) - 1 \Bigg) \varphi^T_{|V_\ell|}\big((x_j)_{j\in V_\ell}\big),\end{aligned} \tag{3.2}$$

where the sum runs over all set partitions $\{V_1,\ldots,V_r\}$ of non-root vertices $\{n+1,\ldots,n+k\}$.

Remark 3.1 The lemma allows for an alternative proof of the alternating sign property from Corollary 3.3, starting from the well-known alternating sign property of the Ursell function instead of the forest-graph equality. Indeed, the sign of every summand in the right-hand side of (3.2) is

$$(-1)^{r+\sum_{i=1}^r (|V_i|-1)} = (-1)^k.$$

Proof of Lemma 3.4 For $n = 1$, the lemma reduces to a well-known equality for the Ursell functions, see e.g. [11, Eq. (5.13)]. For $n \ge 2$, the proof is similar, we provide the details for the reader's convenience. Every multi-rooted graph $G \in \mathcal{D}_{n,n+k}$ can be constructed in the following way. On the root set $\{1,\ldots,n\}$ pick an arbitrary graph G_0. On the complement of the root set do the following construction: Partition the set of the non-root vertices into r sets V_1, ..., V_r, $r \le k$. For every block V_ℓ, pick a connected graph G_ℓ with vertex set V_ℓ, and in addition a non-empty set of edges $E_\ell \subset \big\{\{i,j\} \mid i \in \{1,\ldots,n\},\ j \in V_\ell\big\}$. Then the graph

G with vertices $1,\dots,n+k$ and edge set given by the union of $E_1,\dots,E_\ell$ and of the edge sets of $G_0,G_1,\dots,G_r$ is in $\mathcal{D}_{n,n+k}$, its graph weight is

$$w(G;x_1,\dots,x_{n+k}) = w(G_0;x_1,\dots,x_n)\prod_{\ell=1}^{r}\Big(\prod_{\{i,j\}\in E_\ell} f(x_i,x_j)\Big)w\big(G_\ell;(x_j)_{j\in V_\ell}\big).$$

Summation over G_0 yields the factor $\prod_{1\le i<j\le n}(1+f(x_i,x_j))$. Summation over the connected graphs G_ℓ yields $\varphi^{\mathsf{T}}_{|V_\ell|}(\boldsymbol{x}_{V_\ell})$. Finally, summation over the edge sets E_ℓ yields the factor $\prod_{1\le i\le n,\, j\in V_\ell}(1+f(x_i,x_j))-1$. □

3.2 Kirkwood–Salsburg Equations: Proof of Theorem 2.1

To prove our main result, Theorem 2.1, we will show that the activity expansions $\boldsymbol{\rho}$ satisfy the Kirkwood–Salsburg inequalities and, moreover, that equality holds for non-negative interactions. To do so, we will need to establish a recursive formula for the coefficients $\psi_{n,n+k}$ of ρ_n given in terms of multi-rooted graphs.

Lemma 3.5 *Let $n\in\mathbb{N}$, $k\in\mathbb{N}_0$ and $(x_1,\dots,x_n)\in\mathbb{X}^n$. Abbreviate $s=s(\boldsymbol{x})$. For $L\subset[k]$, let ℓ denote the cardinality of L. Then for all $(y_1,\dots,y_k)\in\mathbb{X}^k$,*

$$\begin{aligned}\psi_{n,n+k}(x_1,\dots,x_n,y_1,\dots,y_k) &= \textstyle\prod_{\substack{1\le i\le n:\\ i\ne s}}\big(1+f(x_s,x_i)\big)\sum_{L\subset[k]}\Big(\prod_{i\in L} f(x_s,y_i)\Big)\\ &\quad\times\psi_{n-1+\ell,n-1+k}\big(x_2',\dots,x_n',(y_i)_{i\in L},(y_j)_{j\in[k]\setminus L}\big).\end{aligned}$$

Furthermore, if $n\ge 2$,

$$\psi_{n,n}(x_1,\dots,x_n) = \prod_{\substack{1\le i\le n:\\ i\ne s}}\big(1+f(x_s,x_i)\big)\,\psi_{n-1,n-1}\big(x_2',\dots,x_n'\big).$$

The lemma is proven in [14, Lemma 4.1] and holds true as well for pair potentials that may take negative values. The index set L corresponds to the non-root vertices adjacent to the selected vertex s. Similar recurrent relation are well-known from the literature and have been employed in the context of both activity and density (virial) expansions for a long time (see, e.g., [20, Eq. 5]).

We now want to translate the recurrence relation for coefficients $\psi_{n,n+k}$ from Lemma 3.5 into integral equations for partial sums and series. For a non-negative activity function, we set

$$\tilde\rho_n(x_1,\dots,x_n;z) := z(x_1)\cdots z(x_n)\sum_{k=0}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k}\big|\psi_{n,n+k}(x_1,\dots,x_n,\boldsymbol{y})\big|\lambda_z^k(\mathrm{d}\boldsymbol{y}).$$

Notice that $\rho_n(x_1,\dots,x_n;z)$ is absolutely convergent if and only if $\tilde\rho_n(x_1,\dots,x_n;z)<\infty$, and, in the case of non-negative potentials,

$$\tilde\rho_n(x_1,\dots,x_n;z) = (-1)^n\rho_n(x_1,\dots,x_n;-z)$$

holds due to the alternating-sign property from Corollary 3.3.

Let $\tilde{\boldsymbol{S}}_N(z) = (\tilde S_{N,n}(\cdot;z))_{n\in\mathbb{N}}$ be the vector of partial sums given by

$$\tilde S_{N,n}(x_1,\dots,x_n;z) := z(x_1)\cdots z(x_n)\sum_{k=0}^{N-n}\frac{1}{k!}\int_{\mathbb{X}^k}\big|\psi_{n,n+k}(x_1,\dots,x_n,\boldsymbol{y})\big|\lambda_z^k(\mathrm{d}\boldsymbol{y})$$

if $N \geq n$, and 0 otherwise. The summand for $k = 0$ is to be read as $|\psi_{n,n}(x_1, \dots, x_n)|$.

Proposition 3.6 *For general pair-interactions, we have*

$$\tilde{\boldsymbol{\rho}}(z) \leq \boldsymbol{e}_z + \tilde{K}_z^s \tilde{\boldsymbol{\rho}}(z) \tag{3.3}$$

and

$$\tilde{\boldsymbol{S}}_1(z) = \boldsymbol{e}_z, \quad \tilde{\boldsymbol{S}}_{N+1} \leq \boldsymbol{e}_z + \tilde{K}_z^s \tilde{\boldsymbol{S}}_N(z) \quad (N \geq 1).$$

Moreover, for non-negative potentials, we get the equalities

$$\tilde{\boldsymbol{\rho}}(z) = \boldsymbol{e}_z + \tilde{K}_z^s \tilde{\boldsymbol{\rho}}(z) \tag{3.4}$$

and

$$\tilde{\boldsymbol{S}}_{N+1} = \boldsymbol{e}_z + \tilde{K}_z^s \tilde{\boldsymbol{S}}_N(z) \quad (N \geq 1).$$

Proof The equality $\tilde{\boldsymbol{S}}_1(z) = \boldsymbol{e}_z$ follows from the definition of $\tilde{\boldsymbol{S}}_1(z)$ and $\psi_{1,1}(x_1) = 1$. For the recurrence relation, we employ arguments from [14, Sect. 4] and combine them with the alternating sign property from Corollary 3.3 to argue equality in the case of non-negative potentials.

Consider first $\tilde{S}_{N+1,n}(z)$ with $2 \leq n \leq N+1$. Define

$$\mathscr{R}_{n,\ell}(x_1, \dots, x_n; y_1, \dots, y_\ell) := z(x_s) \prod_{i=2}^{n} (1 + f(x_s, x_i')) \prod_{i=1}^{\ell} \big| f(x_s, y_i) \big|.$$

Fix $n \geq 2$. Lemma 3.5 and the triangle inequality yield

$$\begin{aligned} &\big|\psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k)\big| \\ &\quad \leq \sum_{L \subset [k]} \mathscr{R}_{n,\ell}(x_1, \dots, x_n; \boldsymbol{y}_L) \left|\psi_{n-1+\ell,n-1+k}\big(x_2', \dots, x_n', \boldsymbol{y}_L, \boldsymbol{y}_{[k]\setminus L}\big)\right|, \end{aligned} \tag{3.5}$$

where $\ell = \# L$. When we integrate over $y_1, \dots, y_k$, all sets L with the same cardinality contribute the same, therefore

$$\begin{aligned} &\frac{1}{k!} \int_{\mathbb{X}^k} \big|\psi_{n,n+k}(x_1, \dots, x_n, y_1, \dots, y_k)\big| \, \lambda_z^k(\mathrm{d}\boldsymbol{y}) \\ &\quad \leq \sum_{\ell=0}^{k} \frac{1}{\ell!(k-\ell)!} \int_{\mathbb{X}^k} \mathscr{R}_{n,\ell}(x_1, \dots, x_n; y_1, \dots, y_\ell) \left|\psi_{n-1+\ell,n-1+k}\big(x_2', \dots, x_n', \boldsymbol{y}\big)\right| \lambda_z^k(\mathrm{d}\boldsymbol{y}). \end{aligned}$$

Summing over $k = 0, \dots, N+1-n$ we obtain a double sum over k and ℓ. A change in summation indices from (ℓ, k) to $(\ell, m) = (\ell, k - \ell)$ yields

$$\tilde{S}_{N+1,n}(x_1, \dots, x_n; z) \leq \sum_{\ell=0}^{N+1-n} \frac{1}{\ell!} \int_{\mathbb{X}^\ell} \mathscr{R}_{n,\ell}(x_1, \dots, x_n; y_1, \dots, y_\ell) \Big\{ \cdots \Big\} \lambda_z^\ell\big(\mathrm{d}(y_1, \dots y_\ell)\big),$$

$$\{\cdots\} = \sum_{m=0}^{N+1-n-\ell} \frac{1}{m!} \int_{\mathbb{X}^m} \left|\psi_{n-1+\ell,n-1+\ell+m}\big(x_2', \dots, x_n', \boldsymbol{y}\big)\right| \lambda_z^m\big(\mathrm{d}(y_{\ell+1}, \dots y_{\ell+m})\big).$$

The term in curly braces is nothing else but $\tilde{S}_{N,n-1+\ell}(x_2', \dots, x_n', y_1, \dots, y_\ell)$. For $\ell \geq N+1-n$, the function $\tilde{S}_{N,n-1+\ell}(\cdot; z)$ is identically zero. It follows that

$$\tilde{S}_{N+1,n}(x_1, \dots, x_n; z) \leq \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \int_{\mathbb{X}^\ell} \mathscr{R}_{n,\ell}(x_1, \dots, x_n; \boldsymbol{y}) \tilde{S}_{N,n-1+\ell}(x_2', \dots, x_n', \boldsymbol{y}) \lambda_z^\ell\big(\mathrm{d}\boldsymbol{y}\big).$$

This proves the inequality $\tilde{S}_{N+1,n}(\cdot; z) \leq \big(\tilde{K}_z^s S_N(z)\big)_n(\cdot)$. The cases $n \geq N+2$ and $n = 1$ are treated in a similar fashion, we leave the details to the reader. For the equality $\tilde{S}_{N+1,n}(\cdot; z) = \big(\tilde{K}_z^s S_N(z)\big)_n(\cdot)$ in the case of non-negative potentials, notice that for such potentials (3.5) holds with an equality — due to the alternating-sign property from Corollary 3.3.

Finally, by passing to the limit $N \to \infty$ in the recurrence relation for $\tilde{\boldsymbol{S}}_N(z)$, the inequality (3.3) for $\tilde{\boldsymbol{\rho}}(z)$ follows (and in the case of non-negative potentials the fixed point equation (3.4) is obtained). Notice that all exchanges of limits, sums and integrals are permitted by monotone convergence and because all terms involved are non-negative. □

Proof of Theorem 2.1 For the implication (i) ⇒ (ii), suppose there exists a sequence $\boldsymbol{\xi} = (\xi_n)_{n\in\mathbb{N}}$ of measurable non-negative functions $\xi_n : \mathbb{X}^n \to \mathbb{R}_+$ such that $\boldsymbol{e}_z + \tilde{K}_z^s \boldsymbol{\xi} \leq \boldsymbol{\xi}$. We prove by induction over N that $\tilde{\boldsymbol{S}}_N(z) \leq \boldsymbol{\xi}$ for all $N \in \mathbb{N}$. For $N = 1$, we have

$$\tilde{\boldsymbol{S}}_1 = \boldsymbol{e}_z \leq \boldsymbol{e}_z + \tilde{K}_z^s \boldsymbol{\xi} \leq \boldsymbol{\xi}.$$

If $\tilde{\boldsymbol{S}}_N \leq \boldsymbol{\xi}$ for some $N \in \mathbb{N}$, then

$$\tilde{\boldsymbol{S}}_{N+1} \leq \boldsymbol{e}_z + \tilde{K}_z^s \tilde{\boldsymbol{S}}_N(z) \leq \boldsymbol{e}_z + \tilde{K}_z^s \boldsymbol{\xi} \leq \boldsymbol{\xi},$$

where the first inequality holds by Proposition 3.6 and the second one due to the inductive hypothesis and the monotonicity of $\tilde{K}_z^s$ on non-negative functions.

This completes the induction and proves $\tilde{\boldsymbol{S}}_N \leq \boldsymbol{\xi}$ for all N. Passing to the limit $N \to \infty$, we find $\tilde{\boldsymbol{\rho}} \leq \boldsymbol{\xi}$. This proves the absolute convergence (ii) as well as the bound (2.6).

Left to show is the implication (ii) ⇒ (i) under the additional assumption that the potential is non-negative. Suppose that $\rho_n(x_1, \ldots, x_n; z)$ is absolutely convergent, for all $n \in \mathbb{N}$ and $(x_1, \ldots, x_n) \in \mathbb{X}^n$. Then $\tilde{\rho}_n(x_1, \ldots, x_n; z)$ is finite everywhere and we may set

$$\xi_n(x_1, \ldots, x_n) := \tilde{\rho}_n(x_1, \ldots, x_n; z).$$

Proposition 3.6 yields $\boldsymbol{\xi} = \boldsymbol{e}_z + \tilde{K}_z^s \boldsymbol{\xi}$ hence a fortiori $\boldsymbol{\xi} \geq \boldsymbol{e}_z + \tilde{K}_z^s \boldsymbol{\xi}$, as a pointwise inequality for all vector entries. This proves (i). □

3.3 Integral Equations for Hard-Core Models: Proof of Theorem 2.6

In this section we specialize to hard-core systems in the continuum as in Sect. 2.3 and use capital letters for objects $X \in \mathbb{X}$. Let $\mathcal{D}_{n,n+k}^{\mathrm{red}} \subset \mathcal{D}_{n,n+k}$ be the collection of graphs $G \in \mathcal{D}_{n,n+k}$ that have no edges linking any two root vertices $i, j \in \{1, \ldots, n\}$. Define $\psi_{n,n+k}^{\mathrm{red}}$ in a similar way as $\psi_{n,n+k}$ but with summation over graphs in $\mathcal{D}_{n,n+k}^{\mathrm{red}}$. It is not difficult to check that

$$\psi_{n,n+k}(X_1, \ldots, X_{n+k}) = \prod_{1\leq i<j\leq n} \big(1 + f(X_i, X_j)\big) \psi_{n,n+k}^{\mathrm{red}}(X_1, \ldots, X_{n+k}).$$

The reduced functions $\psi_{n,n+k}^{\mathrm{red}}$ satisfy recurrence relations similar to Lemma 3.5. Define

$$\begin{aligned} g(X_1; X_2, \ldots, X_n; Y_1, \ldots, Y_k) := \prod_{j=1}^{k} f(X_1, Y_j) \prod_{1\leq i<j\leq k} \big(1 + f(Y_i, Y_j)\big) \\ \prod_{\substack{2\leq i\leq n \\ 1\leq j\leq k}} \big(1 + f(X_i, Y_j)\big). \end{aligned} \tag{3.6}$$

Remember the indicator I from (2.16) and notice

$$g(X_1; X_2, \dots, X_n; Y_1, \dots, Y_k) = (-1)^k I(X_1; X_2 \cup \dots \cup X_n; Y_1, \dots, Y_k). \tag{3.7}$$

Lemma 3.7 *For all $k, n \in \mathbb{N}$, we have*

$$\begin{aligned} &\psi^{\mathrm{red}}_{n,n+k}(X_1, \dots, X_n, Y_1, \dots, Y_k) \\ &\quad = \sum_{L \subset [k]} g\big(X_s; X_2', \dots, X_n'; (Y_j)_{j \in L}\big) \psi^{\mathrm{red}}_{n-1+\ell, n-1+k}\big(X_2', \dots, X_n', (Y_j)_{j\in L}, (Y_j)_{j \in [k]\setminus L}\big). \end{aligned}$$

The proof is based on combinatorial considerations similar to the proof of Lemma 4.1 in [14], we leave the details to the reader. The lemma holds true for arbitrary subsets of $\mathbb{R}^d$, the X_i's and Y_j's need not be in $\mathbb{X}$. That applies to our next result as well.

Lemma 3.8 *For all $k, n \in \mathbb{N}$, we have*

$$\psi^{\mathrm{red}}_{n,n+k}(X_1, \dots, X_n, Y_1, \dots, Y_k) = \varphi^T_{1+k}(X_1 \cup \dots \cup X_n, Y_1, \dots, Y_k).$$

Proof Revisiting the proof of Lemma 3.4, we see that

$$\begin{aligned} &\psi^{\mathrm{red}}_{n,n+k}(X_1, \dots, X_n, Y_1, \dots, Y_k) \\ &\quad = \sum_{\{V_1, \dots, V_r\}} \prod_{\ell=1}^r \Bigg(\prod_{\substack{1 \le i \le n, \\ j \in V_\ell}} \big(1 + f(X_i, Y_j)\big) - 1 \Bigg) \varphi^{\mathsf{T}}_{|V_\ell|}\big((Y_j)_{j \in V_\ell}\big), \end{aligned} \tag{3.8}$$

where the sum runs over all set partitions $\{V_1, \dots, V_r\}$ of non-root vertices $\{n+1, \dots, n+k\}$. For hard-core interactions, the term in parentheses is equal to minus the indicator that $D := X_1 \cup \dots \cup X_n$ is intersected by at least one Y_j, $j \in V_\ell$. Thus

$$\psi^{\mathrm{red}}_{n,n+k}(X_1, \dots, X_n, Y_1, \dots, Y_k) = \sum_{\{V_1, \dots, V_r\}} \prod_{\ell=1}^r \big(-\mathbb{1}_{\{\exists j \in V_\ell :\, Y_j \cap D \neq \varnothing\}}\big) \varphi^{\mathsf{T}}_{|V_\ell|}\big((Y_j)_{j \in V_\ell}\big).$$

Using again (3.8), we deduce

$$\psi^{\mathrm{red}}_{n,n+k}(X_1, \dots, X_n, Y_1, \dots, Y_k) = \psi_{1,1+k}(D, Y_1, \dots, Y_k) = \varphi^{\mathsf{T}}_{1+k}(D, Y_1, \dots, Y_k).$$

□

Lemma 3.9 *Let $k \in \mathbb{N}$ and let D_0, D_1 be two disjoint subsets of $\mathbb{R}^d$, with $D_0 \neq \varnothing$. Then*

$$\begin{aligned} &\varphi^T_{1+k}(D_0 \cup D_1, Y_1, \dots, Y_k) \\ &\quad = \sum_{L \subset [k]} (-1)^\ell I\big(D_0; D_1; (Y_i)_{i \in L}\big) \varphi^T_{1+k-\ell}\Big(D_1 \cup \big(\bigcup_{i \in L} Y_i\big), (Y_j)_{j \in [k]\setminus L}\Big), \end{aligned}$$

where the sum is taken over all subsets $L \subset [k]$ and ℓ denotes the cardinality of L.

Proof The claim of the lemma follows from Eq. (3.7), Lemmas 3.7 and 3.8. □

For $D \in \mathbb{D}_\varepsilon$ with $E_1, \dots, E_n \in \mathbb{E}_\varepsilon$ and $C(D) = \{E_1, \dots, E_n\}$, let

$$\tilde{T}(D; z) := 1 + \sum_{k=1}^\infty \frac{1}{k!} \int_{\mathbb{X}^k} \big|\varphi^{\mathsf{T}}_{1+k}(D, Y_1, \dots, Y_k)\big| \lambda_z^k(\mathrm{d}\boldsymbol{Y}),$$

$\tilde{T}_1(D; z) := \delta_{n,1}(\{E_1, \dots, E_n\})$, and for $N \geq 2$,

$$\tilde{T}_N\big(D; z\big) := \mathbb{1}_{\{n \leq N\}} + \sum_{k=1}^{N-n} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{n+\sum_{i=1}^k |C(Y_i)| \leq N\}} \big|\varphi^{\mathsf{T}}_{1+k}(D, Y_1, \dots, Y_k)\big| \lambda_z^k(\mathrm{d}\boldsymbol{Y}).$$

Although the value of $\tilde{T}(D; z)$ (if the series converges) does not depend on the choice of the chopping map, notice that the value of $\tilde{T}_N(D; z)$ clearly does depend on $C(D) = \{E_1, \dots, E_n\}$ and not only on $E_1 \cup \dots \cup E_n$ — due to the constraint on the number of snippets.

A selection rule is a map s from collections of disjoint snippets $\mathbb{D}_\varepsilon$ to $\mathbb{E}_\varepsilon$ such that

$$s(\{E_1, \dots, E_n\}) \in \{E_1, \dots, E_n\},$$

i.e., $s(\cdot)$ selects one of the snippets. We use the suggestive but somewhat abusive notation E_s for the selected snippet, and let $E'_2, \dots, E'_n$ be any enumeration of the remaining snippets. If $\xi(\cdot)$ is a function from $\mathbb{D}_\varepsilon$ to $\mathbb{R}_+$ that satisfies the measurability assumption from Theorem 2.6, define a new function $\tilde{\kappa}_z^s \xi$ (possibly assuming the value "∞") by setting

$$\begin{aligned}(\tilde{\kappa}_z^s \xi)\big(D\big) &:= \mathbb{1}_{\{n \geq 2\}}\, \xi\big(E'_2 \cup \dots \cup E'_n\big) \\ &+ \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} I(E_s; E'_2 \cup \dots \cup E'_n; Y_1, \dots, Y_k) \xi\Big(E'_2 \cup \dots E'_n \cup Y_1 \cup \dots \cup Y_k\Big) \lambda_z^k(\mathrm{d}\boldsymbol{Y})\end{aligned} \tag{3.9}$$

for $D \in \mathbb{D}_\varepsilon$ with $E_1, \dots, E_n \in \mathbb{E}_\varepsilon$ and $C(D) = \{E_1, \dots, E_n\}$. Furthermore, let $e(D) := \delta_{n,1}(\{E_1, \dots, E_n\})$ be the indicator that D is a single snippet.

Let z be a non-negative activity such that for every non-empty $D \in \mathbb{D}_\epsilon$ the series $\tilde{T}\big(D; z\big)$ converges absolutely. Notice that the topology induced by the Hausdorff distance is equivalent to the myopic topology and the map $\mathscr{K}' \ni F \mapsto \mathbb{1}_{\{F \in \mathscr{K}' |\ F \cap B \neq \varnothing\}}(F) = -f(F, B)$ is measurable with respect to the myopic topology for all compact subsets B (see [21]). Measurability of $\mathscr{K}' \ni F \mapsto \tilde{T}(D \cup F; z)$ for every $D \in \mathscr{K}'$ can be concluded, e.g., by representing the series $\tilde{T}(D \cup F; z)$ as in Eq. (2.15). Since $\tilde{T}(D \cup F; z) = \tilde{T}(\overline{D} \cup F; z)$ for every $D \in \mathbb{D}_\varepsilon$, its topological closure $\overline{D}$ in $\mathbb{R}^d$ and every $F \in \mathscr{K}'$ (by our assumptions that the boundaries of snippets are λ-null set), the measurability of $\mathscr{K}' \ni F \mapsto \tilde{T}(D \cup F; z)$ for all $D \in \mathbb{D}_\varepsilon$ follows.

The next result is an analogue of Proposition 3.6.

Proposition 3.10 *We have*

$$\tilde{T}\big(\cdot; z\big) = e(\cdot) + \tilde{\kappa}_z^s \tilde{T}(\cdot; z).$$

Moreover $\tilde{T}_1(\cdot; z) = e(\cdot)$ *and for* $N \geq 1$

$$\tilde{T}_{N+1}\big(\cdot; z\big) = e(\cdot) + (\tilde{\kappa}_z^s \tilde{T}_N)\big(\cdot; z\big).$$

The proposition follows from Lemma 3.9 by arguments similar to the proof of Proposition 3.6, therefore the proof is omitted.

Proof of Theorem 2.6 To show the implication (ii) $\Rightarrow$ (i) suppose that $T(D; z)$ is absolutely convergent and thus $\tilde{T}(D; z)$ is convergent for all non-empty $D \in \mathbb{D}_\varepsilon$. Moreover, $\tilde{T}(D; z)$ is uniformly bounded from below by 1 and does not depend on the choice of the chopping map C. We set

$$a(D) := \log \tilde{T}(D; z) \geq 0$$

for non-empty $D \in \mathbb{D}_\varepsilon$ and $a(\varnothing) := 0$. Furthermore, for $D \in \mathbb{D}_\varepsilon$ with $E_1, \ldots, E_m \in \mathbb{E}_\varepsilon$ and $C(D) = \{E_1, \ldots, E_m\}$, let $E_s \in \mathbb{E}_\varepsilon$ be given by $E_{s(\{E_1,\ldots,E_m\})}$ for some selection rule $s(\cdot)$ and set $D' := D \backslash E_s$. Exploiting the fixed point equation for $\tilde{T}(\cdot; z)$ from Proposition 3.10, we get

$$\mathrm{e}^{a(D')} + \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} I(E_s; D'; Y_1, \ldots, Y_k) \mathrm{e}^{a(D' \cup Y_1 \cup \cdots \cup Y_k)} \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \leq \mathrm{e}^{a(E_s \cup D')}.$$

Item (i) of Theorem 2.6 follows upon multiplication with $\exp(-a(D'))$ on both sides (in fact, we have shown that the inequality from item (i) holds for every choice of $s \in \{1, \ldots, m\}$). The implication (i) $\Rightarrow$ (ii) follows from Proposition 3.10 by an induction over N similar to the proof of Theorem 2.1 on p. 20. Indeed, check that for the induction step it is sufficient that the corresponding system of Kirkwood–Salsburg inequalities holds for all finite disjoint unions of snippets (for any choice of the chopping map C, the snippet-size $\varepsilon > 0$ and the selection rule s). Bound (2.17) is then established using the triangle inequality, the alternating sign property, and Eq. (2.15). □

3.4 Recurrence Relations for Subset Polymers: Proof of Theorem 2.7

Just as in the continuous case, we will show that the expansions $\tilde{T}$ solve a system of integral equations in the discrete setup. We do so by providing a recursive formula for the corresponding coefficients $\varphi^{\mathsf{T}}_{1+k}$.

Lemma 3.11 *For all finite subsets $D' \subset \mathbb{Z}^d$, all $x \in \mathbb{Z}^d \setminus D'$, and all $k \in \mathbb{N}$,*

$$\begin{aligned}&\varphi^{T}_{1+k}(D' \cup \{x\}, Y_1, \ldots, Y_k)\\ &\quad = \varphi^{T}_{1+k}(D', Y_1, \ldots, Y_k) + \sum_{i=1}^{k} \big(-\mathbb{1}_{\{Y_i \ni x, Y_i \cap D = \varnothing\}}\big) \varphi^{T}_{k}\big(D' \cup Y_i, (Y_j)_{j \neq i}\big)\end{aligned}$$

with $\varphi^{T}_1 \equiv 1$.

Proof Notice that the analogue of Lemma 3.9 holds in the discrete setup of subset polymers as well, in particular for the choice $D_0 := \{x\}$ and $D_1 := D'$. However, since two disjoint polymers Y_1 and Y_2 cannot both intersect the same monomer x, only the summands corresponding to $L = \varnothing$ and $|L| = 1$ in the sum on the right-hand side of the identity in Lemma 3.9 provide non-trivial contributions. □

Let D be a finite non-empty subset of $\mathbb{Z}^d$ and z a non-negative activity function. Set

$$\tilde{T}(D; z) := 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1, \ldots, Y_k) \in \mathbb{X}^k} \big|\varphi^{\mathsf{T}}_{1+k}(D, Y_1, \ldots, Y_k)\big| z(Y_1) \cdots z(Y_k)$$

and for $N \in \mathbb{N}$,

$$\begin{aligned}&\tilde{T}_N(D; z) := \mathbb{1}_{\{|D| \leq N\}}\\ &\quad + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1, \ldots, Y_k) \in \mathbb{X}^k} \mathbb{1}_{\{|D| + \sum_{i=1}^{k} |Y_i| \leq N\}} \varphi^{\mathsf{T}}_{1+k}(D, Y_1, \ldots, Y_k) z(Y_1) \cdots z(Y_k).\end{aligned} \tag{3.10}$$

Furthermore, we use the convention $\tilde{T}_N(\varnothing; z) = 1$ for all $N \geq 1$.

Again, we lift the established recurrent relations on the level of coefficients (given by Lemma 3.11) to the level of partial sums and series, deriving a system of integral equations for those. The following result is an analogue of Proposition 3.6 and Proposition 3.10 for subset polymers.

Proposition 3.12 *Under the assumptions of Lemma 3.11, the identities*

$$\tilde{T}(D' \cup \{x\}; z) = \tilde{T}(D'; z) + \sum_{\substack{Y \ni x:\\ Y \cap D' = \varnothing}} z(Y) \tilde{T}(D' \cup Y; z)$$

and for $N \in \mathbb{N}$,

$$\tilde{T}_{N+1}(D' \cup \{x\}; z) = \tilde{T}_N(D'; z) + \sum_{\substack{Y \ni x:\\ Y \cap D' = \varnothing}} z(Y) \tilde{T}_N(D' \cup Y; z),$$

hold for any non-negative activity z.

Remark 3.2 Notice that the first identity in Proposition 3.12 is just a sign-flipped version of the standard Kirkwood–Salsburg equations for the reduced correlation functions found in [1].

Proof Lemma 3.11 yields

$$\begin{aligned}
&\mathbb{1}_{\{|D' \cup \{x\}| + \sum_{i=1}^k |Y_i| \leq N+1\}} \varphi_{1+k}^{\mathsf{T}}(D' \cup \{x\}, Y_1, \ldots, Y_k) \\
&\quad = \mathbb{1}_{\{|D'| + \sum_{i=1}^k |Y_i| \leq N\}} \varphi_{1+k}^{\mathsf{T}}(D', Y_1, \ldots, Y_k) \\
&\qquad + \sum_{i=1}^k \left(-\mathbb{1}_{\{Y_i \ni x,\ Y_i \cap D' = \varnothing\}}\right) \mathbb{1}_{\{|D' \cup Y_i| + \sum_{j \neq i} |Y_j| \leq N\}} \varphi_k^{\mathsf{T}}(D' \cup Y_i, (Y_j)_{j \neq i}).
\end{aligned}$$

The proof of the recurrence relation for $\tilde{T}_N(\cdot; z)$ is concluded by exploiting the alternating sign property of the Ursell functions, summing over k and $Y_1, \ldots, Y_k$, and exploiting the symmetry of φ_k^{T}. The recurrence relation for $\tilde{T}(\cdot; z)$ follows by passing to the limit $N \to \infty$. □

Proof of Theorem 2.7 To prove the implication $(i) \Rightarrow (ii)$, suppose that condition (i) is satisfied for some set function $a(\cdot)$. Proceeding as in the proof of Theorem 2.1 again, we prove by induction over N that

$$\tilde{T}_N(D; z) \leq \exp(a(D)), \tag{3.11}$$

for all finite subsets $D \subset \mathbb{Z}^d$. For $N = 1$, the inequality reads $\mathbb{1}_{\{|D| \leq 1\}} \leq \exp(a(D))$ and it is true because $a(D) \geq 0$. Now, suppose it holds true for some $N \geq 1$ and all D. Let $\widehat{D} \subset \mathbb{Z}^d$ be finite. If $\widehat{D}$ is empty, then $\tilde{T}_{N+1}(\widehat{D}; z) = 1 \leq \exp(a(\widehat{D}))$. If $\widehat{D}$ is not empty, let x be any element of $\widehat{D}$ and let $D' := \widehat{D} \backslash \{x\}$. Then Proposition 3.12 yields

$$\tilde{T}_{N+1}(\widehat{D}; z) = \tilde{T}_N(D'; z) + \sum_{\substack{Y \ni x,\\ Y \cap D' = \varnothing}} z(Y) \tilde{T}_N(D' \cup Y; z).$$

By the induction hypothesis and condition (2.18),

$$\tilde{T}_{N+1}(\widehat{D}; z) \leq \mathrm{e}^{a(D')} + \sum_{\substack{Y \ni x,\\ Y \cap D' = \varnothing}} z(Y) \mathrm{e}^{a(D' \cup Y)} \leq \mathrm{e}^{a(D' \cup \{x\})} = \mathrm{e}^{a(\widehat{D})}.$$

This completes the inductive proof of (3.11). Passing to the limit $N \to \infty$, we get $\tilde{T}(D; z) \leq \exp(a(D)) < \infty$.
To prove the converse implication $(ii) \Rightarrow (i)$, suppose that $T(D; z)$ is absolutely convergent for all finite subsets D. Then $\tilde{T}(D; z) < \infty$ and Proposition 3.12 yields

$$\tilde{T}(D \cup \{x\}; z) = \tilde{T}(D; z) + \sum_{\substack{Y \ni x, \\ Y \cap D = \varnothing}} z(Y)\tilde{T}(D \cup Y; z).$$

Set $a(D) := \log \tilde{T}(D; z)$. Then $a(D) \geq 0$ because $\tilde{T}(D; z) \geq 1$, moreover

$$\mathrm{e}^{a(D \cup \{x\})} = \mathrm{e}^{a(D)} + \sum_{\substack{Y \ni x, \\ Y \cap D = \varnothing}} z(Y)\mathrm{e}^{a(D \cup Y)}$$

and the inequality (2.18) follows. □

Notice that the preceeding results of Sect. 3.4 can be generalized by proving a more general version of Lemma 3.11—a direct analogue of Lemma 3.9, where we consider two arbitrary finite subsets $D_0 \subset \mathbb{Z}^d$ and $D_1 \subset \mathbb{Z}^d$, instead of the special case where one of the subsets is a monomer. Naturally, one can view configurations of polymers not only as configurations of monomers but as configurations of disjoint snippets of arbitrary shape and derive from the generalized version of Lemma 3.11 a system of Kirkwood–Salsburg equations different from the one in Proposition 3.12, equations which involve terms of higher order in the activity z. Those equations in turn lead to the following alternative for Theorem 2.7:

Theorem 3.13 *Let $(z(X))_{X \in \mathbb{X}}$ be a non-negative activity. The following two conditions are equivalent:*

(i) *There exists a function $a(\cdot)$ from the finite subsets of $\mathbb{Z}^d$ to $[0, \infty)$ such that $a(\varnothing) = 0$ and the following system of inequalities is satisfied: For all finite, non-empty subsets $D \subset \mathbb{Z}^d$ there exists a subset $D_0 \subset D$ such that—setting $D_1 := D \backslash \{D_0\}$ — we have*

$$\sum_{k \geq 1} \sum_{\{Y_1, \ldots, Y_k\} \subset X} z(Y_1) \ldots z(Y_k)\mathrm{e}^{a(D_1 \cup Y_1 \cup \ldots \cup Y_k) - a(D_1)} \leq \mathrm{e}^{a(D_1 \cup D_0) - a(D_1)} - 1,$$

where the sum runs over sets of mutually disjoint polymers $\{Y_1, \ldots, Y_k\} \subset \mathbb{X}$ such that $Y_i \cap D_0 \neq \varnothing$ and $Y_i \cap D_1 = \varnothing$ for all $i \in \{1 \ldots, k\}$.

(ii) *$T(D; z)$ is absolutely convergent for all finite subsets $D \subset \mathbb{Z}^d$.*

Moreover, if one of the equivalent conditions (hence, both) holds true, then, for all finite subsets $D \subset \mathbb{Z}^d$, we have

$$\left|\log T(D; z)\right| \leq \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1, \ldots, Y_k) \in \mathbb{X}^k} \mathbb{1}_{\{\exists i:\, Y_i \cap D \neq \varnothing\}} \left|\varphi_k^T(Y_1, \ldots, Y_k)\right| z(Y_1) \cdots z(Y_k) \leq a(D).$$

The details of the proof are left for the reader as an exercise. Notice that the sufficient condition for convergence given by Theorem 3.13 is more general than the one given by Theorem 2.7. However, all the proofs of sufficient conditions for systems of subset polymers in Sect. 4 are using the special case of Theorem 2.7.

4 Application to Concrete Hard-Core Models

Our main results (Theorems 2.1, 2.6 and 2.7) provide characterizations of the domain of absolute convergence for the activity expansions $\rho_n(x_1, \ldots, x_n; z)$ from which well-known classical criteria are easily recovered (Corollaries 2.2, 2.3 and 2.8). In this section, we illustrate how our convergence conditions provide new, "practitioner-type" sufficient conditions in concrete hard-core models, both discrete and continuous. Our goal here is not to improve on the best available conditions, but to provide upper bounds on the convergence radii that are of reasonable computational feasibility. In the one-dimensional setup of the Tonks gas, however, we are able to go as far as to recover the characterization of absolute convergence from [13].

4.1 Single-Type Subset Polymers in $\mathbb{Z}^d$

Consider the setup of subset polymers from Chapter 2.4. Suppose there is some finite non-empty set $S \subset \mathbb{Z}^d$ and a scalar $z > 0$ such that

$$z(X) = \begin{cases} z, & X \text{ is a translate of } S, \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

We call polymers with non-zero activity *active polymers*. Define

$$V(D) := \left|\{X \in \mathbb{X} \mid z(X) > 0,\ X \cap D \neq \varnothing\}\right|,$$

the number of active polymers intersecting a finite domain $D \subset \mathbb{Z}^d$. Notice $V(\{x\}) = |S|$, for all $x \in \mathbb{Z}^d$.

Theorem 4.1 *Let $z(\cdot)$ be the activity function from (4.1). Suppose there exists $\alpha > 0$ such that*

$$|S|\, e^{\alpha V(S)} z \leq e^{\alpha |S|} - 1. \tag{4.2}$$

Then $T(D; z)$ is absolutely convergent for all finite subsets $D \subset \mathbb{Z}^d$, thus the activity expansions $\rho_n(x_1, \ldots, x_n; z)$ converge absolutely for all $n \in \mathbb{N}$ and all $(x_1, \ldots, x_n) \in \mathbb{X}^n$.

Remark 4.1 Notice that Theorem 4.1 improves on the upper bounds for the convergence radii given by Kotecký–Preiss and by Gruber–Kunz. The improvement over the Gruber–Kunz condition is achieved by a more sophisticated choice of the ansatz function a in the proof of the theorem. However, although we do not have a general proof that the result by Fernández–Procacci is stronger, notice that for all non-pathological examples we considered (e.g., non-overlapping dimers or cubes) Fernández–Procacci provides better bounds than Theorem 4.1.

Example 4.1 (Hypercubes) If $S = \{1, \ldots, k\}^d$ with $k \in \mathbb{N}$, condition (4.2) becomes

$$z \leq \sup_{\alpha > 0} \frac{\exp(\alpha k^d) - 1}{k^d \exp(\alpha (2k-1)^d)}.$$

Carrying out the optimization over α yields the condition

$$(2k-1)^d z \leq \left(1 - \frac{1}{(2 - 1/k)^d}\right)^{(2-1/k)^d - 1}.$$

In the limit $d \to \infty$ at fixed $k \geq 2$, the right-hand side converges from above to the familiar bound $1/\mathrm{e}$.

Proof of Theorem 4.1 We apply Theorem 2.7 with $a(D) := \alpha V(D)$. We check that $V(\cdot)$ is strongly subadditive. Let B, C be finite subsets of $\mathbb{Z}^d$. Then for every polymer X,

$$\mathbb{1}_{\{X\cap B\neq\varnothing\}} + \mathbb{1}_{\{X\cap C\neq\varnothing\}} \geq \mathbb{1}_{\{X\cap(B\cup C)\neq\varnothing\}} + \mathbb{1}_{\{X\cap(B\cap C)\neq\varnothing\}}.$$

Indeed if X intersects B but not C (or C but not B), the inequality reads $1+0 \geq 1+0$ and it is true. If X intersects both B and C, the inequality reads $1+1 \geq 1 + \mathbb{1}_{\{X\cap(B\cap C)\neq\varnothing\}}$ and it is true as well. Finally if X intersects neither B nor C, then both sides of the inequality vanish. Summing over all polymers X, we get

$$V(B) + V(C) \geq V(B\cup C) + V(B\cap C). \tag{4.3}$$

Now we turn to the criterion (i) from Theorem 2.7. Condition (2.18) for $D' = \varnothing$ reads

$$|S|\, z\, \mathrm{e}^{\alpha V(S)} \leq \mathrm{e}^{\alpha V(\{x\})} - 1,$$

it is satisfied because of $V(\{x\}) = |S|$ and the assumption (4.2). For non-empty D', we bound the left-hand side of condition (2.18) with the help of the strong subadditivity. The inequality (4.3) applied to $B = D' \cup \{x\}$ and $C = X$ yields

$$V(D'\cup\{x\}) + V(X) \geq V(D'\cup X) + V(\{x\}), \tag{4.4}$$

for $x \in \mathbb{Z}^d \setminus D'$, $x \in X$, and $X \cap D' = \varnothing$, and

$$\begin{aligned} V(D'\cup X) - V(D') &\leq V(D'\cup\{x\}) + V(X) - V(D') - V(\{x\}) \\ &= V(D'\cup\{x\}) - V(D') + V(S) - |S|. \end{aligned}$$

This provides an X-independent bound for the exponent in the left-hand side of condition (2.18). The number of summands on the left-hand side of condition (2.18) is given by the number of active polymers intersecting x but not D', which is equal to $V(D'\cup\{x\}) - V(D')$. Thus to prove (2.18) it suffices to show that

$$\big(V(D'\cup\{x\}) - V(D')\big) z\, \mathrm{e}^{\alpha[V(D'\cup\{x\}) - V(D') + V(S) - |S|]} \leq \mathrm{e}^{\alpha[V(D'\cup\{x\}) - V(D')]} - 1. \tag{4.5}$$

In view of condition (4.2), the last inequality in turn follows once we check

$$\big(V(D'\cup\{x\}) - V(D')\big)\big(\mathrm{e}^{\alpha|S|} - 1\big)\mathrm{e}^{\alpha[V(D'\cup\{x\}) - V(D') - |S|]} \leq |S|\big(\mathrm{e}^{\alpha[V(D'\cup\{x\}) - V(D')]} - 1\big)$$

or equivalently,

$$\frac{1-\exp(-\alpha|S|)}{|S|} \leq \frac{1-\exp(-\alpha R)}{R}, \quad R := V(D'\cup\{x\}) - V(D').$$

Because of the subadditivity of V, we have $R \geq V(\{x\}) = |S|$. The exponential map $x \mapsto \exp(-\alpha x)$ is convex and therefore the difference quotient is monotone increasing, i.e., $(\exp(-\alpha x) - 1)/x \leq (\exp(-\alpha y) - 1))/y$ whenever $0 \leq x \leq y$. We apply the inequality to $x = R$ and $y = |S|$ and obtain the required bound. □

4.2 Single-Type Hard-Core System in $\mathbb{R}^d$

Consider a bounded convex shape $S \subset \mathbb{R}^d$ which is non-empty, regular closed and balanced (recall: A set $S \subset \mathbb{R}^d$ is called regular closed if and only if it equals the closure of its interior, i.e., $\overline{S^\circ} = S$, and it is called balanced if and only if $\alpha S \subset S$ for all $|\alpha| \leq 1$). We investigate the special case of the hard-core setup in the continuum from Sect. 2.3 where $\mathbb{X}$ consists of all translates $x + S = \{x + y \mid y \in S\}$. Let us further assume that both the activity and the reference measure λ are translationally invariant. Then we may identify $\mathbb{X}$ with $\mathbb{R}^d$, the reference measure λ with the Lebesgue measure, and the activity function with a positive scalar, $z(x) \equiv z > 0$.

For an integrable function $h : \mathbb{X} \to \mathbb{R}$ we write

$$\int_{\mathbb{X}} h(Z)\lambda(\mathrm{d}Z) = \int_{\mathbb{R}^d} h(x + S)\mathrm{d}x.$$

Write $|S|$ for the Lebesgue volume of the shape S and define for Borel sets $D \subset \mathbb{R}^d$

$$V(D) := \int_{\mathbb{R}^d} \mathbb{1}_{\{(x+S)\cap D \neq \varnothing\}}\mathrm{d}x. \tag{4.6}$$

Notice $V(\{y\}) = |S|$, which is positive and finite by our assumptions on S, and $V(S) = |S \oplus S|$ with $A \oplus B := \{a + b \mid a \in A,\ b \in B\}$ the Minkowski sum. The latter identity holds since we assumed the set S to be balanced which implies $\{(x + S) \cap S \neq \varnothing\} = S \oplus S$. Moreover, notice that V—as a function on $\mathbb{D}_\varepsilon$ with the convention $V(\varnothing) = 0$—satisfies the measurability assumption from Theorem 2.6 by the same argument as formulated on p. 22 for $\tilde{T}$.

We refer to such systems as *single-type hard-core systems in the continuum.* In the language of stochastic geometry (see [28, Sects. 3,4]), the associated Gibbs measure is a *hard-core germ–grain model* with deterministic grain S (the germs are the positions x).

Theorem 4.2 *Assume there exists* $\alpha > 0$ *such that*

$$|S|\mathrm{e}^{\alpha V(S)} z < \mathrm{e}^{\alpha|S|} - 1. \tag{4.7}$$

Then the activity expansions $\rho_n(x_1, \ldots, x_n; z)$ *converge absolutely for all* $n \in \mathbb{N}$ *and all* $(x_1, \ldots, x_n) \in \mathbb{X}^n$.

Remark 4.2 Again, notice that while Theorem 4.2—just as its discrete analogue Theorem 4.1)—improves on the Kotecký–Preiss condition, it is in all the cases we considered as examples weaker than Fernández–Procacci (e.g., for systems of hard spheres the bounds on the radius of convergence obtained in [9, 22] are slightly better). However, the advantage of our criterion is that an explicit bound is provided directly, with no need for numerical computation regardless of the dimension.

Example 4.2 (Hard spheres) If $S = B_R(0)$ is the closed ball of radius $R > 0$ around the origin, condition (4.7) becomes

$$z \leq \sup_{\alpha>0} \frac{\exp(\alpha|B_R(0)|) - 1}{|B_R(0)|\ \exp(\alpha|B_{2R}(0)|)}.$$

Carrying out the optimization over α yields the condition

$$|B_{2R}(0)|\, z \leq \left(1 - \frac{1}{2^d}\right)^{2^d - 1}.$$

In the limit $d \to \infty$ at fixed $R > 0$, the right-hand side converges from above to the familiar bound $1/\mathrm{e}$.

Proof Let $\alpha > 0$ satisfy condition (4.7). Set $a(\widehat{D}) := \alpha V(\widehat{D})$ for $\widehat{D} \in \mathbb{D}_\varepsilon$, choose some chopping map C and let $D \in \mathbb{D}_\varepsilon$ (the snippet-size $\varepsilon > 0$ will be specified later in the proof). For the simplest selection rule s choosing always the first snippet E_1, condition (i) in Theorem 2.6 reads

$$\sum_{k=1}^{\infty} \frac{z^k}{k!} \int_{\mathbb{X}^k} I(E_1; D'; Y_1, \ldots, Y_k)\, \mathrm{e}^{\alpha[V(D' \cup (\cup_{i=1}^k Y_i)) - V(D')]} \lambda^k(\mathrm{d}\boldsymbol{Y}) \\ \leq \mathrm{e}^{\alpha[V(D' \cup E_1) - V(D')]} - 1, \tag{4.8}$$

where $D' = D \backslash E_1$.

Notice that—unlike in the discrete case—terms of order higher than one in z do not necessarily vanish in the series in (4.8). Inspired by the proof of Theorem 4.7, we try first to bound the exponent on the left-hand side of (4.8), seeking a bound that separates $Y := Y_1 \cup \cdots \cup Y_k$ from E_1 and D'. If the constraint were that $E_1 \subset Y$, we would conclude with strong subadditivity applied to $B := Y$ and $C := D' \cup E_1$ that $V(E_1) + V(D' \cup Y) \leq V(Y) + V(D' \cup E_1)$. For the weaker constraint $E_1 \cap Y \neq \varnothing$, this is no longer true. Let $Z \in \mathbb{X}$. A straightforward case distinction reveals that under the indicator I, the inequality

$$\mathbb{1}_{\{Z \cap E_1 \neq \varnothing\}} + \mathbb{1}_{\{Z \cap (D' \cup Y)\} \neq \varnothing\}} \leq \mathbb{1}_{\{Z \cap Y \neq \varnothing\}} + \mathbb{1}_{\{Z \cap (D' \cup E_1) \neq \varnothing\}}$$

is correct for all possible values of the left-hand side except possibly $1 + 1$. Indeed it may happen that Z intersects both E_1 and $D' \cup Y$, hence a fortiori $D' \cup E_1$, but not Y, so that the right-hand side becomes $0 + 1$. This happens precisely when Z intersects D' and E_1 but not Y. The inequality becomes correct if we add the indicator of this event to the right-hand side. Integrating over Z, we obtain

$$V(E_1) + V(D' \cup Y) \leq V(Y) + V(D' \cup E_1) + \int_{\mathbb{X}} \mathbb{1}_{\{Z \cap E_1 \neq \varnothing,\, Z \cap Y = \varnothing,\, Z \cap D' \neq \varnothing\}} \lambda(\mathrm{d}Z).$$

Moreover, there exists a constant $C = C(S, d) > 0$ that depends only on the dimension d and the shape S such that if $k = 1$ and $Y = Y_1 \in \mathbb{X}$ is a translate of S, then

$$\int_{\mathbb{X}} \mathbb{1}_{\{Z \cap E_1 \neq \varnothing,\, Z \cap Y = \varnothing,\, Z \cap D' \neq \varnothing\}} \lambda(\mathrm{d}Z) \leq C\varepsilon.$$

Indeed, on the left side we may drop the indicator that Z intersects D' and see that it is sufficient to check

$$\int_{\mathbb{X}} \mathbb{1}_{\{Z \cap E_1 \neq \varnothing,\, Z \cap Y = \varnothing\}} \lambda(\mathrm{d}Z) \leq C\varepsilon. \tag{4.9}$$

To see that such an estimate holds let $B_\varepsilon(c)$ be a closed ball of radius ε around some $c \in \mathbb{R}^d$ containing the snippet E_1 and let $x \in Y \cap E_1$ (by assumption this intersection is non-empty). Then $x \in B_\varepsilon(c)$ and the inequality

$$\mathbb{1}_{\{Z \cap E_1 \neq \varnothing,\, Z \cap Y = \varnothing\}} \leq \mathbb{1}_{\{Z \cap B_\varepsilon(c) \neq \varnothing,\, x \notin Z\}}$$

holds pointwise in Z, thus also

$$\int_{\mathbb{X}} \mathbb{1}_{\{Z \cap E_1 \neq \varnothing,\, Z \cap Y = \varnothing\}} \lambda(\mathrm{d}Z) \leq \int_{\mathbb{X}} \mathbb{1}_{\{Z \cap B_\varepsilon(c) \neq \varnothing,\, x \notin Z\}} \lambda(\mathrm{d}Z).$$

Notice that

$$\int_{\mathbb{X}} \mathbb{1}_{\{Z\cap B_\varepsilon(c)\neq\varnothing,\, x\notin Z\}}\lambda(\mathrm{d}Z) = |\{y\in\mathbb{R}^d | (y+S)\cap B_\varepsilon(0)\neq\varnothing, \tilde{x}\notin y+S\}|,$$

where $B_\varepsilon(0)$ is the closed ball of radius ε around 0 and $\tilde{x} := x - c$. For the set on the right-hand side of that equation, the identity

$$\{y\in\mathbb{R}^d | (y+S)\cap B_\varepsilon(0)\neq\varnothing, \tilde{x}\notin y+S\} = (S\oplus B_\varepsilon(0))\setminus(\tilde{x}+S),$$

holds since S being balanced directly implies

$$\{y\in\mathbb{R}^d | (y+S)\cap B_\varepsilon(0)\neq\varnothing\} = S\oplus B_\varepsilon(0)$$

and

$$\{y\in\mathbb{R}^d | \tilde{x}\in y+S\} = \tilde{x}+S.$$

Furthermore, observe that the inclusion

$$(S\oplus B_\varepsilon(0))\setminus(\tilde{x}+S) \subset ((S\oplus B_\varepsilon(0))\setminus S)\cup((S\oplus B_\varepsilon(\tilde{x}))\setminus(\tilde{x}+S))$$

holds since $\tilde{x}\in B_\varepsilon(0)$ and S is balanced set. Moreover, $(S\oplus B_\varepsilon(\tilde{x}))\setminus(\tilde{x}+S)$ is the translate of $(S\oplus B_\varepsilon(0))\setminus S$ by $\tilde{x}$, hence it has the same Lebesgue volume and

$$\begin{aligned}
|(S\oplus B_\varepsilon(0))\setminus(\tilde{x}+S)| &\leq |(S\oplus B_\varepsilon(0))\setminus S| + |(S\oplus B_\varepsilon(\tilde{x}))\setminus(\tilde{x}+S)|\\
&= 2|(S\oplus B_\varepsilon(0))\setminus S|\\
&= 2\left(|S\oplus B_\varepsilon(0)| - |S|\right).
\end{aligned}$$

Finally, by Steiner's formula for compact convex sets (see [28]), $|S\oplus B_\varepsilon(0)| - |S|$ is given by a non-constant polynomial in ε, which yields a bound of the form given by the right-hand side of (4.9) (where the constant $C > 0$ can be expressed in terms of the intrinsic volumes of S following the formula).

Consequently, we obtain the bound

$$V(E_1) + V(D'\cup Y) \leq V(Y) + V(D'\cup E_1) + C\varepsilon \tag{4.10}$$

which corresponds to the bound (4.4) in the proof of Theorem 4.1.

The inequality (4.10) immediately yields the following upper bound for the left-hand side of (4.8):

$$\mathrm{e}^{\alpha C\varepsilon}\,\mathrm{e}^{\alpha[V(D'\cup E_1)-V(D')-V(E_1)]}\sum_{k=1}^{\infty}\frac{z^k}{k!}\int_{\mathbb{X}^k} I(E_1; D'; Y_1,\dots,Y_k)\,\mathrm{e}^{\alpha V(Y)}\lambda^k(d\boldsymbol{Y}).$$

The summand for $k=1$ is equal to

$$z\,\mathrm{e}^{\alpha V(S)}\int_{\mathbb{X}} \mathbb{1}_{\{Y_1\cap E_1\neq\varnothing, Y_1\cap D'=\varnothing\}}\lambda(\mathrm{d}Y_1) = z[V(D'\cup E_1) - V(D')]\mathrm{e}^{\alpha V(S)}.$$

For $k\geq 2$, we bound $V(Y)\leq\sum_{i=1}^k V(Y_i) = kV(S)$, drop the indicator that the Y_i's do not intersect D', and get the upper bound

$$z^k\mathrm{e}^{\alpha kV(S)}\int_{\mathbb{X}^k}\prod_{i=1}^k \mathbb{1}_{\{Y_i\cap E_1\neq\varnothing\}}\mathbb{1}_{\{Y_1,\dots,Y_k \text{ disjoint}\}}\lambda^k(\mathrm{d}\boldsymbol{Y}).$$

Notice that there exists $N\in\mathbb{N}$ such that for all $k\geq N+1$ the integral vanishes. To see this, assume that there are infinitely many disjoint objects $Y\in\mathbb{X}$ intersecting the snippet E_1

(and therefore some open ε-ball B_ε in which the snippet is contained). Since all the objects Y are translates of S, we can choose the same radius $r > 0$ for all of the infinitely many disjoint objects Y intersecting E_1 such that $Y = x + S \subset B_r(x)$. Naturally, every such r-ball must intersect B_ε and therefore their union is again a bounded Borel subset of $\mathbb{R}^d$. But — by our assumptions on the shape S—every $Y \in \mathbb{X}$ has the same fixed, strictly positive Lebesgue measure, thus their disjoint union must have infinite Lebesgue measure, which is a contradiction to its boundedness.

For $k \leq N$, we drop the indicator that $Y_3, \dots, Y_k$ are disjoint and find that the integral is bounded by

$$V(E_1)^{k-2} \int_{\mathbb{X}} \mathbb{1}_{\{Y_1 \cap E_1 \neq \varnothing\}} \Big(\int_{\mathbb{X}} \mathbb{1}_{\{Y_2 \cap E_1 \neq \varnothing,\, Y_2 \cap Y_1 = \varnothing\}} \lambda_z(\mathrm{d}Y_2) \Big) \lambda_z(\mathrm{d}Y_1).$$

The inner integral is bounded by $C\varepsilon$ because of (4.9), the outer integral gives an additional factor $V(E_1)$. Altogether, the left-hand side of (4.8) is bounded by

$$\mathrm{e}^{\alpha C\varepsilon}\, \mathrm{e}^{\alpha[V(D' \cup E_1) - V(D') - V(E_1)]} \Big(z[V(D' \cup E_1) - V(D')]\mathrm{e}^{\alpha V(S)} + C\varepsilon \sum_{k=2}^{N} z^k V(E_1)^{k-1} \mathrm{e}^{\alpha k V(S)} \Big).$$

Proceeding as in the proof of Theorem 4.2, but taking into account the strict inequality from assumption (4.7), we find that there exist $\alpha > 0$ such that

$$\mathrm{e}^{\alpha[V(D' \cup E_1) - V(D') - V(E_1)]} z[V(D' \cup E_1) - V(D')]\mathrm{e}^{\alpha V(S)} < \mathrm{e}^{\alpha[V(D' \cup E_1) - V(D)]} - 1,$$

compare to (4.5). Therefore, picking ε small enough, we see that (4.8), hence also condition (i) in Theorem 2.6 is satisfied and all $T(D; z)$, $D \in \mathbb{D}_\varepsilon$, are absolutely convergent. The claim of the theorem follows immediately. □

4.3 Multi-Type Hard Spheres in $\mathbb{R}^d$

Let $(r_n)_{n \in \mathbb{N}}$ be an increasing sequence of positive real numbers and let $B_{r_n}(0) \subset \mathbb{R}^d$, $n \in \mathbb{N}$, be the family of d-dimensional closed balls around 0 with the corresponding radii. To the sequence $(r_n)_{n \in \mathbb{N}}$ of radii is associated the sequence of non-negative activities $(z_n)_{n \in \mathbb{N}}$ such that the ball B_{r_i} has the activity z_i. In the setup of hard-core systems in the continuum from Sect. 2.3, let $\mathbb{X}$ be given by all possible translates of these objects. Notice that the closed balls are compact convex sets that are non-empty and regular closed. We refer to this special case of a hard-core system as a system of *multi-type hard spheres* in $\mathbb{R}^d$. We will show a new sufficient condition for absolute convergence of the activity expansions in these types of models.

For an integrable function $h : \mathbb{X} \to \mathbb{R}$ we write

$$\int_{\mathbb{X}} h(Z)\lambda(\mathrm{d}Z) = \sum_{\ell \geq 1} \int_{\mathbb{R}^d} h(x + B_{r_\ell}(0))\mathrm{d}x.$$

We define the family of functions $(V_r)_{r>0}$ by setting for Borel sets $D \subset \mathbb{R}^d$

$$V_r(D) := \int_{\mathbb{R}^d} \mathbb{1}_{\{(x + B_r(0)) \cap D \neq \varnothing\}} \mathrm{d}x, \tag{4.11}$$

where $B_r(0)$ is the d-dimensional closed ball of radius $r > 0$ around 0. Naturally, the map V_r coincides with the map V from (4.6) for the grain S given by the closed ball $B_r(0)$ and therefore satisfies the measurability assumption from Theorem 2.6 (as a function on $\mathbb{D}_\varepsilon$ with the convention $V(\varnothing) = 0$). Furthermore, we have $V_r(\{y\}) = |B_r(0)|$ and $V_r(B_s(y)) = |B_s(y) \oplus B_r(0)| = |B_{s+r}(0)|$ (where $\oplus$ denotes the Minkowski sum) for any $y \in \mathbb{R}^d$ and any numbers $r, s > 0$.

The following auxiliary result turns out to be essential for the proof of the new sufficient condition:

Lemma 4.3 *Let D_1 be a finite union of bounded convex regular closed subsets of $\mathbb{R}^d$ and let D_2 be a d-dimensional ball in $\mathbb{R}^d$. The map $(0, \infty) \ni r \mapsto \frac{V_r(D_1 \cup D_2) - V_r(D_1)}{V_r(D_2)}$ is monotonically decreasing in r.*

Proof First of all, observe that for sets D given by a finite union of convex, regular closed subsets of $\mathbb{R}^d$ the volume $V_r(D)$ can be written as

$$V_r(D) = |D| + S(D)r + o(r), \tag{4.12}$$

where $S(D)$ denotes the surface area of D. This follows from a generalized version of the classical Steiner's formula (see [27, Sect. 4.4]). In particular, we see that the map $r \mapsto \frac{V_r(D_1 \cup D_2) - V_r(D_1)}{V_r(D_2)}$ is differentiable in $r = 0$.

Next, we notice that the map satisfies the following *semi-group property*:

$$V_{r+\varepsilon}(D) = V_\varepsilon(D \oplus B_r(0)).$$

Therefore, to prove the claim of the lemma, it suffices to consider the differential at zero:

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \left(\frac{V_\varepsilon(A \cup B) - V_\varepsilon(A)}{V_\varepsilon(B)} - \frac{|A \cup B| - |A|}{|B|} \right),$$

where $A := D_1 \oplus B_r(0)$ and $B := D_2 \oplus B_r(0)$.

Using the formula (4.12), a simple computation shows that this limit is equal to

$$\frac{|B|\,(S(A \cup B) - S(A)) - S(B)\,(|A \cup B| - |A|)}{|B|^2}.$$

The monotonicity in the claim of the lemma is then equivalent to

$$|B|\,(S(A \cup B) - S(A)) - S(B)\,(|A \cup B| - |A|) \le 0$$

or, equivalently,

$$\frac{|B|}{S(B)} \le \frac{|A \cup B| - |A|}{S(A \cup B) - S(A)}.$$

Using the obvious identities $|A \cup B| - |A| = |B| - |A \cap B|$ and $S(A \cup B) - S(A) = S(B) - S(A \cap B)$, we can rewrite the last inequality as

$$\frac{S(B)}{|B|} \le \frac{S(A \cap B)}{|A \cap B|},$$

which holds by the isoperimetric inequality since $B = D_2 \oplus B_r(0)$ is a ball in $\mathbb{R}^d$ ("the ball is the shape that minimizes the surface area for given volume", see [7, 3.2.43]). □

The following sufficient condition is, in some sense, a "continuous version" of the Gruber–Kunz criterion in the setup of hard spheres in $\mathbb{R}^d$. The similarity in the form arises as follows: To establish the recurrence relations underlying the proof of Gruber–Kunz we selected a monomer, a single point in $\mathbb{Z}^d$, from a configuration of polymers. We follow this idea in the proof of the following result, choosing the chopping map C and the selection rule s such that a tiny snippet that approximates a single point in the continuous space sufficiently well is selected. At the same time we choose an ansatz function a that can be interpreted as the continuous analogue of the ansatz function from the proof of Corollary 2.8.

Theorem 4.4 *In the setup of multi-type hard objects on $\mathbb{R}^d$, assume that the activity z satisfies*

$$\exists \alpha > 0: \quad \sum_{\ell \geq 1} |B_{r_\ell}| \mathrm{e}^{\alpha |B_{r_\ell + r_1}|} z_\ell < e^{\alpha |B_{r_1}|} - 1, \tag{4.13}$$

where by $|B_r|$ we denote the (Lebesgue) volume of a ball of radius $r > 0$. Then the activity expansions $\rho_n(X_1, \ldots, X_n; z)$ converge absolutely.

Remark 4.3 Activity expansions for systems with infinitely many types of objects are not particularly well-studied in statistical mechanics. In the case of finitely many types, we expect our result to exceed Kotecký–Preiss but to be weaker then Fernández–Procacci—as in the special case of a single type treated above (see Remark 4.2). The general case, for $r_n \to \infty$ in particular, remains to be investigated.

Proof Again, our strategy is to show that condition (i) from Theorem 2.6 is satisfied for an appropriate ansatz function a. By assumption, r_1 is the radius of the smallest ball present in the system. Set $a(\widehat{D}) := \alpha V_{r_1}(\widehat{D})$ for $\widehat{D} \in \mathbb{D}_\varepsilon$, where V_{r_1} is given by (4.11) and α satisfies (4.13). Choose some chopping map C and let $D \in \mathbb{D}_\varepsilon$ (the snippet-size $\varepsilon > 0$ is to be specified later in the proof). Just as in the proof of Theorem 4.2, independently of the choice of the snippet E_1 (i.e., independently of the selection rule s), we obtain the following upper bound for the left-hand side of the inequality from condition (i):

$$\mathrm{e}^{\alpha C_1(\varepsilon)} \, \mathrm{e}^{\alpha [V(D' \cup E_1) - V(D') - V(E_1)]} \sum_{k=1}^{\infty} \frac{z^k}{k!} \int_{\mathbb{X}^k} I(E_1; D', Y_1, \ldots, Y_k) \, \mathrm{e}^{\alpha V(\boldsymbol{Y})} \lambda^k(\mathrm{d}\boldsymbol{Y}). \tag{4.14}$$

The positive number $C_1(\varepsilon)$ converges towards 0 for $\varepsilon \searrow 0$ and is precisely the bound from (4.9) in the proof of Theorem 4.2 for Y given by a translate of $B_{r_1}(0)$, i.e., by a sphere of minimal volume present in the system.

The summand for $k = 1$ in (4.14) is equal to

$$\sum_{\ell \geq 1} z_\ell \mathrm{e}^{\alpha |B_{r_\ell + r_1}|} \int_{\mathbb{X}} \mathbb{1}_{\{r(Y_1) = r_\ell\}} \mathbb{1}_{\{Y_1 \cap E_1 \neq \varnothing, Y_1 \cap D' = \varnothing\}} \lambda(\mathrm{d}Y_1).$$

Notice that the integrals in the last expression are equal to $[V_{r_\ell}(D' \cup E_1) - V_{r_\ell}(D')]$ for every $\ell \in \mathbb{N}$.

The summand for any $k \geq 2$ in (4.14) is bounded from above by

$$\sum_{\ell_1, \ldots, \ell_k} z_{\ell_1} \ldots z_{\ell_k} \mathrm{e}^{\alpha \sum_{i=1}^{k} |B_{r_{\ell_i} + r_1}|} \int_{\mathbb{X}^k} \prod_{i=1}^{k} \mathbb{1}_{\{r(Y_i) = r_{\ell_i}\}}$$
$$\prod_{i=1}^{k} \mathbb{1}_{\{Y_i \cap E_1 \neq \varnothing, Y_i \cap D' = \varnothing\}} \mathbb{1}_{\{Y_1, \ldots, Y_k \text{ disjoint}\}} \lambda^k(\mathrm{d}\boldsymbol{Y}), \tag{4.15}$$

which—by arguments similar to the ones used for the bound $C_1(\varepsilon)$—is again bounded by

$$C_2(\varepsilon) \sum_{\ell_1, \ldots, \ell_k} z_{\ell_1} \ldots z_{\ell_k} \mathrm{e}^{\alpha \sum_{i=1}^{k} |B_{r_{\ell_i} + r_1}|} |B_{r_{\ell_1}}| \ldots |B_{r_{\ell_k}}| \tag{4.16}$$

for a positive constant $C_2(\varepsilon)$ that is independent of k and satisfies $C_2(\varepsilon) \searrow 0$ for $\varepsilon \searrow 0$. Notice that the sum in (4.16) is finite for every $k \in \mathbb{N}$ by assumption (4.13) [since it is simply given by the k-th power of the left-hand side of the inequality in (4.13)]. Moreover, by the same argument as in the proof of Theorem 4.2, the expression in (4.15) does vanish for all but finitely many $k \in \mathbb{N}$, i.e., there exists a number $N \in \mathbb{N}$ such that (4.15) is equal to zero for all $k \geq N+1$.

Altogether we get the upper bound

$$\mathrm{e}^{\alpha C_1(\varepsilon)}\, \mathrm{e}^{\alpha[V(D'\cup E_1)-V(D')-V(E_1)]} \times \Big(\sum_{\ell\geq 1} z_\ell \mathrm{e}^{\alpha|B_{r_\ell+r_1}|}[V_{r_\ell}(D'\cup E_1) - V_{r_\ell}(D')] + \\ C_2(\varepsilon) \sum_{2\leq k\leq N} \sum_{\ell_1,\dots,\ell_k} z_{\ell_1}\dots z_{\ell_k} \mathrm{e}^{\alpha \sum_{i=1}^{k} |B_{r_{\ell_i}+r_1}|} |B_{r_{\ell_1}}|\dots|B_{r_{\ell_k}}| \Big).$$

As in the single-type case, we see that it is sufficient to prove the strict inequality

$$\mathrm{e}^{\alpha[V(D'\cup E_1)-V(D')-V(E_1)]} \sum_{\ell\geq 1} z_\ell \mathrm{e}^{\alpha|B_{r_\ell+r_1}|}[V_{r_\ell}(D'\cup E_1) - V_{r_\ell}(D')] < \mathrm{e}^{\alpha[V(D'\cup E_1)-V(D)]} - 1$$

for small values of $\varepsilon > 0$ and, consequently, for small volumes of the snippet E_1 contained in an ε-ball.

To do so, we bound $[V_{r_\ell}(D'\cup E_1) - V_{r_\ell}(D')]$ from above by $[V_{r_\ell}(D'\cup B_\varepsilon) - V_{r_\ell}(D')]$ for every $\ell \in \mathbb{N}$, where B_ε is the ball of radius ε containing the snippet E_1. Then we use Lemma 4.3 to obtain

$$[V_{r_\ell}(D'\cup B_\varepsilon) - V_{r_\ell}(D')] \leq [V_{r_m}(D'\cup B_\varepsilon) - V_{r_m}(D')] \frac{V_{r_\ell}(B_\varepsilon)}{V_{r_m}(B_\varepsilon)}$$

for $m \leq \ell$ (since in that case $r_m \leq r_\ell$ holds by assumption) and therefore

$$\sum_{\ell\geq 1} z_\ell \mathrm{e}^{\alpha|B_{r_\ell+r_1}|}[V_{r_\ell}(D'\cup E_1) - V_{r_\ell}(D')] \leq \frac{V_{r_1}(D'\cup B_\varepsilon) - V_{r_1}(D')}{V_{r_1}(B_\varepsilon)} \sum_{\ell\geq 1} z_\ell \mathrm{e}^{\alpha|B_{r_\ell+r_1}|} V_{r_\ell}(B_\varepsilon).$$

By dominated convergence and assumption (4.13) we can choose $\varepsilon > 0$ small enough to strictly bound the right-hand side of the last equation by

$$\frac{V_{r_1}(D'\cup E_1) - V_{r_1}(D')}{V_{r_1}(E_1)} (\mathrm{e}^{\alpha V_{r_1}(E_1)} - 1).$$

Finally, it suffices to show the inequality

$$\mathrm{e}^{\alpha[V(D'\cup E_1)-V(D')-V(E_1)]} \frac{V_{r_1}(D'\cup E_1) - V_{r_1}(D')}{V_{r_1}(E_1)} (\mathrm{e}^{\alpha|V_{r_1}(E_1)|} - 1) \leq \mathrm{e}^{\alpha[V(D'\cup E_1)-V(D)]} - 1 \tag{4.17}$$

as in (4.5) to conclude the proof. □

4.4 Tonks Gas on $\mathbb{Z}$

Next we turn to the discrete one-dimensional Tonks gas with translationally invariant activities. That is, in the setup of subset polymers from Sect. 2.4 for $d = 1$, let $(z_\ell)_{\ell\in\mathbb{N}}$ be a sequence of non-negative numbers and consider the activity

$$z(X) = \begin{cases} z_\ell, & X = \{m, m+1, \dots, m+\ell-1\} \text{ for some } m \in \mathbb{Z}, \\ 0, & \text{else.} \end{cases} \tag{4.18}$$

Theorem 4.5 *Let $d = 1$ and let $(z_\ell)_{\ell\in\mathbb{N}}$ be a sequence of non-negative activities.*

(a) *Suppose there exists $\alpha > 0$ such that*

$$\sum_{\ell=1}^{\infty} \mathrm{e}^{\alpha\ell} z_\ell \leq \mathrm{e}^{\alpha} - 1. \tag{4.19}$$

Then $T(D; z)$ is absolutely convergent for all finite subsets $D \subset \mathbb{Z}$.

(b) *Conversely, if $T(D; z)$ is absolutely convergent for all finite subsets $D \subset \mathbb{Z}$, then there exists $\alpha > 0$ such that (4.19) holds true.*

Remark 4.4 The condition (4.19) is exactly the necessary and sufficient criterion for absolute convergence of the activity expansion of the pressure in the system derived in [13]. While the result itself is not novel, we consider the proof to be instructive since it demonstrates how our approach can provide conditions improving on the Fernández–Procacci criterion. In this concrete setup even the optimal result—recovering the whole domain of convergence—can be achieved.

The proof of the sufficient condition relies on a refinement of Theorem 2.7. Roughly, we weaken condition (i) to consider the Kirkwood–Salsburg inequalities being satisfied only for single rods rather than for arbitrary configurations of rods; at the same time we specify the selection rule by assuming that the leftmost (or, alternatively, the rightmost) element $\{x\}$ is always picked from any given domain.

Proposition 4.6 *Suppose there exists a non-negative function $a(\cdot)$ from the finite intervals of $\mathbb{Z}$ to $[0, \infty)$ with $a(\varnothing) = 0$ and for every finite interval D of $\mathbb{Z}$ with $x = \min D$ such that*

$$\sum_{\substack{Y \ni x, \\ Y \cap D' = \varnothing}} z(Y)\, \mathrm{e}^{a(D'\cup Y) - a(D')} \leq \mathrm{e}^{a(D'\cup\{x\}) - a(D')} - 1, \tag{4.20}$$

where we set $D' = D\backslash\{x\}$. Then $T(D, z)$ is absolutely convergent, for all finite $D \subset \mathbb{Z}$ (interval or not).

Proof We revisit the proof of the implication (i) $\Rightarrow$ (ii) of Theorem 2.7 given on p. 24 and prove first by induction over N that $\tilde{T}_N(D; z) \leq \exp(a(D))$, for all finite discrete intervals $D \subset \mathbb{Z}$. For $N = 1$, the inequality is trivial because $\tilde{T}_1(D; z) = \mathbb{1}_{\{|D|\leq 1\}} \leq 1 \leq \exp(a(\widehat{D}))$. Now, suppose $\tilde{T}_N(D; z) \leq \exp(a(D))$ for some $N \in \mathbb{N}$ and all discrete intervals $D \subset \mathbb{Z}$. Let $\widehat{D} \subset \mathbb{Z}$ be any discrete interval. If $\widehat{D} = \varnothing$, then $\tilde{T}_{N+1}(\widehat{D}; z) = 1 \leq \exp(a(\widehat{D}))$. If $\widehat{D}$ is non-empty, let $x := \min \widehat{D}$ (or, alternatively, $x := \max \widehat{D}$) and set $D' = \widehat{D}\backslash\{x\}$, then Proposition 3.12 yields

$$\tilde{T}_{N+1}(\widehat{D}; z) = \tilde{T}_N(D'; z) + \sum_{\substack{Y \ni x: \\ Y \cap D' = \varnothing}} z(Y) \tilde{T}_N(D' \cup Y; z).$$

Since all the arguments D' and $D' \cup Y$ of $\tilde{T}_N$ on the right side of this identity are again finite discrete intervals, the inductive hypothesis and our assumption (4.20) imply that

$$\tilde{T}_{N+1}(\widehat{D}; z) \leq \mathrm{e}^{a(D')} + \sum_{\substack{Y \ni x: \\ Y \cap D' = \varnothing}} z(Y) \mathrm{e}^{a(D'\cup Y)} \leq \mathrm{e}^{a(\widehat{D})}.$$

This completes the inductive proof of the inequality $\tilde{T}_N(D;z) \le \exp(a(D))$. Passing to the limit $N \to \infty$, we get $\tilde{T}(D;z) \le \exp(a(D)) < \infty$ for all intervals $D \subset \mathbb{Z}$. The convergence extends to all finite sets because $\log \tilde{T}(\cdot;z)$ is subadditive. □

Proof of Theorem 4.5(a) Consider the selection rule $s(D) := \min D$ that picks the left-most point of a finite set. For $\alpha > 0$ and $L \in \mathbb{N}$ let

$$V_L(D) := \big|\{X \subset \mathbb{Z} \mid X \text{ is an } L\text{-rod}, \, X \cap D \neq \varnothing\}\big|$$

and $a(D) \equiv a_{\alpha,L}(D) := \alpha\, V_L(D)$. The choice of α and L is specified later. For a non-empty interval D, write $x := s(D) = \min(D)$, and $D' := D \setminus \{x\}$. If D' is non-empty, then condition (4.20) reads

$$\sum_{\ell=1}^{\infty} z_\ell \, e^{\alpha \ell} \le e^{\alpha} - 1. \tag{4.21}$$

Indeed in that case for each ℓ there is a single ℓ-rod X that contains x but does not intersect D' (note that $x+1 \in D'$ because of the assumption that D is an interval and $x = \min D$). The rod is simply the ℓ-rod with right-most endpoint x. Moreover $V_L(D' \cup X) - V_L(D') = \ell$ and $V_L(D' \cup \{x\}) - V_L(D') = 1$.

On the other hand if D' is empty, then the number of ℓ-rods that contain any given site $x \in \mathbb{Z}^d$ is equal to ℓ and the number of L-rods intersecting an ℓ-rod is equal to $L + \ell - 1$, therefore condition (4.20) reads instead

$$\sum_{\ell=1}^{\infty} \ell z_\ell \, e^{\alpha(\ell + L - 1)} \le e^{\alpha L} - 1. \tag{4.22}$$

The proof of Theorem 4.5 is complete once we check the existence of $\alpha > 0$ and $L \in \mathbb{N}$ such that the inequalities (4.21) and (4.22) hold true. Set

$$h(u) := 1 + \sum_{\ell=1}^{\infty} z_\ell \, u^\ell \qquad (u \in \mathbb{R}_+).$$

Conditions (4.21) and (4.22) are equivalent to

$$h(e^\alpha) \le e^\alpha, \quad h'(e^\alpha) \le 1 - e^{-\alpha L}. \tag{4.23}$$

Notice that h is convex and monotone increasing with $h(0) = 1$. The assumption (4.19) yields the existence of some $u = e^\alpha > 0$ such that $h(u) < u$. On the other hand, clearly $h(0) = 1 > 0$. Therefore the mean-value theorem yields the existence of a point $\tilde{u} \in (0,u)$ such that $h(\tilde{u}) = \tilde{u}$. The point $\tilde{u}$ is necessarily larger then 1 because $h(\tilde{u})$ is. Suppose by contradiction that $h'(\tilde{u}) \ge 1$. Then the convexity of h implies

$$h(u) \ge h(\tilde{u}) + h'(\tilde{u})(u - \tilde{u}) \ge h(\tilde{u}) + (u - \tilde{u}) = u,$$

which contradicts the assumption $h(u) < u$. Therefore $h(\tilde{u}) = \tilde{u} > 1$ and $h'(\tilde{u}) < 1$. Replacing α with $\tilde{\alpha} := \log \tilde{u}$ if needed, and picking $L = L(\alpha)$ large enough, we find that (4.23) is satisfied for some $\alpha > 0$. This concludes the proof. □

Proof of Theorem 4.5(b) Let $a(D) := \log \tilde{T}(D;z) = \log T(D;-z)$. In view of Eq. (2.15) and the alternating sign property, we have

$$a(D) = \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1,\dots,Y_k) \in \mathbb{X}^k} \mathbb{1}_{\{\exists i:\, Y_i \cap D \neq \varnothing\}} \big|\varphi_k^{\mathsf{T}}(Y_1,\dots,Y_k)\big| \, z(Y_1)\cdots z(Y_k).$$

By Proposition 3.12, for every $D \subset \mathbb{Z} \setminus \{1\}$, we have

$$\sum_{\substack{Y \ni 1,\\ Y \cap D = \varnothing}} z(X) \mathrm{e}^{a(D \cup Y) - a(D)} \leq \mathrm{e}^{a(D \cup \{1\}) - a(D)} - 1. \tag{4.24}$$

Let us choose $D \subset \mathbb{Z} \cap (-\infty, 0]$ with $0 \in D$. Then for every given $\ell \in \mathbb{N}$, the unique rod of length ℓ that contains 1 but does not intersect D is the rod $\{1, \dots, \ell\}$, and we obtain

$$\sum_{\ell=1}^{\infty} z_\ell\, \mathrm{e}^{a(D \cup \{1, \dots, \ell\}) - a(D)} \leq \mathrm{e}^{a(D \cup \{1\}) - a(D)} - 1. \tag{4.25}$$

Let $D_0 := D$ and for $m \geq 1$ set $D_m := D \cup \{1, \dots, m\}$. The exponent on the left-hand side in (4.25) may be written as

$$a(D \cup \{1, \dots, \ell\}) - a(D) = \sum_{m=1}^{\ell} \big(a(D_m) - a(D_{m-1}) \big). \tag{4.26}$$

Now

$$\begin{aligned} &a(D_m) - a(D_{m-1}) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1, \dots, Y_k) \in \mathbb{X}^k} \Big(1\!\!1_{\{\exists i:\, Y_i \cap D_m \neq \varnothing\}} - 1\!\!1_{\{\exists i:\, Y_i \cap D_{m-1} \neq \varnothing\}} \Big) \big|\varphi_k^\mathsf{T}(Y_1, \dots, Y_k)\big| z(Y_1) \cdots z(Y_k). \end{aligned}$$

The only clusters $(Y_1, \dots, Y_k)$ that contribute to the sum are those that intersect D_m but do not intersect D_{m-1}. This is only possible if one of the Y_i's contains m and all of them are contained in $\mathbb{Z} \cap [m, \infty)$. Thus

$$\begin{aligned} &a(D_m) - a(D_{m-1}) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(Y_1, \dots, Y_k) \in \mathbb{X}^k} 1\!\!1_{\{\exists i:\, Y_i \ni m\}} 1\!\!1_{\{\forall i:\, Y_i \subset [m, \infty)\}} \big|\varphi_k^\mathsf{T}(Y_1, \dots, Y_k)\big| z(Y_1) \cdots z(Y_k). \end{aligned}$$

Because of the translational invariance, the value of the sum does not depend on m. Thus $a(D_m) - a(D_{m-1}) = \alpha > 0$ for all $m \geq 1$ and some $\alpha > 0$. Turning back to (4.26), we obtain

$$a(D \cup \{1, \dots, \ell\}) - a(D) = \ell \alpha$$

and then (4.25) yields $\sum_{\ell=1}^{\infty} z_\ell \exp(\alpha \ell) \leq \exp(\alpha) - 1$. □

4.5 Tonks Gas on $\mathbb{R}$

Next, we want to consider the continuous version of the one-dimensional Tonks gas. Let $(L_\ell)_{\ell \in \mathbb{N}}$ be sequence of strictly positive numbers and $\mathbb{X}$ the space of compact intervals $I \subset \mathbb{R}$ with lengths $|I| \in \{L_\ell \mid \ell \in \mathbb{N}\}$. The map $\mathbb{R} \times \mathbb{N}$, $(x, \ell) \mapsto [x - L_\ell/2, x + L_\ell/2]$ is a bijection between $\mathbb{R} \times \mathbb{N}$ and $\mathbb{X}$. The reference measure λ is defined by the equality

$$\int_{\mathbb{X}} h(X) \lambda(\mathrm{d}X) = \sum_{\ell=1}^{\infty} \int_{-\infty}^{\infty} h\big(\big[x - \tfrac{L_\ell}{2}, x + \tfrac{L_\ell}{2}\big]\big) \mathrm{d}x$$

for all non-negative measurable functions $h : \mathbb{X} \to \mathbb{R}_+$. We assume that the activity is of the form

$$z(X) = \begin{cases} z_\ell, & X = [x, x + L_\ell] \text{ for some } \ell \in \mathbb{N}, x \in \mathbb{R}, \\ 0, & \text{else} \end{cases}$$

for some sequence $(z_\ell)_{\ell \in \mathbb{N}}$ of non-negative numbers. We assume that rod lengths are bounded from below, i.e., there exists $\delta > 0$ such that

$$\inf_{\ell \in \mathbb{N}} L_\ell \geq \delta. \tag{4.27}$$

From here on, we will consider the following chopping map: For $X = [x, x+L_\ell] \in \mathbb{X}$, let $C(X) = \{E_1, \dots, E_m\}$ consist of the intersections of X with the intervals $[x+(k-1)\varepsilon, x+k\varepsilon)$ with $k \in \mathbb{Z}$, where $\varepsilon \in (0, \delta)$. The space of snippets $\mathbb{E}_\varepsilon$ consists of intervals $[a, b]$ and $[a, b)$ of length $b - a \leq \varepsilon$.

Theorem 4.7 *In the setup of multi-type Tonks gas on $\mathbb{R}$, under the assumption (4.27):*

(a) *Suppose there exists $\alpha > 0$ such that*

$$\sum_{\ell=1}^{\infty} \mathrm{e}^{\alpha L_\ell} z_\ell < \alpha. \tag{4.28}$$

Then the expansion for $T(D; z)$ is absolutely convergent, for all bounded sets $D \subset \mathbb{R}$.

(b) *Conversely, if $T(D; z)$ is absolutely convergent for all bounded subsets $D \subset \mathbb{R}$, then there exists $\alpha > 0$ such that (4.28) holds true with "$\leq$" instead of "$<$".*

Remark 4.5 The theorem essentially recovers the necessary and sufficient convergence criterion from [13] (derived there for the activity expansion of the pressure in the system). The sufficient condition in [13] is (4.28) with "$\leq$" instead of "$<$". Again, while the result itself is not novel, its proof demonstrates the potential of our approach to go beyond the Fernández-Procacci criterion—also in continuous setups.

First we prove an auxiliary result, the analogue of Proposition 4.6 for the continuous setup, which is not quite as trivial. We introduce the following notion: Define the ε-gap-filling operation $\widehat{\cdot}$ by setting $\widehat{D} := D \cup \{x \in \mathbb{R} | \ \exists y, z \in D \textit{ with } y < x < z \textit{ such that } z - y \leq \varepsilon\}$ for any $D \subset \mathbb{R}$. Let $\mathscr{P}$ be some subset of the power set of $\mathbb{R}$, we say that a function $\xi : \mathscr{P} \to \mathbb{R}$ does not see gaps of diameter at most ε if it is invariant under the ε-gap-filling operation, i.e., if $\xi(D) = \xi(\widehat{D})$ for all $D \in \mathscr{P}$.

Proposition 4.8 *Suppose that there exists a non-negative, measurable map $a(\cdot)$ defined on finite unions of (bounded) intervals which does not see gaps of diameter at most ε and satisfies the following system of inequalities: For any (bounded) interval D with $C(D) = \{E_1, \dots, E_n\}$, $E_1, \dots, E_n \in \mathbb{E}_\varepsilon$, where the chopping map C is defined as above, there is a subinterval $E_s \subset D$ of length at most ε, such that*

$$\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} I(E_s; D'; Y_1, \dots, Y_k) \mathrm{e}^{a(D' \cup Y_1 \cup \dots \cup Y_k) - a(D')} \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \leq \mathrm{e}^{a(E_s \cup D') - a(D')} - 1, \tag{4.29}$$

where we set $D' := D \backslash E_s$ and $I(E_s; D'; Y_1, \dots, Y_k)$ is the indicator from Eq. (2.16). Then $T(D; z)$ is absolutely convergent for all bounded subsets $D \subset \mathbb{R}$.

Proof We can modify the Kirkwood–Salsburg-type equations $\tilde{\kappa}_z^s$ from Chapter 3.3 as follows: If $\xi(\cdot)$ is a function from $\mathbb{D}_\varepsilon$ to $\mathbb{R}_+$ that does not see gaps of diameter at most ε and satisfies the measurability assumption from Theorem 2.6, define the function $\tilde{\mathscr{K}}_z^s\xi$ (possibly assuming the value "∞") by

$$(\tilde{\mathscr{K}}_z^s\xi)(D) := \mathbb{1}_{\{n\geq 2\}}\,\xi\Big(\widehat{E_2'\cup\ldots\cup E_n'}\Big)$$
$$+\sum_{k=1}^{\infty}\frac{1}{k!}\int_{\mathbb{X}^k} I(E_s;E_2'\cup\cdots\cup E_n';Y_1,\ldots,Y_k)\xi\Big(\widehat{E_2'\cup\ldots\cup E_n'\cup Y_1\cup\ldots\cup Y_k}\Big)\lambda_z^k(\mathrm{d}\boldsymbol{Y}),$$

for $D\in\mathbb{D}_\varepsilon$ with $E_1,\ldots,E_n\subset\mathbb{E}_\varepsilon$ and $C(D)=\{E_1,\ldots,E_n\}$, where $\widehat{D}$ is given by "filling gaps" of diameter at most ε in $D\subset\mathbb{R}$ as defined above.

Notice that for any such function $\xi(\cdot)$ (that does not see gaps of diameter at most ε and satisfies the measurability assumption from Theorem 2.6)

$$\tilde{\mathscr{K}}_z^s\xi=\tilde{\kappa}_z^s\xi$$

holds, where $\tilde{\kappa}_z^s\xi$ is the function defined by (3.9). In particular, the left hand side of the equation is well-defined. Since the functions $\tilde{T}_N(\cdot;z)$, $N\in\mathbb{N}$, and $\tilde{T}(\cdot;z)$ do not see gaps of diameter at most ε (by our assumption $\varepsilon<\delta$ and the respective definitions), Proposition 3.10 implies

$$\tilde{T}\big(\cdot;z\big)=e(\cdot)+\tilde{\mathscr{K}}_z^s\tilde{T}(\cdot;z).$$

and

$$\tilde{T}_{N+1}\big(\cdot;z\big)=e(\cdot)+(\tilde{\mathscr{K}}_z^s\tilde{T}_N)\big(\cdot;z\big).$$

Assumption (4.29) is equivalent to $e(D)+(\tilde{\mathscr{K}}_z^s\mathrm{e}^a)(D)\leq\mathrm{e}^{a(D)}$ for any interval $D\subset\mathbb{R}$. We prove by induction over N that $\tilde{T}_N(D;z)\leq\mathrm{e}^{a(D)}$ for all $N\in\mathbb{N}$ and all intervals $D\subset\mathbb{R}$. For $N=1$, we have by our assumption

$$\tilde{T}_1(D,z)=e(D)\leq e(D)+(\tilde{\mathscr{K}}_z^s\mathrm{e}^a)(D)\leq\mathrm{e}^{a(D)}$$

for all intervals $D\subset\mathbb{R}$. Next, assume for some $N\in\mathbb{N}$ that $\tilde{T}_N(D;z)\leq\mathrm{e}^{a(D)}$ for all intervals $D\subset\mathbb{R}$, then

$$\tilde{T}_{N+1}(D;z)=e(D)+\tilde{\mathscr{K}}_z^s\tilde{T}_N(D;z)\leq e(D)+(\tilde{\mathscr{K}}_z^s\mathrm{e}^a)(D)\leq\mathrm{e}^{a(D)},$$

where the first inequality holds by the inductive hypothesis, by monotonicity of $\tilde{\mathscr{K}}_z^s$ on non-negative functions and by the observation that for intervals $D\subset\mathbb{R}$ all the arguments of ξ appearing in the definition of $(\tilde{\mathscr{K}}_z^s\xi)(D)$ are again intervals.

This completes the induction and proves $T_N(D;z)\leq\mathrm{e}^{a(D)}$ for all $N\in\mathbb{N}$ and all intervals $D\subset\mathbb{R}$. Taking the limit $N\to\infty$ yields the corresponding bound for $\tilde{T}(D;z)$. The claim for arbitrary bounded subsets follows since every bounded subset is contained in some compact interval and

$$\tilde{T}(D_1;z)\leq\tilde{T}(D_2,z)$$

for $D_1\subset D_2\subset\mathbb{R}$. □

Proof of Theorem 4.7(a) In analogy to the discrete case, for $\alpha>0$ and $L>0$ let

$$V_L(D):=\int_{-\infty}^{\infty}\mathbb{1}_{\{[x,x+L]\cap D\neq\varnothing\}}\,\mathrm{d}x.$$

and $a(D) \equiv a_{\alpha,L}(D) := \alpha\, V_L(D)$. The choice of α and L is specified later in the proof. We apply Proposition 4.8 with the choice of the chopping map introduced at the beginning of this subsection and the selection rule s that picks the leftmost snippet.

Remember the indicator $I(E_1; D'; Y_1, \dots, Y_k)$ from Eq. (2.16). We show that there exists $\alpha > 0$ such that

$$\sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} I(E_1; D'; Y_1, \dots, Y_k) \mathrm{e}^{\alpha[V_L(D'\cup Y_1\cup\dots\cup Y_k) - V_L(D')]} \lambda_z^k(\mathrm{d}\boldsymbol{Y}) \\ \leq \mathrm{e}^{\alpha V_L(D'\cup E_1) - V_L(D')} - 1 \tag{4.30}$$

for all intervals $D' = [a, b) \subset \mathbb{R}$ or $D' = [a, b]$, including the empty set $D' = \varnothing$, and all snippets $E_1 = [(k-1)\varepsilon, a) \in \mathbb{E}_\varepsilon$.

If D' is non-empty, then because of $\inf_{\ell\in\mathbb{N}} L_\ell \geq \varepsilon$ and $|E_1| \leq \varepsilon$ there cannot be two or more disjoint rods in X that intersect E_1 but do not intersect D', so the inequality to be proven reduces to

$$\int_{\mathbb{X}} \mathrm{e}^{\alpha(V_L(D'\cup Y) - V_L(D'))} \mathbb{1}_{\{Y\cap E_1 \neq \varnothing,\ Y\cap D' = \varnothing\}} \lambda_z(\mathrm{d}Y) \leq \mathrm{e}^{\alpha(V_L(D'\cup E_1) - V_L(D'))} - 1. \tag{4.31}$$

Assuming that $L \geq \varepsilon$, this is equivalent to

$$\sum_{\ell=1}^{\infty} z_\ell \int_0^{|E_1|} \mathrm{e}^{\alpha(L_\ell + x)} \mathrm{d}x \leq \mathrm{e}^{\alpha|E_1|} - 1. \tag{4.32}$$

The integral on the left-hand side is equal to $\exp(\alpha L_\ell)[\exp(\alpha|E_1|) - 1]/\alpha$, so we find that (4.30) is equivalent to

$$\sum_{\ell=1}^{\infty} z_\ell\, \mathrm{e}^{\alpha L_\ell} \leq \alpha,$$

which holds true because of the assumption (4.28).

If D' is empty, we note that there can be at most two disjoint rods in X that intersect the snippet E_1, hence (4.30) becomes

$$\int_{\mathbb{X}} \mathrm{e}^{\alpha V_L(Y)} \mathbb{1}_{\{Y\cap E_1\neq\varnothing\}} \lambda_z(\mathrm{d}Y) + \frac{1}{2}\int_{\mathbb{X}^2} \mathrm{e}^{\alpha V_L(Y_1\cup Y_2)} \mathbb{1}_{\{Y_1\cap E_1\neq\varnothing,\ Y_2\cap E_1\neq\varnothing,\ Y_1\cap Y_2=\varnothing\}} \lambda_z^2\big(\mathrm{d}(Y_1, Y_2)\big) \\ \leq \mathrm{e}^{\alpha V_L(E_1)} - 1. \tag{4.33}$$

The right-hand side is equal to $\exp(\alpha(L + |E_1|)) - 1$. The first term on the left-hand side is equal to

$$\sum_{\ell=1}^{\infty} z_\ell (L_\ell + |E_1|) \mathrm{e}^{\alpha(L+L_\ell)} = \sum_{\ell=1}^{\infty} z_\ell\, L_\ell\, \mathrm{e}^{\alpha(L+L_\ell)} + O(\varepsilon).$$

The second term on the left-hand side of (4.33) is equal to

$$\sum_{\ell,r=1}^{\infty} z_\ell z_r \int_{E_1^2} \mathbb{1}_{\{x<y\}}\, \mathrm{e}^{\alpha V_L([x-L_\ell, y+L_r])} \mathrm{d}x\mathrm{d}y$$

which is bounded by

$$\Big(\sum_{\ell=1}^{\infty} z_\ell \mathrm{e}^{\alpha L_\ell}\Big)^2 \mathrm{e}^{\alpha(\varepsilon+L)} |E_1|^2 = O(\varepsilon^2).$$

For the inequality (4.33) to be satisfied, it is sufficient that

$$\sum_{\ell=1}^{\infty} L_\ell z_\ell \mathrm{e}^{\alpha L_\ell} + O(\varepsilon) \le \mathrm{e}^{\alpha|E_1|} - \mathrm{e}^{-\alpha L}. \tag{4.34}$$

Arguments similar to the proof of Theorem 4.5(b), applied to the convex function $h : \mathbb{R}_+ \to \mathbb{R}$, $h(u) := 1 + \sum_{\ell=1}^{\infty} z_\ell u^{L_\ell}$, show that under condition (4.28) there exists $\alpha > 0$ such that not only condition (4.28) holds true but in addition

$$h'(\mathrm{e}^\alpha) = \sum_{\ell=1}^{\infty} L_\ell z_\ell \mathrm{e}^{\alpha L_\ell} < 1.$$

Thus one can choose $L = L(\alpha)$ large enough and ε small enough so that (4.34) and hence (4.33) hold true. □

Proof of Theorem 4.7(b) We proceed as in the proof of Theorem 4.5(b). Suppose that the expansions are absolutely convergent and define

$$a(D) := \log T(D; -z) = \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{\exists i \in [n]:\, Y_i \cap D \neq \varnothing\}} \left|\varphi_k^{\mathsf{T}}(Y_1, \dots, Y_k)\right| \lambda_z^k(\mathrm{d}\boldsymbol{Y}).$$

Then by Proposition 3.13 and since $\inf_{\ell\in\mathbb{N}} L_\ell$ is bounded from below by $\varepsilon > 0$,

$$\int_{\mathbb{X}} \mathbb{1}_{\{X\cap E_1 \neq \varnothing,\, X \cap D' = \varnothing\}} \mathrm{e}^{a(D'\cup X) - a(D')} \lambda_z(\mathrm{d}X) \le \mathrm{e}^{a(D'\cup E_1) - a(D')} - 1 \tag{4.35}$$

for example for $E_1 = [0, \varepsilon)$ and $D' = [\varepsilon, \varepsilon + L]$ with $L > 0$ and ε sufficiently small.

Before we evaluate the two sides of the inequality, we note two useful properties of $a(\cdot)$. First, the map a does not see gaps of diameter at most ε. Precisely, if $X = [x - L_\ell, x]$ with $x \in [0, \varepsilon)$ and D' is as above, then

$$a(D' \cup X) = a([x - L_\ell, \varepsilon + L]).$$

Indeed, any rod $Y_i \in \mathbb{X}$ that intersects $[0, \varepsilon)$ must also intersect $D' \cup X$ because it has a length $|Y_i| \ge \varepsilon$. Second, because of translational invariance, the weight $a(D)$ of a non-empty interval depends only on its length $|D|$. We check that in addition, it is an affine function of the length. For $x \in \mathbb{R}$, define

$$\alpha(x) := \sum_{\ell=1}^{\infty} z_\ell \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{\forall i \in [k]:\, Y_i \subset (-\infty, x]\}} \left|\varphi_{1+k}^{\mathsf{T}}(Y_1, \dots, Y_k, [x - L_\ell, x])\right| \lambda_z^k(\mathrm{d}\boldsymbol{Y}).$$

The quantity $\alpha(x)$ is best thought of as an integral over clusters in which the right-most rod $[x - L_\ell, x]$ has its right end pinned at x. By translational invariance, $\alpha(x)$ is actually independent of x and we may write $\alpha(x) \equiv \alpha$ for some scalar $\alpha \ge 0$. Now let $I = [a, b]$ and $J = [b, c]$ with $a < b < c$. Then

$$a(I \cup J) - a(J) = \sum_{k=1}^{\infty} \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}_{\{\exists i \in [k]:\, Y_i \cap I \neq \varnothing\}} \mathbb{1}_{\{\forall i \in [k]:\, Y_i \cap J = \varnothing\}} \left|\varphi_k^{\mathsf{T}}(Y_1, \dots, Y_k)\right| \lambda_z^k(\mathrm{d}\boldsymbol{Y}).$$

Any cluster $(Y_1, \dots, Y_k)$ that intersects I but not J has its right-most end in $[a, b)$, therefore

$$a(I \cup J) - a(J) = \int_I \alpha(x)\mathrm{d}x = \alpha\, |I|.$$

With these two observations, the left-hand side of (4.35) becomes

$$\sum_{\ell=1}^{\infty} z_\ell \int_0^\varepsilon \mathrm{e}^{a([x-L_\ell,x]\cup D')-a(D')}\mathrm{d}x = \sum_{\ell=1}^{\infty} z_\ell \int_0^\varepsilon \mathrm{e}^{\alpha(x+L_\ell)}\mathrm{d}x = \sum_{\ell=1}^{\infty} z_\ell \mathrm{e}^{\alpha L_\ell}\frac{1}{\alpha}(\mathrm{e}^{\alpha\varepsilon}-1)$$

while the right-hand side of (4.35) is $\exp(\alpha\varepsilon) - 1$. It follows that

$$\sum_{\ell=1}^{\infty} z_\ell \mathrm{e}^{\alpha L_\ell} \le \alpha.$$

□

Funding Open Access funding enabled and organized by Projekt DEAL. The authors did not receive support from any funding body for the submitted work.

Data Availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

A Proof of Lemma 2.5

Proof of Lemma 2.5 We show that the system of inequalities

$$\frac{1+\sum\limits_{k\ge1}\sum\limits_{\substack{Y=\{y_1,\ldots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}}\prod\limits_{i=1}^{k}\mu(y_i)\prod\limits_{w\in\Gamma(Y)}\mathrm{e}^{\mu(w)}}{\prod\limits_{q\in\Gamma(x_1)\cap\Gamma(X)}\mathrm{e}^{\mu(q)}}$$

$$\ge 1+\sum_{k\ge1}\sum_{\substack{Y=\{y_1,\ldots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j\\ y_i\sim X}}\prod_{i=1}^{k}\mu(y_i)\prod_{w\in\Gamma(Y)\cap\Gamma(X)^C}\mathrm{e}^{\mu(w)}, \tag{A.1}$$

which is equivalent to (2.12), holds under the assumptions of the lemma. We do so by proving the following three claims. However, first we would like to introduce some additional notation to complement the notation from Sect. 2.2.

For given $x_1 \in \mathbb{X}$ and $X = \{x_2, \ldots, x_p\} \subset \mathbb{X}$ let Q denote the set $\Gamma(x_1)\cap\Gamma(X)$ and let $\mathcal{C}$ denote the set of (non-empty) compatible subsets of Q. Furthermore, we define the family

$(A_U)_{U\subset Q}$, $A_U = A_U(x_1, Q, \mu)$, indexed by all the subsets $U \subset Q$ (including the empty set), by

$$A_U = A_U(x_1, Q, \mu) := \sum_{k\geq 1} \sum_{Y=\{y_1,\dots,y_k\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)\setminus U} \mathrm{e}^{\mu(w)},$$

where the sum is over subsets $Y = \{y_1, \dots, y_k\} \subset \mathbb{X}$ such that the following constraints are satisfied: Y is an compatible set, $Y \subset \Gamma(x_1)$, $Y \cap Q = \varnothing$ and $U = \Gamma(Y) \cap Q$.
Finally, define the family of coefficients $(\beta_U)_{U\subset Q}$, $\beta_U = \beta_U(x_1, Q, \mu)$, also indexed by all the subsets $U \subset Q$ (including the empty set), by

$$\beta_U = \beta_U(x_1, Q, \mu) := \prod_{q\in Q\setminus U} \mathrm{e}^{-\mu(q)} + \sum_{\substack{C\in\mathcal{C}\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}.$$

Then the following statements hold true:

Claim A.1 *The right-hand side of* (A.1) *is bounded from above by*

$$1 + \sum_{U\subset Q} A_U.$$

Claim A.2 *The left-hand side of* (A.1) *is bounded from below by*

$$\beta_\varnothing + \sum_{U\subset Q} \beta_U A_U. \tag{A.2}$$

Claim A.3 *The lower bounds* $\beta_U \geq 1$ *hold for every* $U \subset Q$ *and thus*

$$\beta_\varnothing + \sum_{U\subset Q} \beta_U A_U \geq 1 + \sum_{U\subset Q} A_U.$$

All the sums in the three claims above run over subsets of $Q = \Gamma(x_1) \cap \Gamma(X)$ including the empty set (so that the coefficient $\beta_\varnothing$ appears in (A.2) twice). The inequalities (A.1) follow directly from the three claims. Proving the claims is thus sufficient to conclude the proof of the lemma:

Proof of Claim A.1 Reorder the sum in the right-hand side of (A.1) by putting the summand together which belong to the same $U := \Gamma(Y) \cap Q$. Notice that the constraint $Y \sim X$ implies $Y \cap Q = \varnothing$ since $Y \subset \Gamma(x_1)$. The claim now follows directly from the simple observation that $\Gamma(Y) \cap \Gamma(X)^C \subset \Gamma(Y)\setminus U$ for any Y and thus

$$\prod_{w\in\Gamma(Y)\setminus U} \mathrm{e}^{\mu(w)} \geq \prod_{w\in\Gamma(Y)\cap\Gamma(X)^C} \mathrm{e}^{\mu(w)}.$$

□

Proof of Claim A.2 To see that the bounds stated in the claim hold, decompose the sum in the left-hand side of (A.1) as

$$\begin{aligned} &1 + \sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)} \mathrm{e}^{\mu(w)} \\ &= 1 + \sum_{C\in\mathcal{C}\cup\{\varnothing\}} \sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}} \mathbb{1}_{\{C=Q\cap Y\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)} \mathrm{e}^{\mu(w)}. \end{aligned} \tag{A.3}$$

Notice that for any $C \in \mathcal{C}$

$$\sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{C=Q\cap Y\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)} \mathrm{e}^{\mu(w)}$$
$$\geq \prod_{c\in C} \mu(c) \prod_{w\in\Gamma(C)\cap Q} \mathrm{e}^{\mu(w)}$$
$$\left(1+\sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{Y\cap Q=\varnothing\}} \mathbb{1}_{\{Y\sim C\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)\setminus(\Gamma(C)\cap Q)} \mathrm{e}^{\mu(w)}\right).$$

This estimate is established by discarding the exponential weights corresponding to a subset of $\Gamma(C)$: Multiplying the lower bound in the last line with

$$\prod_{w\in\Gamma(C)\setminus\Gamma(Y)\setminus Q} \mathrm{e}^{\mu(w)} \geq 1$$

yields equality. No further estimates are necessary to prove the claim; we simply plug the obtained lower bound into the right-hand side of (A.3) and get

$$1+\sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{Y\cap Q=\varnothing\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)} \mathrm{e}^{\mu(w)} + \sum_{C\in\mathcal{C}} \prod_{c\in C} \mu(c) \prod_{w\in\Gamma(C)\cap Q} \mathrm{e}^{\mu(w)}$$
$$\times\left(1+\sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{Y\cap Q=\varnothing\}} \mathbb{1}_{\{Y\sim C\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)\setminus(\Gamma(C)\cap Q)} \mathrm{e}^{\mu(w)}\right)$$

or, equivalently,

$$1+\sum_{C\in\mathcal{C}} \prod_{c\in C} \mu(c) \prod_{w\in\Gamma(C)\cap Q} \mathrm{e}^{\mu(w)} + \sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{Y\cap Q=\varnothing\}} \prod_{i=1}^{k} \mu(y_i) \prod_{w\in\Gamma(Y)} \mathrm{e}^{\mu(w)}$$
$$+\sum_{C\in\mathcal{C}} \prod_{c\in C} \mu(c) \prod_{w\in\Gamma(C)\cap Q} \mathrm{e}^{\mu(w)} \sum_{k\geq 1} \sum_{\substack{Y=\{y_1,\dots,y_k\} \\ y_i \nsim x_1,\ y_i \sim y_j}} \mathbb{1}_{\{Y\cap Q=\varnothing\}} \mathbb{1}_{\{Y\sim C\}} \prod_{i=1}^{k} \mu(y_i)$$
$$\prod_{w\in\Gamma(Y)\setminus(\Gamma(C)\cap Q)} \mathrm{e}^{\mu(w)}$$

as a lower bound for the left-hand side of (A.3).

Reordering the last expression by summing over Y first, one realizes that it is equal to

$$1+\sum_{C\in\mathcal{C}}\prod_{c\in C}\mu(c)\prod_{w\in\Gamma(C)\cap Q}\mathrm{e}^{\mu(w)}+\sum_{k\geq 1}\sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}}\mathbb{1}_{\{Y\cap Q=\varnothing\}}\prod_{i=1}^{k}\mu(y_i)\prod_{w\in\Gamma(Y)}\mathrm{e}^{\mu(w)}$$
$$\times\left(1+\sum_{\substack{C\in\mathcal{C}\\ C\sim Y}}\prod_{c\in C}\mu(c)\prod_{w\in(\Gamma(C)\cap Q)\setminus\Gamma(Y)}\mathrm{e}^{\mu(w)}\right),$$

which may be rewritten as

$$1+\sum_{C\in\mathcal{C}}\prod_{c\in C}\mu(c)\prod_{w\in\Gamma(C)\cap Q}\mathrm{e}^{\mu(w)}+\sum_{k\geq 1}\sum_{\substack{Y=\{y_1,\dots,y_k\}\\ y_i\nsim x_1,\ y_i\sim y_j}}\mathbb{1}_{\{Y\cap Q=\varnothing\}}\prod_{i=1}^{k}\mu(y_i)\prod_{w\in\Gamma(Y)\setminus(\Gamma(Y)\cap Q)}\mathrm{e}^{\mu(w)}$$
$$\times\left(\prod_{w\in\Gamma(Y)\cap Q}\mathrm{e}^{\mu(w)}+\sum_{\substack{C\in\mathcal{C}\\ C\sim Y}}\prod_{c\in C}\mu(c)\prod_{w\in(\Gamma(C)\cup\Gamma(Y))\cap Q}\mathrm{e}^{\mu(w)}\right).$$

From the last expression, we obtain precisely the sum in (A.2) by putting the summands in the last expression which belong to the same $U=\Gamma(Y)\cap Q$ together and dividing by $\prod_{q\in Q}\mathrm{e}^{\mu(q)}$. This yields the claimed lower bound. □

Proof of Claim A.3 Without loss of generality, we may assume that $Q=\Gamma(x_1)\cap\Gamma(X)$ is a finite non-empty set. Then the claim can be proven via induction over the cardinality of Q.

To start the induction consider the case $Q=\{q\}$, $q\in\mathbb{X}$. Then $\beta_\varnothing=\mathrm{e}^{-\mu(q)}+\mu(q)\geq 1$ and $\beta_{\{q\}}=1$ by definition.

For the inductive step, let $n\in\mathbb{N}$, $Q_n=\{q_1,\dots,q_n\}\subset\mathbb{X}$ and let $\mathcal{C}_n$ be the set of compatible subsets of Q_n. Furthermore, let $q_{n+1}\in\mathbb{X}\setminus Q_n$ and let $Q_{n+1}=Q_n\cup\{q_{n+1}\}$. Naturally, there exists a family of subsets $\Theta_n\subset\mathcal{C}_n$, such that the set $\mathcal{C}_{n+1}$ of compatible subsets of Q_{n+1} is given by $\mathcal{C}_{n+1}=\{C\cup\{q_{n+1}\}|C\in\Theta_n\ or\ C=\varnothing\}\cup\mathcal{C}_n=:\overline{\Theta}_n\cup\mathcal{C}_n$.

Under the assumption that $\beta_U(Q_n)\geq 1$ for all $U\subset Q_n$ it is to show that $\beta_U(Q_{n+1})\geq 1$ for all $U\subset Q_{n+1}$. Therefore let $U\subset Q_{n+1}$. If $q_{n+1}\in U$ then $\beta_U(Q_{n+1})=\beta_{U\setminus\{q_{n+1}\}}(Q_n)\geq 1$ by the inductive hypothesis. Left to consider is the case $q_{n+1}\notin U$ (and thus $U\subset Q_n$). Recall that we defined the coefficient $\beta_U(Q_{n+1})$ by

$$\beta_U(Q_{n+1})=\prod_{q\in Q_{n+1}\setminus U}\mathrm{e}^{-\mu(q)}+\sum_{\substack{C\in\mathcal{C}_{n+1}\\ C\cap U=\varnothing}}\prod_{c\in C}\mu(c)\prod_{w\in(Q_{n+1}\setminus U)\setminus\Gamma(C)}\mathrm{e}^{-\mu(w)}.$$

Using the decomposition $\mathcal{C}_{n+1} = \overline{\Theta}_n \cup \mathcal{C}_n$, we get

$$\begin{aligned}
\beta_U(Q_{n+1}) &= \mathrm{e}^{-\mu(q_{n+1})} \prod_{q\in Q_n\setminus U} \mathrm{e}^{-\mu(q)} + \sum_{\substack{C\in\overline{\Theta}_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_{n+1}\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}\\
&\quad + \sum_{\substack{C\in\mathcal{C}_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_{n+1}\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}\\
&= \mathrm{e}^{-\mu(q_{n+1})} \prod_{q\in Q_n\setminus U} \mathrm{e}^{-\mu(q)} + \mu(q_{n+1}) \prod_{q\in(Q_n\setminus U)\setminus\Gamma(q_{n+1})} \mathrm{e}^{-\mu(q)}\\
&\quad + \mu(q_{n+1}) \sum_{\substack{C\in\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_{n+1}\setminus U)\setminus\Gamma(C\cup\{q_{n+1}\})} \mathrm{e}^{-\mu(w)}\\
&\quad + \mathrm{e}^{-\mu(q_{n+1})} \sum_{\substack{C\in\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}\\
&\quad + \sum_{\substack{C\in\mathcal{C}_n\setminus\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}.
\end{aligned}$$

Multiplying the second summand in the last expression by the products of negative exponential weights

$$\prod_{w\in\Gamma(\{q_{n+1}\})} \mathrm{e}^{-\mu(w)}) \le 1$$

and the third summand by

$$\prod_{w\in\Gamma(\{q_{n+1}\})\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}) \le 1,$$

we obtain the following lower bound:

$$\begin{aligned}
&\beta_U(Q_{n+1})\\
&\ge \left(\mathrm{e}^{-\mu(q_{n+1})} + \mu(q_{n+1})\right)\left(\prod_{q\in Q_n\setminus U} \mathrm{e}^{-\mu(q)} + \sum_{\substack{C\in\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}\right)\\
&\quad + \sum_{\substack{C\in\mathcal{C}_n\setminus\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}
\end{aligned}$$

Since $\mathrm{e}^{-\mu(q_{n+1})} + \mu(q_{n+1}) \ge 1$ for any μ, this last expression is in turn bounded from below by

$$\begin{aligned}
&\prod_{q\in Q_n\setminus U} \mathrm{e}^{-\mu(q)} + \sum_{\substack{C\in\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)} + \sum_{\substack{C\in\mathcal{C}_n\setminus\Theta_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)}\\
&\quad = \prod_{q\in Q_n\setminus U} \mathrm{e}^{-\mu(q)} + \sum_{\substack{C\in\mathcal{C}_n\\ C\cap U=\varnothing}} \prod_{c\in C} \mu(c) \prod_{w\in(Q_n\setminus U)\setminus\Gamma(C)} \mathrm{e}^{-\mu(w)} = \beta_U(Q_n).
\end{aligned}$$

By the inductive hypothesis $\beta_U(Q_n)$ is bounded from below by 1, hence we have shown $\beta_U(Q_{n+1}) \ge 1$. This concludes the induction and therefore also the proof of Claim A.3. □

Combining the three statements from the claims A.1, A.2 and A.3 immediately yields the claim of Lemma 2.5. □

References

1. Bissacot, R., Fernández, R., Procacci, A.: On the convergence of cluster expansions for polymer gases. J. Stat. Phys. **139**(4), 598–617 (2010)
2. Bissacot, R., Fernández, R., Procacci, A., Scoppola, B.: An improvement of the Lovász local lemma via cluster expansion. Comb. Probab. Comput. **20**(5), 709–719 (2011)
3. Brydges, D.C.: A short course on cluster expansions. In: Phénomènes critiques, systèmes aléatoires, théories de jauge, Part I, II (Les Houches, 1984), pp. 129–183. North-Holland, Amsterdam (1986)
4. Dobrushin, R.L.: Estimates of semi-invariants for the Ising model at low temperatures topics in statistics and theoretical physics. Am. Math. Soc. Transl. **177**(2), 59–81 (1996)
5. Faris, W.G.: A connected graph identity and convergence of cluster expansions. J. Math. Phys. **49**(11), 113302 (2008)
6. Faris, W.G.: Combinatorial species and cluster expansions. Mosc. Math. J. **10**(4), 713–727 (2010)
7. Federer, H.: Geometric Measure Theory. Springer, New York (1969)
8. Fernández, R., Procacci, A.: Cluster expansion for abstract polymer models. New bounds from an old approach. Commun. Math. Phys. **274**(1), 123–140 (2007)
9. Fernández, R., Procacci, A., Scoppola, B.: The analyticity region of the hard sphere gas. Improved bounds. J. Stat. Phys. **128**(5), 1139–1143 (2007)
10. Fialho, P.M.S.: Abstract polymer gas: a simple inductive proof of the Fernández–Procacci criterion. J. Stat. Phys. **178**, 1354–1361 (2020)
11. Friedli, S., Velenik, Y.: Statistical Mechanics of Lattice Systems. Cambridge University Press, Cambridge (2018).. (**A concrete mathematical introduction**)
12. Gruber, C., Kunz, H.: General properties of polymer systems. Commun. Math. Phys. **22**(2), 133–161 (1971)
13. Jansen, S.: Cluster and virial expansions for the multi-species Tonks gas. J. Stat. Phys. **161**(5), 1299–1323 (2015)
14. Jansen, S.: Cluster expansions for Gibbs point processes. Adv. Appl. Probab. **51**, 1129–1178 (2019)
15. Jansen, S.: Thermodynamics of a hierarchical mixture of cubes. J. Stat. Phys. **179**, 309–340 (2020)
16. Jansen, S., Kolesnikov, L.: Activity expansions for Gibbs correlation functions. In: Proceedings of the XI International Conference Stochastic and Analytic Methods in Mathematical Physics, Number 6 in Lectures in Pure and Applied Mathematics. Universitätsverlag Potsdam (2020)
17. Jansen, S., Kuna, T., Tsagkarogiannis, D.: Virial inversion and density functionals. Online preprint arXiv:1906.02322 [math-ph] (2019)
18. Kotecký, R., Preiss, D.: Cluster expansion for abstract polymer models. Commun. Math. Phys. **103**(3), 491–498 (1986)
19. Matheron, G.: Random sets and integral geometry. With a Foreword by Geoffrey S. Watson, Wiley Series in Probability and Mathematical Statistics. Wiley, New York, London, Sydney (1975)
20. Minlos, R.A., Pogosjan, S.K.: Estimates of Ursell functions, group functions and their derivatives. Teoret. Mat. Fiz. **31**(2), 199–213 (1977). (**English translation: Theor. Math. Phys. 31(2), 408–418 (1977)**)
21. Molchanov, I.: Theory of Random Sets. Springer, London (2005)
22. Nguyen, T.X., Fernández, R.: Convergence of cluster and virial expansions for repulsive classical gases. J. Stat. Phys. **179**(2), 448–484 (2020)
23. Penrose, O.: Convergence of fugacity expansions for classical systems. Stat. Mech. Found. Appl. **139**, 598–617 (1967)
24. Poghosyan, S., Ueltschi, D.: Abstract cluster expansion with applications to statistical mechanical systems. J. Math. Phys. **50**(5), 053509 (2009)
25. Procacci, A., Yuhjtman, S.A.: Convergence of Mayer and virial expansions and the Penrose tree-graph identity. Lett. Math. Phys. **107**(1), 31–46 (2017)
26. Ruelle, D.: Statistical Mechanics: Rigorous Results. World Scientific, Singapore (1969)
27. Schneider, R.: Convex Bodies: The Brunn-Minkowski Theory. Encyclopedia of Mathematics and Its Applications, 2nd edn. Cambridge University Press, Cambridge (2013)
28. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Springer, Berlin (2008)
29. Scott, A.D., Sokal, A.D.: The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. J. Stat. Phys. **118**(5–6), 1151–1261 (2005)

30. Stell, G.: Cluster expansions for classical systems in equilibrium. In: Frisch, H.L., Lebowitz, J.L. (eds.) The Equilibrium Theory of Classical Fluids, pp. 171–261. Benjamin, New York (1964)
31. Temmel, C.: Sufficient conditions for uniform bounds in abstract polymer systems and explorative partition schemes. J. Stat. Phys. **157**(6), 1225–1254 (2014)
32. Ueltschi, D.: Cluster expansions and correlation functions. Mosc. Math. J. **4**(2), 511–522 (2004)
33. Ueltschi, D.: An improved tree-graph bound. Oberwolfach Rep., 14, 2017. In: Miniworkshop: Cluster Expansions: From Combinatorics to Analysis Through Probability (R. Fernández, S. Jansen, D. Tsagkarogiannis, eds.)

3 Logarithms of Catalan generating functions: A combinatorial approach

This section contains a preprint version of:
S. Jansen and L. Kolesnikov. Logarithms of Catalan generating functions: A combinatorial approach. arXiv e-prints, 2022.

Personal contribution:
This article is a collaborative effort based on discussions with my supervisor, Sabine Jansen, who proposed this combinatorial puzzle. I developed and implemented many ideas incorporated in our main results, with occasional input from Sabine. I also took a leading role in drafting and writing the article.

Logarithms of Catalan generating functions: A combinatorial approach

Sabine Jansen*† Leonid Kolesnikov*

Abstract

We analyze the combinatorics behind the operation of taking the logarithm of the generating function G_k for k^{th} generalized Catalan numbers. We provide combinatorial interpretations in terms of lattice paths and in terms of tree graphs. Using explicit bijections, we are able to recover known closed expressions for the coefficients of $\log G_k$ by purely combinatorial means of enumeration. The non-algebraic proof easily generalizes to higher powers $\log^a G_k$, $a \geq 2$.

Keywords: Catalan numbers, logarithms of generating functions, combinatorial interpretation, lattice paths, Dyck paths, plane trees, cycle-rooted trees, exact enumeration.

MSC 2020 Classification: 05A15, 05A10.

1 Introduction

The present article originated in the following question: given $k \in \mathbb{N}$, what is the combinatorial interpretation of the power series $F(x)$ that solves the equation

$$\mathrm{e}^{F(x)} = 1 + x\,\mathrm{e}^{kF(x)}, \tag{1.1}$$

and is there a way of computing the coefficients of $F(x)$ by counting suitable labeled combinatorial structures? The question was raised in the context of the statistical mechanics for a one-dimensional system of non-overlapping rods on a line [6, Section 5.2]; up to sign flips, the function $F(x)$ corresponds to the pressure of a gas of rods of length k and activity x on the discrete lattice $\mathbb{Z}$.

The exponential $\exp(F(x))$ is easily recognized as the generating function for (generalized) Catalan numbers, whose definition we recall below. Thus we are looking for a combinatorial interpretation of the logarithm of the generating function for (generalized) Catalan numbers. Logarithms of Catalan generating functions have in fact attracted interest since Knuth's Christmas lecture [7]; to the best of our knowledge, the focus has been on the computation of coefficients, with the question of combinatorial interpretation left open. We provide several such interpretations, among them one with cycle-rooted labeled trees. For the interpretation

*Mathematisches Institut, Ludwig-Maximilians-Universität, Theresienstr. 39, 80333 München, Germany.
†Munich Center for Quantum Science and Technology (MCQST), Schellingstr. 4, 80799 München.

it is essential that we work with *labeled* combinatorial species, as is manifest already for a simple special case: For $k = 1$, the solution to (1.1) is

$$F(x) = -\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n} = \sum_{n=1}^{\infty} \frac{x^n}{n!}(n-1)!.$$

As $1/n$ is not an integer, the function F is not an ordinary generating function, but it is the exponential generating function for a labeled structure, namely for cycles.

Let us recall some facts about Catalan numbers. The sequence of natural numbers $(C_n)_{n\geq 0}$ with

$$C_n := \frac{2n!}{(n+1)!n!}, \quad n \geq 0,$$

is commonly referred to as Catalan numbers since the 1970's. The name goes back to Eugène Charles Catalan who was the first to introduce Catalan numbers in the above form, after they already appeared in literature as far back as the 18th century, most prominently in the work of Leonhard Euler.

Catalan numbers emerge in a huge variety of different counting problems: Over 200 possible interpretations are listed in the monograph [12] by R. P. Stanley alone; many of those are of great significance in the field of combinatorics. Two especially prominent types of structures enumerated by Catalan numbers are discrete paths (e.g., Dyck or Motzkin paths) and tree graphs (e.g., binary or plane trees) under certain restrictions, see items $4-56$ in [12, Chapter 2]).

The generating function G_2 of Catalan numbers $(C_n)_{n\geq 0}$ is given by the formal power series

$$G_2(x) := \sum_{n\geq 0} C_n x^n = 1 + \sum_{n\geq 1} \frac{x^n}{n!}\frac{(2n)!}{(n+1)!}.$$

Naturally, one can view G_2 as the ordinary generating function for any of the over 200 unlabeled structures in [12] or as the exponential generating function for any of the corresponding labeled structures (in the sense of combinatorial species and associated generating functions, see [1]). In particular, we will view G_2 as the exponential generating function for labeled lattice paths (see Section 2) or for labeled binary trees (see Section 3).

The generating function G_2 can be generalized to the following formal power series: For $k \geq 2$, consider the power series G_k, sometimes called the *binomial series* [2, 11], given by

$$G_k(x) := 1 + \sum_{n\geq 1} \frac{x^n}{n}\binom{kn}{n-1} = 1 + \sum_{n\geq 1} \frac{x^n}{n!}(n-1)!\binom{kn}{n-1}$$

and let us refer to the coefficients

$$\frac{1}{n}\binom{kn}{n-1}, \quad n \geq 1,$$

as *generalized* k^{th} *Catalan numbers* following the terminology in [5] (also known under the name of Fuss-Catalan numbers [9]); notice that the Catalan numbers $(C_n)_{n\geq 1}$ are indeed recovered for $k = 2$. The power series G_k satisfies

$$G_k(x) = 1 + xG_k(x)^k.$$

It is well-known (see [5]) that generalized k^{th} Catalan numbers enumerate monotone lattice paths, the so called k-good paths, or alternatively plane k-ary trees. Therefore, we can and will interpret G_k as the exponential generating function for labeled lattice paths (see Section 2) or for labeled plane k-ary trees (see Section 3).

The main object of study in this paper is the logarithm of the generating function G_k for $k \geq 2$, which again can be represented by a formal power series. Explicit expressions for the coefficients are already known from the literature: The expansion of $\log G_2$ was presented 2014 in the annual Christmas lecture by Donald Knuth [7] — who subsequently posed an elegant conjecture for the expansion of $\log^2 G_2$ as a problem in [8] to be solved by various authors soon after:

$$\log^2 G_2 = \sum_{n\geq 2} \frac{x^n}{n} \binom{2n}{n} (H_{2n-1} - H_n),$$

where the *harmonic numbers* $(H_m)_{m\geq 0}$ are given by $H_m := \sum_{i=1}^{m} \frac{1}{i}$ for $m \in \mathbb{N}$.

Higher powers $\log^a G_k$, $a \geq 2$, were examined in [2] and [11], explicit formulas for the coefficients were derived — in terms of harmonic numbers in the former and in terms of Stirling cycle numbers in the latter work. The proofs are of algebraic nature and involve general inversion formulas — in particular, the Lagrange inversion formula.

Here, we present a combinatorial, bijective proof providing explicit expressions for the coefficients of $\log G_k$ by means of exact enumeration. The proof easily generalizes to the case of the higher powers $\log^a G_k$, $a \geq 2$. For example, in the aforementioned case of the squared logarithm $\log^2 C_k$, we obtain

$$[x^n] \log^2 G_k(x) = 2 \sum_{p=2}^{n} \frac{k-1}{kn-p} \binom{kn-p}{n-p} H_{p-1},$$

where $H_m := \sum_{i=1}^{m} \frac{1}{i}$ for $m \in \mathbb{N}$. Naturally, this expression for the coefficients can be rewritten to match the one by Knuth presented above. In the general case $a \geq 1$, we get the formula

$$[x^n] \log^a G_k = \sum_{p=a}^{n} c_{k,n}^{(p)} N_{p,a},$$

where

$$c_{k,n}^{(p)} = \frac{kp-p}{kn-p} \binom{kn-p}{n-p}$$

and

$$N_{p,a} := \sum_{\substack{(q_1,\ldots,q_a)\in[p]^a \\ q_1+\ldots+q_a=p}} \frac{1}{\prod_{i=1}^{a} q_i}.$$

While identifying the coefficients of $\log^a C_k$ for $k \geq 2$ and $a \geq 1$ is not a novel result (since those are known from [2] and [11]), we think that our proof itself is of interest — as we are not aware of any alternative proof that is essentially non-algebraic in nature.

The article is organized as follows: In Section 2, we provide a combinatorial interpretation of $\log G_k$ in terms of lattice paths (Theorem 2.5) by using bijective results identifying lattice paths with sets of certain paths or path-like structures (Lemma 2.4 and Lemma 2.6). Additionally, we use this interpretation to provide a closed expression for coefficients of $\log G_k$ (Theorem 2.8) via a purely combinatorial proof, which can be easily generalized to higher powers $\log^a G_k$, $a \geq 2$ (Theorem 2.9). In Section 3 we provide an alternative interpretation of $\log G_k$ in terms of plane trees (Theorem 3.5). Again, at the heart of this interpretation is a bijective result identifying k-ary trees with sets of certain trees or tree-like structures (Lemma 3.4 and Lemma 3.6). Finally, in the appendix, a method to encode both lattice paths and plane trees via *cyclically ordered multisets* is introduced, providing a bijection between the two combinatorial species and establishing a direct connection between the two combinatorial interpretations of $\log G_k$.

2 Combinatorial interpretation via lattice paths

2.1 Lattice paths and associated generating functions

In this section, we want to consider a combinatorial interpretation of (generalized) Catalan numbers in terms of monotone lattice paths and understand the logarithm of the corresponding generating functions on the level of these combinatorial structures. We concentrate on item 24 in [12, Chapter 2], but consider labeled structures instead of unlabeled.

Definition 2.1 (Labeled good paths)**.** *Let $n \in \mathbb{N}$ and let $k \geq 2$. Let $V \subset \mathbb{N}$ be a finite label set with $|V| = n$. A path on the quadratic lattice $\mathbb{Z}^2$ from $(0,0)$ to $(n,(k-1)n)$ with steps $(0,1)$ or $(1,0)$, together with a labeling of the heights $\{(k-1)j\}_{0\leq j\leq n-1}$ by elements of V (as visualized in Figure 2), is called a V-labeled k-good path if it never rises above the line $y = (k-1)x$. Denote the set of all such paths by $\mathscr{P}_k(V)$ and write $\mathscr{P}_k(n) := \mathscr{P}_k([n])$.*

Remark 2.1. By labeling we mean a bijective map from $\{(k-1)j\}_{0\leq j\leq n-1}$ to V. Notice these heights are exactly those on which the path can potentially intersect the diagonal $y = (k-1)x$.

Remark 2.2. Our notion of (unlabeled) good paths is essentially the same as introduced in [5], up to a vertical shift of the path by 1. Notice that, by [5], G_k — as the generating function for k^{th} generalized Catalan numbers — is equal to the exponential generating function for $(\mathscr{P}_k(n))_{n\in\mathbb{N}_0}$, i.e.,

$$G_k(x) = 1 + \sum_{n\geq 1} \frac{x^n}{n!}|\mathscr{P}_k(n)|.$$

Next, we want to introduce combinatorial structures that are enumerated by the coefficients of $\log G_k$.

Definition 2.2 (Label-minimal good paths)**.** *Let $n \in \mathbb{N}$ and let $k \geq 2$. Let $V \subset \mathbb{N}$ be a finite label set with $|V| = n$. A V-labeled k-good path P is called label-minimal if the label of the height 0 is minimal under all labels labeling heights at which P intersects the diagonal $y = (k-1)x$.*

Denote the set of V-labeled k-good paths that are label-minimal by $\mathscr{P}_k^{\min}(V)$ and write $\mathscr{P}_k^{\min}(n) := \mathscr{P}_k^{\min}([n])$. The corresponding exponential generating function is defined by the following formal power series:

$$G_k^{\min}(x) = \sum_{n\geq 1} \frac{x^n}{n!}|\mathscr{P}_k^{\min}(n)|.$$

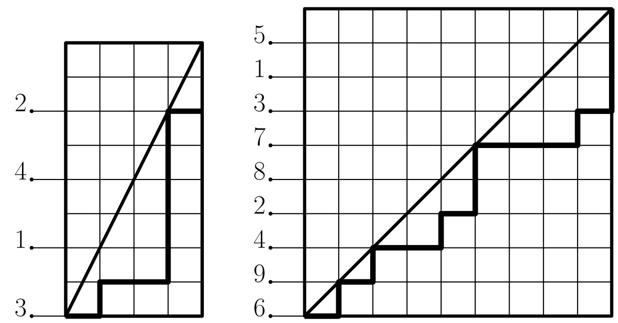

Figure 1: On the left side, we see a 3-good lattice path of length 12 (labeled by $\{1,2,3,4\}$); on the right side, we see a 2-good lattice path of length 18 (labeled by $\{1,\dots,9\}$).

Let $1 \le \ell \le n$ and let $B_1 \cup \ldots \cup B_\ell$ be a partition of $[n]$. For every $i \in [\ell]$, let $P_i \in \mathscr{P}_k^{\min}(B_i)$. The set $\{P_1,\dots,P_\ell\}$ is called a label-minimal k-field on $[n]$. Denote the set of all label-minimal k-fields on $[n]$ by $\mathscr{F}_k^{\min}(n)$.

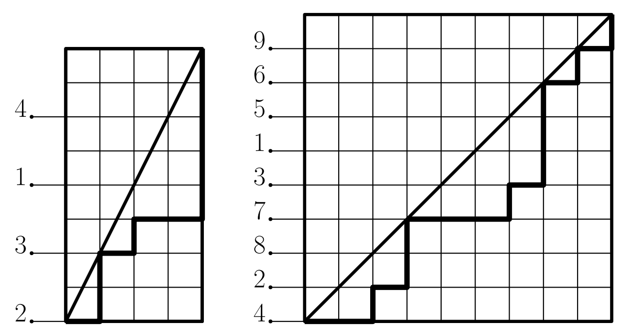

Figure 2: On the left side, we see a label-minimal 3-good lattice path of length 12 (labeled by $\{1,2,3,4\}$); on the right side, we see a label-minimal 2-good lattice path of length 18 (labeled by $\{1,\dots,9\}$).

Alternatively, just like the logarithm of the exponential generating function for permutations can be interpreted as the exponential generating function for cycles (as explained in the introduction), one can interpret $\log G_k$ via certain cyclic structures as well. Informally speaking, those cyclic structures can be obtained by "bending k-good paths into circles", i.e., by identifying endpoints of $[n]$-labeled k-good paths with their starting points and keeping the labelings (which thus become cycles on $[n]$).

Definition 2.3 (Labeled ornaments)**.** *Let $n \in \mathbb{N}$ and let $k \ge 2$. Let V be a finite label set with $|V| = n$. For $P \in \mathscr{P}_k(V)$ construct a labeled infinite lattice path $\hat{P}$ by taking (infinitely many) labeled paths $j(n,(k-1)n) + P$, $j \in \mathbb{Z}$, and concatenating them (while keeping the labeling).*

An equivalence relation on the set $\mathscr{P}_k(V)$ can be defined as follows: Let two V-labeled k-good paths P_1 and P_2 be equivalent if and only if $\hat{P}_1$ is a translate of $\hat{P}_2$ along the line $y = (k-1)x$ (including the labeling).

The corresponding equivalence classes $[P]$ can be identified with the shape of the infinite periodic paths $\hat{P}$ together with an infinite periodic labeling (i.e., a cycle on $[n]$) which are *obtained by identifying the endpoint and the starting point of P.*

Denote the set of the equivalence classes, called V-labeled k-ornaments, by $\mathscr{P}_k^\circ(V)$ and write $\mathscr{P}_k^\circ(n) := \mathscr{P}_k^\circ([n])$. The corresponding exponential generating function is defined by the following formal power series:

$$G_k^\circ(x) = \sum_{n \geq 1} \frac{x^n}{n!} |\mathscr{P}_k^\circ(n)|.$$

Let $1 \leq \ell \leq n$ and let $B_1 \cup \ldots \cup B_\ell$ be a partition of $[n]$. For every $i \in [\ell]$, let O_i be a B_i-labeled k-ornament. The set $\{O_1, \ldots, O_\ell\}$ is called a k-ornament field on $[n]$. Denote the set of all k-ornament fields on $[n]$ by $\mathscr{F}_k^\circ(n)$.

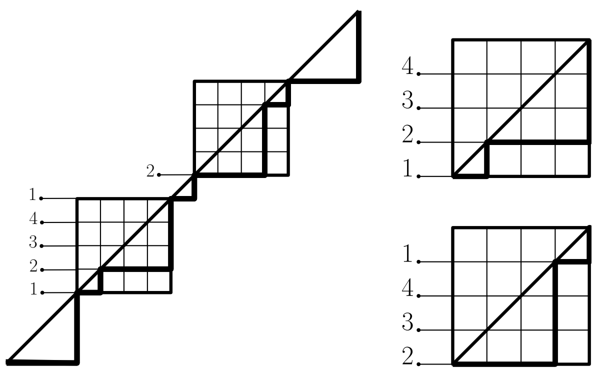

Figure 3: Both 2-good paths of length 8 depicted on the right side are representatives of the 2-ornament depicted on the left side.

Now, with these definitions at hand, we are ready to give a combinatorial interpretation for $\log G_k$ in terms of label-minimal k-good paths or, alternatively, in terms of k-ornaments.

2.2 Bijective results

The following lemma provides the combinatorial insight essential to the proofs of the main results in this section: It enables us to identify labeled good paths with sets of label-minimal good paths.

Lemma 2.4. *Let $n \in \mathbb{N}$ and let $k \geq 2$. There is a bijection between $\mathscr{P}_k(n)$ and $\mathscr{F}_k^{\min}(n)$.*

Proof. Let us define a bijection $m : \mathscr{P}_k(n) \to \mathscr{F}_k^{\min}(n)$. For a k-good path $P \in \mathscr{P}_k(n)$, we obtain a label-minimal k-field $m(P) \in \mathscr{F}_k^{\min}(n)$ from P by the following inductive procedure:

Step 0: Set $\Pi = P$.

Step $N \geq 1$: Let $0 = y_1 < \ldots < y_\ell$ denote the heights at which the path Π intersects the line $y = (k-1)x$ and let $i_1, \ldots, i_\ell \in [n]$ denote the corresponding labels.

- If there exists a $j \in [\ell]$ such that $i_j < i_1$ holds, set $y := \min\{y_j : j \in [\ell],\ i_j < i_1\}$ and set $\Pi = P$. Cut the path Π at the height y, obtaining two paths — a path

Π_1 from $(0,0)$ to $(\frac{y}{k-1}, y)$ and a path Π_2 starting at $(\frac{y}{k-1}, y)$ which inherit their labelings from Π. Π_1 and Π_2 are again k-good paths — up to a translation of Π_2. Replace Π with the translate of Π_2 starting in $(0,0)$ and GOTO Step N+1.

- Otherwise STOP.

Naturally, this procedure produces a label-minimal k-field on $[n]$.

Conversely, given a label-minimal k-field $F \in \mathscr{F}_k^\circ(n)$, construct an $[n]$-labeled k-good path $m^{-1}(F) \in \mathscr{P}_k(n)$ as follows: Order the labeled k-good paths from F decreasing in the label at $y = 0$. Successively, glue the predecessor path to the successor path by concatenation (identifying the endpoint of the former with the starting point of the latter). Naturally, the resulting lattice path is an $[n]$-labeled k-good path and the described procedure does indeed define the inverse of the map m introduced above. □

Remark 2.3. Clearly, our choice of label-minimal paths is somewhat arbitrary in the following sense: In the inductive procedure from Lemma 2.4 defining the map m, one can choose different rules to "cut" the path P at its intersections with the diagonal. E.g., one could instead consider "label-maximal" paths (or, more generally, define y as the height labeled minimally with respect to an arbitrary order on the labels instead of the canonical one).

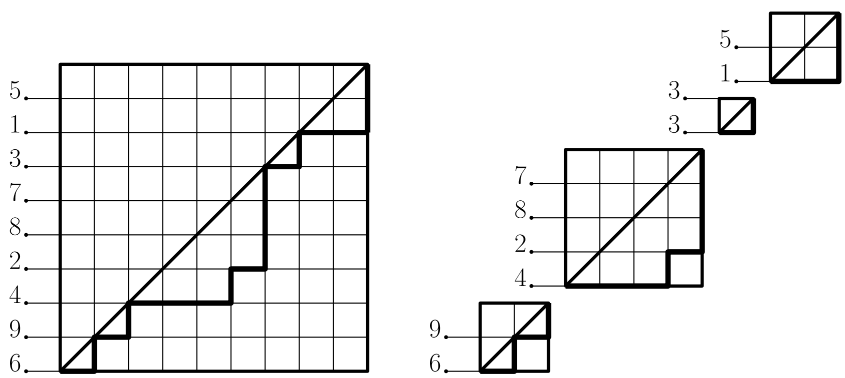

Figure 4: On the left, we see a 2-good path of length 18, on the right we see the label-minimal 2-field corresponding to it in the sense of the proof of Lemma 2.4.

This bijective result allows us to interpret $\log G_k$ as the exponential generating function for label-minimal good paths:

Theorem 2.5. *Let $k \geq 2$. The following holds as an identity between formal power series:*

$$\log G_k = G_k^{\min}.$$

Proof. The claim follows directly from Lemma 2.4 via a standard combinatorial argument (see, e.g., [4], for the argument formulated in the framework of combinatorial species). □

Lemma 2.6. *For $n \in \mathbb{N}$ and $k \geq 2$, there is a bijection between the sets $\mathscr{P}_k^{\min}(n)$ and $\mathscr{P}_k^\circ(n)$.*

Proof. The bijection is given by assigning to the label-minimal path $P \in \mathscr{P}_k^{\min}(n)$ its equivalence class $[P] \in \mathscr{P}_k^\circ(n)$. This map is clearly invertible since every element of $\mathscr{P}_k^\circ(n)$ has a unique representative $P \in \mathscr{P}_k^{\min}(n)$ that is label-minimal. □

Remark 2.4. Again, we see that the choice of label-minimal paths was somewhat arbitrary: In the above proof, one could identify $[P] \in \mathscr{P}_k^\circ(n)$ with a representative different from P, e.g., with the "label-maximal" path in $[P]$, see Remark 2.3.

The lemma allows us to identify $\log G_k$ with the exponential generating function for labeled k-ornaments:

Theorem 2.7. *Let $k \geq 2$. The following holds as an identity between formal power series:*

$$\log G_k = G_k^\circ.$$

Proof. The claim follows from Theorem 2.5 and Lemma 2.6 since the latter implies that $G_k^\circ = G_k^{min}$ for $k \geq 2$. □

We have shown how taking the logarithm of the generating function for k^{th} Catalan numbers G_k can be interpreted on the level of lattice paths. By Theorem 2.5, it can be interpreted as the exponential generating function for label-minimal k-good paths — so that taking the logarithm of G_k corresponds to discarding those k-good paths that have labels at height 0 which are not minimal among the labels labeling intersections of the path with the diagonal $y = (k-1)x$. Alternatively, by Theorem 2.7, $\log G_k$ can be interpreted as the exponential generating function for k-ornaments — so that taking the logarithm corresponds to identifying those k-good paths that result in the same k-ornament when they are "bent into a circle".

2.3 Identifying the coefficients

Lemma 2.4 also provides an elementary way to recover the explicit expressions for the coefficients of $\log G_k$ for every $k \geq 2$ (known from [6, 7]) — by simply counting k-ornaments.

Theorem 2.8. *Let $k \geq 2$ and $n \in \mathbb{N}$. We have*

$$\log(G_k(x)) = \sum_{n \geq 1} \frac{x^n}{n!} \frac{(kn-1)!}{(kn-n)!}.$$

Proof. A well-known result (see, e.g., [3]) provides the number $c_{k,n}^{(p)}$ of Dyck paths of length kn with exactly $p \in \mathbb{N}$ returns to zero (which corresponds to the number of unlabeled k-good paths of length kn that intersect the diagonal $y = (k-1)x$ exactly $p+1$ times):

$$c_{k,n}^{(p)} = \frac{kp-p}{kn-p} \binom{kn-p}{n-p}.$$

Notice that if some k-good lattice path P intersects the line $y = (k-1)x$ exactly $p+1$ times then the same holds for every path in $[P]$ and $|[P]| = p$ (since choosing a representative of $[P]$ is equivalent to choosing which intersection point to place at $y = 0$). Therefore, the number of $[n]$-labeled k-ornaments intersecting the diagonal $y = (k-1)x$ exactly p times (for any representative, counting starting point and endpoint as one intersection) is given by

$$n! \frac{c_{k,n}^{(p)}}{p}$$

and thus we get

$$[x^n]\log G_k = \sum_{p=1}^{n} \frac{c_{k,n}^{(p)}}{p} = \sum_{p=1}^{n} \frac{k-1}{kn-p}\binom{kn-p}{n-p} = \frac{1}{n!}\frac{(kn-1)!}{(kn-n)!}.$$

$\square$

The presented proof of the preceding theorem has the following advantage: It can be easily modified to investigate the coefficients of $\log^a G_k$ for higher powers $a \geq 2$. As mentioned in the introduction, the result itself is not novel and similar expressions for the coefficients are known from [2, 11].

Theorem 2.9. *Let $a, k \geq 2$ and $n \in \mathbb{N}$. We have*

$$[x^n]\log^a G_k = \sum_{p=a}^{n} c_{k,n}^{(p)} N_{p,a},$$

where

$$c_{k,n}^{(p)} = \frac{kp-p}{kn-p}\binom{kn-p}{n-p}$$

and

$$N_{p,a} := \sum_{\substack{(q_1,\ldots,q_a)\in[p]^a \\ q_1+\ldots+q_a=p}} \frac{1}{\prod_{i=1}^{a} q_i}.$$

Remark 2.5. In the special case $a = 2$, considered by Knuth in [8], we get

$$[x^n]\log^2 G_k = \sum_{p=2}^{n} c_{k,n}^{(p)} \sum_{1\leq q\leq p-1} \frac{1}{q(p-q)} = 2\sum_{p=2}^{n} \frac{k-1}{kn-p}\binom{kn-p}{n-p} H_{p-1},$$

where $H_m := \sum_{i=1}^{m} \frac{1}{i}$ for $m \in \mathbb{N}$.

Proof. By Theorem 2.5 and by a standard combinatorial argument, $\log^a G_k$ is the exponential generating function for k-ornament fileds consisting of $a \geq 2$ k-ornaments. For every $n \in \mathbb{N}$, we need to determine the number of such k-ornament fields on $[n]$. To do so, we employ the same decomposition as in the proof of Theorem 2.8 sorting the k-ornament fields by the total number of intersections with the diagonal $y = (k-1)x$ (in any corresponding set of representatives). So, let $\hat{N}_{p,a}^{(k,n)}$ denote the number of k-ornament fields on $[n]$ consisting of precisely $a \geq 2$ k-ornaments such that in total there are p intersections with the diagonal $y = (k-1)x$ (for any representative, counting starting point and endpoint as one intersection). Then

$$[x^n]\log^a G_k = \frac{a}{n!}\sum_{p=a}^{n} \hat{N}_{p,a}^{(k,n)}.$$

In the proof of Theorem 2.8, we already established that

$$n!\frac{c_{k,n}^{(p)}}{p}$$

is the number of $[n]$-labeled k-ornaments O intersecting the diagonal $y = (k-1)x$ exactly p times (for any representative, counting starting point and endpoint as one intersection). We now want to determine how many k-ornament fields of precisely $a \geq 2$ k-ornaments correspond to each such k-ornament O — in the sense that they can be obtain by cutting O at precisely $a \geq 2$ intersections with the diagonal $y = (k-1)x$. This number is exactly the number of possible decompositions of a cycle of length p into $a \geq 2$ segments which is given by

$$\sum_{\substack{(q_1,\ldots,q_a)\in[p]^a \\ q_1+\ldots+q_a=p}} \frac{p}{a},$$

where the tuple $(q_1, \ldots, q_a)$ corresponds to the lengths of the segments, the factor p corresponds to the possible choice of the starting point for the first segment and the factor $\frac{1}{a}$ is due to the fact that there are $a \geq 2$ sequences $(q_1, \ldots, q_a)$ corresponding to the same cycle on $\{q_1, \ldots, q_a\}$.

Left to notice is the following: Consider a k-ornament field of $a \geq 2$ k-ornaments and let the corresponding numbers of intersections with the diagonal $y = (k-1)x$ be given by a fixed sequence $(q_1, \ldots, q_a)$ with $q_1 + \ldots + q_a = p$. From how many distinct k-ornaments intersecting the diagonal $y = (k-1)x$ precisely p times can this k-ornament be obtained by the cutting procedure described above? Naturally, this is equivalent to asking how many different cycles on $[p]$ can be cut to obtain a set of $a \geq 2$ cycles with lengths $(q_1, \ldots, q_a)$ and the answer is just given by the number $\prod_{i=1}^{k} q_i$.

Thus the number $\hat{N}_{p,a}^{(k,n)}$ is given by

$$\hat{N}_{p,a}^{(k,n)} = n!\frac{c_{k,n}^{(p)}}{p} \sum_{\substack{(q_1,\ldots,q_a)\in[p]^a \\ q_1+\ldots+q_a=p}} \frac{p}{a\prod_{i=1}^{a} q_i}$$

and, plugging that in the above expression, we obtain

$$[x^n] \log^a G_k = \frac{a}{n!} \sum_{p=a}^{n} \left(n!\frac{c_{k,n}^{(p)}}{p} \sum_{\substack{(q_1,\ldots,q_a)\in[p]^a \\ q_1+\ldots+q_a=p}} \frac{p}{a\prod_{i=1}^{a} q_i} \right) = \sum_{p=a}^{n} c_{k,n}^{(p)} N_{p,a}.$$

□

3 Combinatorial interpretation via tree graphs

3.1 Tree graphs and associated generating functions

In this section, we provide an alternative combinatorial interpretation for the logarithm of the binomial series G_k in terms of tree graph structures. To do so, we introduce several sets of labeled graphs.

Definition 3.1 (Rooted plane trees)**.** *Let $k \geq 2$. For a finite set $V \subset \mathbb{N}$, we define a rooted plane k-ary tree with the vertex set V as follows: Consider a quadruple $(V, E, r, (\ell(v))_{v\in V})$ such that*

1. $r \in V$, $E \subset \binom{V}{2}$,

2. *the graph* (V, E, r) *is a tree rooted in* r,

3. *for each vertex* $v \in V$, *the set* $C(v) \subset V$ *of children of* v *in* (V, E, r) *satisfies the constraint* $|C(v)| \leq k$,

4. *for each vertex* $v \in V$, $\ell(v) : C(v) \to \{1, \ldots, k\}$ *is an injective map.*

For each vertex $v \in V$, *we interpret the numbers* $\{1, \ldots, k\}$ *as an ordered list of slots potentially available for the children of* v. *We say that the* p^{th} v*-slot is occupied by a vertex* $j \in V$, *if* $j \in C(v)$ *and* $l(v)(j) = p \in \{1, \ldots, k\}$. *We say that the* p^{th} v*-slot is vacant, if such a* j *does not exist. The slots* $\{1, \ldots, k\}$ *are visualized in an increasing order from left to right and vacant slots are depicted by small solid (unlabeled) nodes.*

We denote the set of rooted plane k*-ary trees with the vertex set* V *by* $\mathscr{T}_k(V)$.

Remark 3.1. Vacant slots can be interpreted as unlabeled leaf vertices (compare to the *full binary trees* as in [5]).

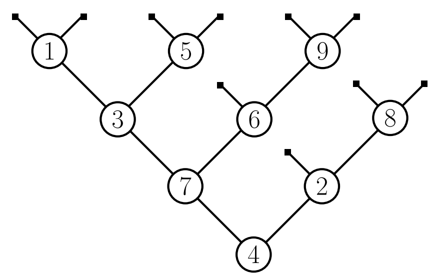

Figure 5: Binary ($d \equiv 2$) tree with $n = 9$ vertices.

Definition 3.2 (Root-minimal plane trees)**.** *Let* $k \geq 2$. *For a finite set* $V \subset \mathbb{N}$, *let* t *be a rooted plane* k*-ary tree with the vertex set* V, *i.e.,* $t \in \mathscr{T}_k(V)$. *We say that a vertex* $v \in V$ *is* on the rightmost branch of t *if* v *is an element of the vertex set* $B \subset V$ *defined via the following induction:*

1. *Let the root* $r \in V$ *be in* B.

2. *If a vertex* $v \in V$ *is in* B, *then the vertex occupying the* k^{th} *(rightmost)* v*-slot is in* B.

We call t root-minimal *if the root* $r \in V$ *is smaller (with respect to the canonical order on the natural numbers) than any of the other vertices on the rightmost branch of the tree. The set of root-minimal plane* k*-ary trees is denoted by* $\mathscr{T}_k^{\min}(V)$. *We denote the set of root-minimal plane* k*-ary forests with the vertex set* V *by* $\mathscr{F}_k^{\min}(V)$.

Definition 3.3 (Cycle-rooted plane trees)**.** *Let* $k \geq 2$. *For a finite set* $V \subset \mathbb{N}$, *we define a cycle-rooted plane* k*-ary tree with the vertex set* V *as follows: Consider a quintuple* $(V, E, R, o, (\ell(v))_{v \in V})$ *such that*

1. $R \subset V$, $E \subset \binom{V}{2}$,

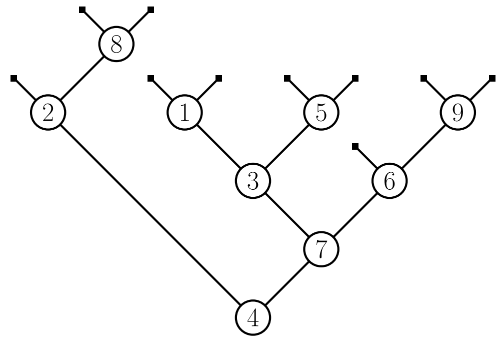

Figure 6: Root-minimal binary tree with $n = 9$ vertices.

2. $(R, E \cap \binom{R}{2})$ *is the cycle graph associated with the cyclic permutation* o *on* R *and is visualized as oriented clockwise,*

3. *the graph* $(V, E \backslash \binom{R}{2}, R)$ *is a forest of* $|R|$ *trees rooted in vertices from* R*,*

4. *for each vertex* $v \in V$*, the set* $C(v) \subset V$ *of children of* v *in* $(V, E \backslash \binom{R}{2}, R)$ *satisfies the constraint* $|C(v)| \leq k$*,*

5. *for each vertex* $v \in V$*,* $\ell(v) : C(v) \to \{1, \ldots, k\}$ *is an injective map; we use the same vocabulary and interpret* $\ell(v)$ *in the same manner as in Definition 3.1,*

6. *For every* $r \in R$*, the* k^{th} *(rightmost)* r*-slot is vacant.*

We denote the set of cycle-rooted plane k*-ary trees with the vertex set* V *by* $\mathscr{T}_k^\circ(V)$ *and the set of cycle-rooted* k*-ary forests with the vertex set* V *by* $\mathscr{F}_d^\circ(V)$*.*

Remark 3.2. Cycle-rooted trees can be interpreted as equivalence classes of rooted plane trees: Two rooted plane trees are equivalent if and only if they result in the same cycle-rooted tree when we identify the root of the tree with its right-most leaf (the right-most branch therefore becoming the cycle sub-graph in the resulting cycle-rooted tree). In this way a cycle-rooted tree with a cycle of length r corresponds to an equivalence class consisting of r rooted plane trees. Compare this to the definition of k-ornaments (Definition 2.3 in Section 2).

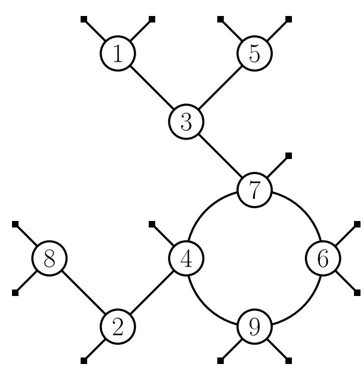

Figure 7: Cycle-rooted binary tree with $n = 9$ internal vertices.

Let $n \in \mathbb{N}$ and let $k \geq 2$. For notational convenience, we set $\mathscr{T}_k(n) := \mathscr{T}_k([n])$ and, analogously, write $\mathscr{T}_k^{\min}(n)$, $\mathscr{T}_k^{\circ}(n)$, $\mathscr{F}_k^{\min}(n)$ and $\mathscr{F}_k^{\circ}(n)$ for any $n \in \mathbb{N}$.

As mentioned in the introduction, the exponential generating function for $(\mathscr{T}_k(n))_{n\in\mathbb{N}}$ is given by the series G_k (see [5]), i.e.,

$$G_k(x) = 1 + \sum_{n\geq 1} \frac{x^n}{n!} |\mathscr{T}_k(n)|.$$

Moreover, we denote

- by $\hat{G}_k^{\min}$ the exponential generating function for $(\mathscr{T}_k^{\min}(n))_{n\in\mathbb{N}}$ given by

$$\hat{G}_k^{\min}(x) = \sum_{n\geq 1} \frac{x^n}{n!} |\mathscr{T}_k^{\min}(n)|,$$

- by $\hat{G}_d^{\circ}$ the exponential generating function for $(\mathscr{T}_d^{\circ}(n))_{n\in\mathbb{N}}$ given by

$$\hat{G}_d^{\circ}(x) = \sum_{n\geq 1} \frac{x^n}{n!} |\mathscr{T}_d^{\circ}(n)|.$$

3.2 Bijective results.

The following lemma is the tree analogue of Lemma 2.4.

Lemma 3.4. *Let $n \in \mathbb{N}$ and $k \geq 2$. There is a bijection between the set of k-ary trees with n vertices $\mathscr{T}_k(n)$ and the set of root-minimal k-ary forests with n vertices $\mathscr{F}_d^{\min}(n)$.*

Proof. We consider the following map m from $\mathscr{T}_k(n)$ to $\mathscr{F}_k^{\min}(n)$. Let $t \in \mathscr{T}_k(n)$, then we obtain the forest $m(t) \in \mathscr{F}_k^{\min}(n)$ from t by the following inductive procedure:

Step 0: Set $i = r$. Set $\ell = i$.

Step $N \geq 1$:

- If the k^{th} (rightmost) ℓ-slot is vacant, STOP.
- If the k^{th} (rightmost) ℓ-slot is occupied by a vertex $j \in V$ and $j < i$, then delete the edge $\{\ell, j\}$, obtaining $N+1$ trees, and leave the k^{th} ℓ-slot vacant. Let all vertices that were roots in the previous step remain roots and let j become the root in the tree to which it belongs. Set $i = j$, $\ell = i$ and GOTO Step $N+1$.
- If the k^{th} (rightmost) ℓ-slot is occupied by a vertex $j \in V$ and $j > i$, then do nothing and the number of trees remains N. All vertices that were roots in the previous step remain roots. If the j-slot $d(j)$ is vacant, STOP. If the k^{th} j-slot is occupied by some vertex, set $\ell = j$ and GOTO Step $N+1$.

Naturally, this procedure produces a forest of root-minimal trees while preserving the vertex set and the offspring constraint k, thus the map $m : \mathscr{T}_k(n) \to \mathscr{F}_k^{\min}(n)$ is well-defined.

Conversely, given a k-ary forest in $\mathscr{F}_k^{\min}(n)$, one can obtain a tree from it by the following procedure: Order the trees of the forest decreasing in the root numbers (with respect to the canonical order on the natural numbers). From this sequence of trees, we obtain a single tree (with the root given by the largest of the initial roots) by successively attaching the successor tree to the predecessor tree as follows: Let j be the last vertex on the rightmost branch of the predecessor tree. We place the root of the successor tree in the vacant k^{th} (rightmost) j-slot, leaving the offspring structure unchanged otherwise. Naturally, this procedure preserves the vertex set and the offspring constraint k as well, and thus defines a map from $\mathscr{F}_k^{\min}(n)$ to $\mathscr{T}_k(n)$ — which clearly is the inverse for the map m defined above. □

Remark 3.3. Naturally, our choice of root-minimal trees is somewhat arbitrary in the following sense: In the proof Lemma 3.4, one can choose a different rule to compare the labels i and j. E.g., one could instead consider "maximal-rooted" trees (or, more generally, use any arbitrary order on the natural numbers instead of the canonical one).

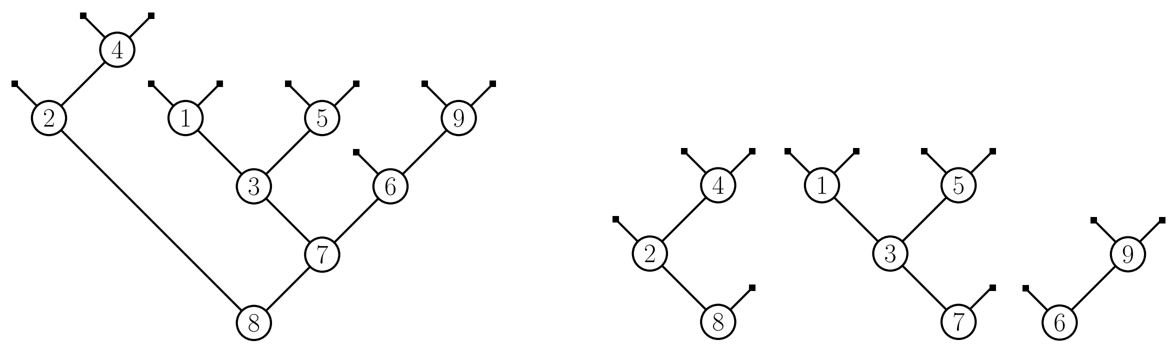

Figure 8: On the left side, we see a binary tree with $n = 9$ vertices; on the right side, we see the root-minimal binary forest corresponding to it in the sense of the proof of Lemma 3.4.

The following theorem is the tree analogue of Theorem 2.5 and a direct consequence of the preceding lemma.

Theorem 3.5. *Let $k \geq 2$. The following holds as an identity between formal power series:*

$$\log \hat{G}_k = \hat{G}_k^{\min}.$$

Proof. Analogously to the proof of Theorem 2.5, the claim follows directly from Lemma 3.4 via a standard combinatorial argument (see, e.g., [4], for the argument formulated in the framework of combinatorial species). □

The following lemma is the tree analogue of Lemma 2.6.

Lemma 3.6. *Let $n \in \mathbb{N}$ and $k \geq 2$. There is a bijection between the sets $\mathscr{T}_k^{\circ}(n)$ and $\mathscr{T}_k^{\min}(n)$.*

Proof. We consider the following map p from $\mathscr{T}_k^{\circ}(n)$ to $\mathscr{T}_k^{\min}(n)$. Starting with a cycle-rooted tree $c \in \mathscr{T}_k^{\circ}(n)$, one obtains a root-minimal tree $p(c) \in \mathscr{T}_k^{\min}(n)$ by the following procedure: For every vertex $r \in R$ on the unique cycle in c, the k^{th} (rightmost) r-slot is vacant by definition. Delete the edge $\{i, j\}$ of the cycle which connects the minimal cycle vertex $i \in R$ with its neighbor in the counter-clockwise direction $j \in R$. Let the minimal cycle vertex i now be the root of the resulting tree and, for every $r \in R \backslash \{j\}$, let the k^{th} r-slot be occupied by the former clockwise neighbor of r on the cycle while leaving the k^{th} j-slot vacant. That way, the former cycle becomes the rightmost brunch of the resulting tree. Otherwise, let the offspring structure be inherited from c. Notice that the resulting tree is indeed in $\mathscr{T}_k^{\min}(n)$, the map p is thus well-defined.

Conversely, to obtain from a root-minimal tree $t \in \mathscr{T}_k^{\min}(n)$ a cycle-rooted tree in $\mathscr{T}_k^{\circ}(n)$ consider the following procedure: Add an edge between the root r of t and the last vertex of the rightmost branch of t, obtaining a cycle. Set $R \subset V$ to be the cycle nodes (that are precisely the vertices on the right-most branch of the original root-minimal tree t). For every cycle node $r \in R$, let the k^{th} (rightmost) r-slot be vacant. Otherwise, for every $v \in V$, let the offspring structure of v be inherited from the map $l(v)$ defining t. Clearly, this procedure provides the inverse to the map p defined above. □

Remark 3.4. If we identify the cycle-rooted trees with equivalence classes of trees as hinted in Remark 3.2, then a bijection is given by just assigning to a root-minimal tree t its equivalence class $[t]$. The map is indeed invertible, since every equivalence class has a unique representative which is root-minimal (compare to the proof of Lemma 2.6).

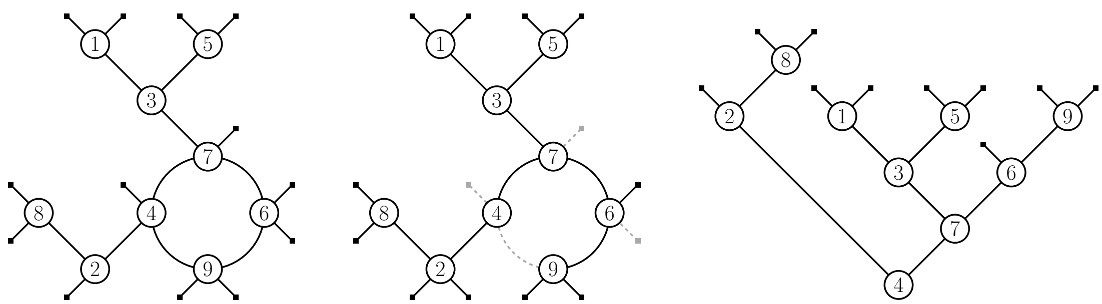

Figure 9: The cycle-rooted tree from Figure 7 (depicted on the left side) corresponds to the root-minimal tree from Figure 6 (depicted on the right side) in the sense of the proof of Lemma 3.6. The construction is illustrated in the middle.

The following theorem follows immediately from Lemma 3.6 and Theorem 3.5. It is the tree analogue of Theorem 2.7:

Theorem 3.7. *Let $k \geq 2$. The following holds as an identity between formal power series:*

$$\log \hat{G}_k = \hat{G}_k^{\circ}.$$

Proof. The claim follows from Theorem 3.5 and Lemma 3.6 since the latter implies that $\hat{G}_k^{\circ} = \hat{G}_k^{\min}$ for $k \geq 2$. □

We have shown how taking the logarithm of the generating function for k^{th} Catalan numbers G_k can be interpreted on the level of trees. By Theorem 3.5, $\log G_k$ can be interpreted as the exponential generating function for root-minimal plane k-ary trees — i.e., taking the logarithm of G_k corresponds to discarding those k-ary trees that have roots that are not minimal among the vertices on the right-most branch of the tree. Alternatively, by Theorem 3.7, $\log G_k$ can be interpreted as the exponential generating function for cycle-rooted k-ary trees — so that taking the logarithm corresponds to identifying those trees that result in the same cycle-rooted tree when their right-most branch is "bent into a circle".

A Cyclic multisets: Encoding lattice ornaments and trees

Here we introduce a way to encode both k-ornaments and cycle-rooted k-ary trees by structures we call cyclically ordered multisets. The rough idea of the encoding is best explained starting from binary rooted trees. Each internal vertex (except for the root) sits on a branch connecting one of its leaf-descendants to the root, and is at the origin of a new branch emanating from it. Enumerating the vertices in the order in which they are visited by a depth-first search, along with the lengths of the associated emanating branches, we obtain sequences $(v(1), \dots, v(n))$, $(f(1), \dots, f(n))$ of labels and branch lengths, with the branch lengths summing up to the total number of vertices. In turn, the branch lengths may be reinterpreted as step heights of lattice paths. Alternatively, we may view the branch lengths $f(j)$ as multiplicities of the element $v(j)$ in some multiset. The precise constructions are more involved as k-ary trees may have more than one branch emanating from internal vertices and the natural structure for cycle-rooted trees is a cycle, rather than an ordered list, of the vertex labels.

For every $n \in \mathbb{N}$ and $k \geq 2$, we will introduce a bijective map π encoding $[n]$-labeled k-ornaments and a bijective map τ encoding cycle-rooted k-ary trees on $[n]$ using the same set of cyclically ordered multisets. Naturally, those maps τ and π induce a bijection between the sets $\mathscr{T}_k^\circ(n)$ and $\mathscr{P}_k^\circ(n)$ for every $n \in \mathbb{N}$ and $k \geq 2$ which can be interpreted as a way to encode k-ary trees by monotone lattice paths and is similar the well-known encoding of binary trees by Dyck paths from [10]. Moreover, the bijections π and τ provide an alternative approach to finding the coefficients of $\log G_k$ — by simply counting cyclically ordered multisets in the image of τ and π. Before we further discuss the encoding, we would like to introduce the set of cyclically ordered multisets rigorously:

Definition A.1 (Cyclically ordered multisets)**.** *Let $k \geq 2$. A cyclically ordered k-multiset (σ, f) on $[n]$ consists of a cycle (cyclic permutation) σ on $[n]$ together with a map $f : [n] \to \mathbb{N}_0^{k-1}$ given by*

$$[n] \ni i \mapsto (f_1(i), \dots, f_{k-1}(i)) \in \mathbb{N}_0^{k-1}$$

such that $\sum_{i=1}^n \sum_{q=1}^{k-1} f_q(i) = n$. To the cycle σ, assign the cycle graph $C_\sigma = (V, E)$, given by

$$V = [n] \times [k-1]$$

and

$$E = \{\{(i,q),(j,p)\} |\ i = j \text{ and } |q-p| = 1 \text{ or } i \text{ is the } \sigma\text{-predecessor of } j,\ q = k-1 \text{ and } p = 1\}.$$

Alternatively, one can view f as a function on the nodes of C_σ, i.e., $f : [n] \times [k-1] \to \mathbb{N}_0$, $(i, q) \mapsto f_q(i)$. We denote the set of cyclically ordered k-multisets on $[n]$ by $\mathscr{M}_k^\circ(n)$.

Let $n \in \mathbb{N}$. In the binary case $k = 2$, one needs the whole set $\mathscr{M}_2^\circ(n)$ to encode the corresponding 2-ornaments or binary trees. For $k \geq 3$, however, the set $\mathscr{M}_k^\circ(n)$ is too big. We introduce a subset of $\mathscr{M}_k^\circ(n)$ which is naturally suited to encode the structures from $\mathscr{P}_k^\circ(n)$ and $\mathscr{T}_k^\circ(n)$:

Definition A.2 (Multisets with root vertices)**.** *Let $m = (\sigma, f) \in \mathscr{M}_k^\circ(n)$, let $i, j \in [n]$ and let $1 \leq k_i, k_j \leq k-1$. We call a simple path on the circle graph C_σ starting in (i, k_i) and ending in (j, k_j) a* segment *of C_σ if $i = j$ and $k_i \leq k_j$ or if it is consistent with the orientation of σ,* *i.e., if $i \neq j$ and for every pair of consecutive points $(\ell_1, k-1), (\ell_2, 1)$ in s we have that ℓ_2 is the σ-sucessor of ℓ_1. To any segment s of C_σ we assign the scope of s given by*

$$\lambda(s) = |\{i \in [n] | (i, m) \in s \text{ for some } 1 \leq m \leq k-1\}|$$

and the weight of s in m given by

$$w^{(m)}(s) = \sum_{(i,m) \in s} f_m(i).$$

For $m \in \mathscr{M}_k^\circ(n)$, we define the set of root vertices $W(m)$ by

$$W(m) := \{i \in [n] | \text{ every segment } s \text{ of } C_\sigma \text{ starting in } (i,1) \text{ satisfies } w^{(m)}(s) \geq \lambda(s)\}.$$

We denote the set of those multisets in $\mathscr{M}_k^\circ(n)$ that possess root vertices by $M(k,n)$, i.e.,

$$M(k,n) := \{m \in \mathscr{M}_k^\circ(n) | W(m) \neq \emptyset\}.$$

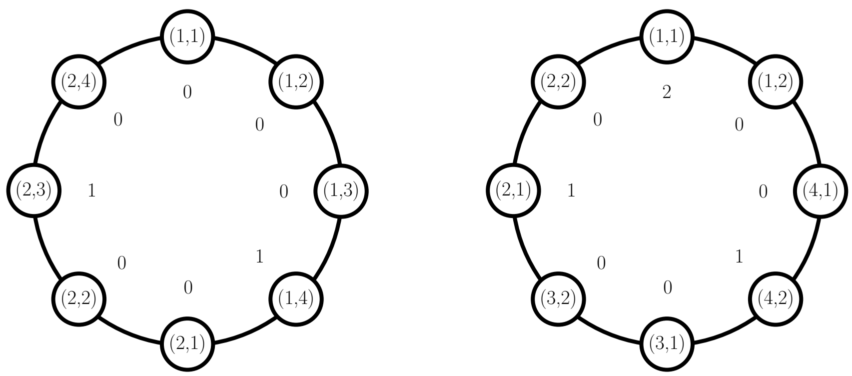

Figure 10: On the left side $m \in \mathscr{M}_5^\circ(2)$ is depicted, on the right side $m' \in \mathscr{M}_3^\circ(4)$. The numbers inside the circle graph depict the multiplicities of the vertices of the circle graph C_σ closest to them. Notice that $m \notin M(5,2)$, but $m' \in M(3,4)$, since $1, 2 \in W(m')$.

Now we can introduce a map encoding lattice ornaments by cyclically ordered multisets:

Definition A.3 (Map π encoding lattice ornaments by multisets)**.** *Let $k \geq 2$ and $n \in \mathbb{N}$. We define the embedding $\pi : \mathscr{P}_k^\circ(n) \to \mathscr{M}_k^\circ(n)$ as follows: For a k-ornament $O \in \mathscr{P}_k^\circ(n)$, we set $\pi(O) =: (\sigma, f)$, where σ is simply given by the labeling of O. To obtain the map f, take any representative of O and set $f_q(i)$, $q \in [k-1]$, $i \in [n]$, to be the number of steps to the right at the height $y = y_i + q - 1$, where y_i is the height labeled by i in O.*

Remark A.1. Naturally, the map π is indeed injective. The property of the path O to not rise above the diagonal $y = (k-1)x$ corresponds to the property $W(\pi(O)) \neq \varnothing$ on the level of multisets. Moreover, the set of labels marking the heights at which O intersects the diagonal becomes the set $W(\pi(O))$. Thus the range $\pi(\mathscr{P}_k^\circ(n))$ of π is given by $M(k,n) := \{m \in \mathscr{M}_k^\circ(n) | \, W(m) \neq \varnothing\}$ so that $|\mathscr{P}_k^\circ(n)| = |M(k,n)|$. For $k = 2$, we have $M(k,n) = \mathscr{M}_k^\circ(n)$ and π is a bijection.

Now we investigate how cycle-rooted trees can be encoded by cyclically ordered multiset. To this end, we introduce the following map:

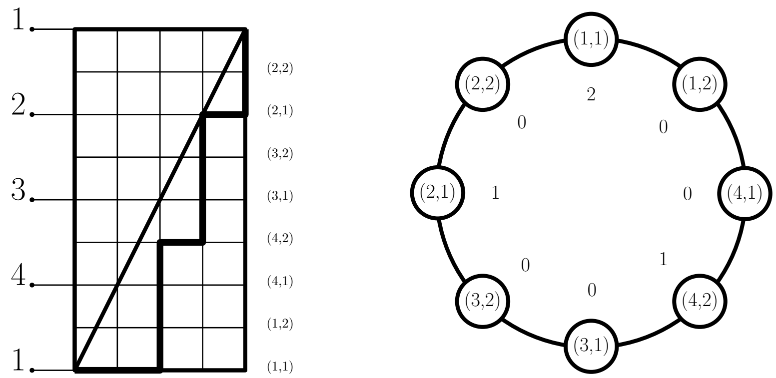

Figure 11: The [4]-labeled 3-ornament corresponding to the 3-good path from Figure 2 (depicted on the left side) is mapped by π to the multiset from $\mathscr{M}_3^\circ(4)$ (depicted on the right side).

Definition A.4 (Map τ encoding cycle-rooted trees by multisets)**.** *Let $k \geq 2$ and $n \in \mathbb{N}$. We introduce an embedding $\tau : \mathscr{T}_k^\circ(n) \to \mathscr{M}_k^\circ(n)$. Given a cycle-rooted tree $t \in \mathscr{T}_k^\circ(n)$, we construct the cyclically ordered multiset $\tau(t) = (f, \sigma) \in \mathscr{M}_k^\circ(n)$ by the following two-step procedure:*

- *Step 1 (Constructing the cycle σ by exploration of vertices in t): Starting at any root of $t \in \mathscr{T}_k^\circ(n)$, the cycle σ is obtained by the following exploration procedure: In every step of the exploration, we uncover a single vertex of t. In the first step, we uncover an arbitrary root r of t. In every further step, as long as there are unexplored vertices in the maximal k-ary subtree of t rooted in r, we go to the last explored vertex that has an unexplored child and uncover its leftmost unexplored child. When the maximal k-ary subtree of t rooted in r is explored, we move to the next root in t according to the cyclic order induced by the oriented cycle of roots t and repeat the procedure. We stop when all vertices of t are explored and define σ as the cycle induced directly by the linear order in which the vertices of t were uncovered.*

- *Step 2 (Define the function f by re-distributing multiplicities of vertices in t): Initially every vertex of t is assigned a single multiplicity. Then the multiplicities are re-distributed between the vertices of t by "rolling-down" (viewed drawing the trees growing upwards with equiangular branches, see Figure 12): Let $i \in [n]$ be an arbitrary vertex of t. For $q \in [k]$, consider the path $\Theta_q(i)$ given by the unique simple path starting in i and ending in its leaf-descendant such that every vertex $j \neq i$ on the path occupies slot q of its parent. Denote by $|\Theta_q(i)|$ the number of vertices on the path $\Theta_q(i)$.*

  If i is a root, i.e. $i \in R$, set
  $$f_q(i) := |\Theta_q(i)| + \delta_{q,1}$$
  for $1 \leq q \leq k-1$.

  If i is not a root, then i is the child of a vertex, say i occupies slot p of its parent. For $1 \leq q \leq k-1$, let q' denote the q-smallest element of $[k]\backslash\{p\}$ and set
  $$f_q(j) := |\Theta_{q'}(j)|.$$

  Notice that $\sum_{j=1}^n \sum_{q=1}^{k-1} f_q(j) = n$ indeed holds for the function f defined above.

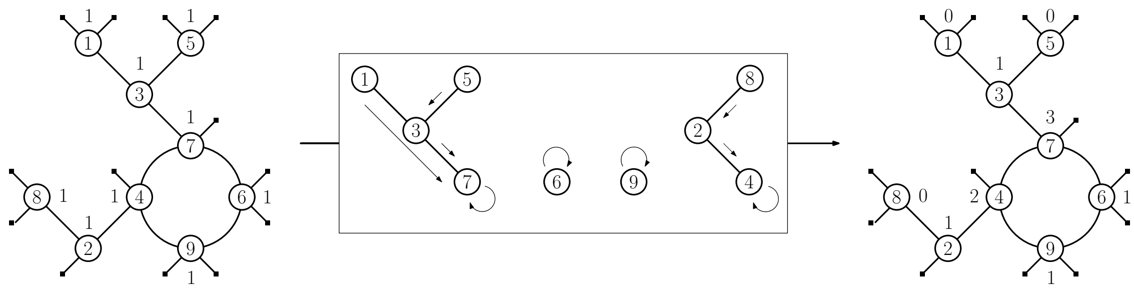

Figure 12: Redistribution of multiplicities from Step 2 of Definition A.4 in the binary case: The multiplicities of non-root vertices "roll down" and the multiplicities of roots do not move.

Remark A.2. The map τ is indeed injective. The set R of roots of t is mapped under τ precisely onto the set $W(\tau(t))$ on the level of multisets. Again, the range $\tau(\mathscr{T}_k^\circ(n))$ of τ is given by $M(k,n)$ so that $|\mathscr{T}_k^\circ(n)| = |M(k,n)|$. In the binary case $k=2$, we have $M(k,n) = \mathscr{M}_k^\circ(n)$ and τ is a bijection.

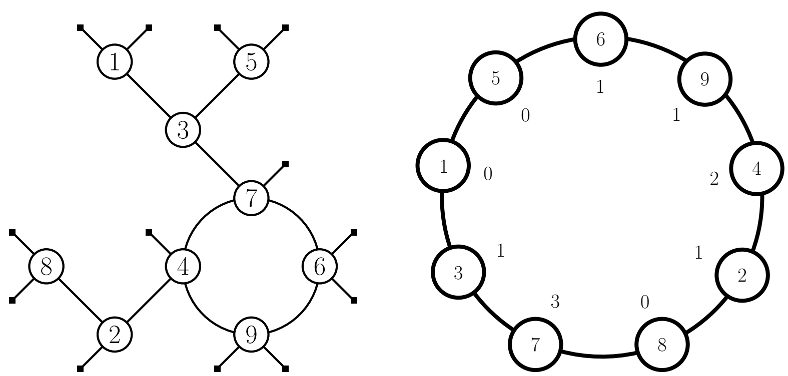

Figure 13: Final result: The cycle-rooted tree from Figure 7 (depicted on the left side) is mapped by τ to the multiset from $\mathscr{M}_2^\circ(9)$ (depicted on the right side).

Let $k \geq 2$ and $n \in \mathbb{N}$. By Remark A.1 and Remark A.2, a bijection between the sets $\mathscr{T}_k^\circ(n)$ and $\mathscr{P}_k^\circ(n)$ is given by the composition $\hat{\pi}^{-1} \circ \tau$, where $\hat{\pi} : \mathscr{P}_k^\circ(n) \to M(k,n)$ is given by $\hat{\pi}(O) = \pi(O)$ for $O \in \mathscr{P}_k^\circ(n)$. Moreover, let $t \in \mathscr{T}_k^\circ(n)$ and $O_t := \hat{\pi}^{-1}(\tau(t))$, then there is a one-to-one correspondence between the roots of t (vertices R of the cycle subgraph of t) and the labels at which O_t intersects the diagonal $y = (k-1)x$. The bijection can be viewed as an alternative to the well-known encoding of binary trees by Dyck paths presented in [10, Chapter 6.3] which also involves a depth-first exploration of the tree (as described in Step 2 of Definition A.4).

Finally, notice the following: It can be shown that the set $M(k,n)$ contains exactly the fraction $\frac{1}{k-1}$ of all elements in $\mathscr{M}_k^\circ(n)$. Since by definition $|\mathscr{M}_k^\circ(n)| = (n-1)! \left(\!\!\binom{(k-1)n}{n}\!\!\right)$ holds, where $\left(\!\!\binom{i}{j}\!\!\right)$ denotes the multiset coefficient and can be written as $\left(\!\!\binom{i}{j}\!\!\right) = \binom{i+j-1}{j}$ for $i, j \in \mathbb{N}$, we have

$$|M(k,n)| = \frac{|\mathscr{M}_k^\circ(n)|}{k-1} = \frac{(n-1!)}{k-1}\binom{kn-1}{n} = \frac{(kn-1)!}{(kn-n)!}.$$

This outlines an alternative proof for Theorem 2.8, since we have $|M(k,n)| = |\mathscr{P}_k^\circ(n)| =$

$|\mathscr{T}_k^\circ(n)|$ and thus $\log G_k$ is the exponential generating function for $(M_k(n))_{n\in\mathbb{N}}$, i.e.,

$$\log G_k(x) = \sum_{n\geq 1} \frac{x^n}{n!}|M_k(n)| = \sum_{n\geq 1} \frac{x^n}{n!}\frac{(kn-1)!}{(kn-n)!}.$$

References

[1] Bergeron, F., Labelle, G., and Leroux, P. *Combinatorial species and tree-like structures*, vol. 67 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1998.

[2] Chu, W. Logarithms of a binomial series: Extension of a series of Knuth. *Math. Commun.*, 24 (2019), 83–90.

[3] Drube, P. Raised k-Dyck paths, 2022. arXiv.2206.01194.

[4] Faris, W. G. Combinatorics and cluster expansions. *Prob. Surveys 7* (2010), 157–206.

[5] Hilton, P., and Pedersen, J. Catalan Numbers, Their Generalization, and Their Uses. *Math. Intell. 13* (03 1991), 64–75.

[6] Jansen, S. Cluster and virial expansions for the multi-species Tonks gas. *J. Stat. Phys. 161*, 5 (2015), 1299–1323.

[7] Knuth, D. E. 3/2-ary trees. Annual Christmas lecture, 2014. https://www.youtube.com/watch?v=P4AaGQIo0HY.

[8] Knuth, D. E. Log-squared of the Catalan generating function. *Amer. Math. Monthly 122:390 (Problem 11832)* (2015).

[9] Lin, C.-H. Some combinatorial interpretations and applications of Fuss-Catalan numbers. *ISRN Discrete Math. 2011* (01 2011).

[10] Pitman, J. *Combinatorial stochastic processes: École d'été de probabilités de Saint-Flour XXXII-2002*. Springer, 2006.

[11] Prodinger, H. Logarithms of a binomial series: A Stirling number approach. *Ars Math. Contemp. 17* (10 2019), 271–275.

[12] Stanley, R. P. *Catalan Numbers*. Cambridge University Press, 2015.

4 The direct-connectedness function in the random connection model

This section contains the published version of:
S. Jansen, L. Kolesnikov, and K. Matzke. The direct-connectedness function in the random connection model. *Adv. in Appl. Probab.*, 1–44, 2022.

Personal contribution:
This article is mainly a collaborative effort with Kilian Matzke (who was a PhD student at the time), with occasional input from my supervisor, Sabine Jansen. Kilian proposed the question and initiated the collaboration in which he had a leading role. In numerous discussions, we developed the ideas for the main results — to which I could contribute my expertise in cluster expansion methods and resummation of series.

Adv. Appl. Probab. 1–44 (2022)
*doi:*10.1017/apr.2022.22

THE DIRECT-CONNECTEDNESS FUNCTION IN THE RANDOM CONNECTION MODEL

SABINE JANSEN,* **
LEONID KOLESNIKOV,* *** AND
KILIAN MATZKE,* **** *Ludwig-Maximilians-Universität München*

Abstract

We investigate expansions for connectedness functions in the random connection model of continuum percolation in powers of the intensity. Precisely, we study the pair-connectedness and the direct-connectedness functions, related to each other via the Ornstein–Zernike equation. We exhibit the fact that the coefficients of the expansions consist of sums over connected and 2-connected graphs. In the physics literature, this is known to be the case more generally for percolation models based on Gibbs point processes and stands in analogy to the formalism developed for correlation functions in liquid-state statistical mechanics.

We find a representation of the direct-connectedness function and bounds on the intensity which allow us to pass to the thermodynamic limit. In some cases (e.g., in high dimensions), the results are valid in almost the entire subcritical regime. Moreover, we relate these expansions to the physics literature and we show how they coincide with the expression provided by the lace expansion.

Keywords: Ornstein–Zernike equation; random connection model; connectedness functions; Poisson process; percolation; graphical expansions; lace expansion

2020 Mathematics Subject Classification: Primary 60K35
Secondary 60G55; 82B43; 60D05

1. Introduction and main result

Perturbation analysis plays an important role in both stochastic geometry [14, Chapter 19] and statistical mechanics. For Gibbs point processes (grand-canonical Gibbs measures in statistical mechanics), quantities like factorial moment densities (also called correlation functions) are highly nontrivial functions of the intensity of the Gibbs point process itself (density) or the intensity of an underlying Poisson point process (activity). When interactions are pairwise, it is well known that the coefficients of these expansions are given by sums over geometric, weighted graphs. There is a vast literature addressing the convergence of these expansions; see, for example, [2, 16]. Some attempts have been made at exploiting power series expansions

Received 29 November 2020; revision received 25 March 2022.

* Postal address: Mathematisches Institut, Ludwig-Maximilians-Universität München, Theresienstraße 39, 80333 München, Germany.
** Email address: jansen@math.lmu.de
*** Email address: kolesni@math.lmu.de
**** Email address: matzke@math.lmu.de

from statistical mechanics for likelihood analysis of spatial point patterns in spatial statistics; see [19].

The physics literature provides similar power series expansions for connectedness functions in a class of percolation models driven by Gibbs point processes, the so-called random connection models (RCMs) [6]. The expansion coefficients for the pair-connectedness function can be written in terms of a sum of certain connected graphs (see (3.1)) and the coefficients for the direct-connectedness function in terms of a sum over certain 2-connected graphs (see (4.1)). The two functions are related via the Ornstein–Zernike equation (OZE) [20], an integral equation which is of paramount importance in physical chemistry and soft matter physics and which enters some approaches to percolation theory; see [25, Chapter 10]. For Bernoulli bond percolation on $\mathbb{Z}^d$, the OZE encodes a renewal structure and is used to prove Ornstein–Zernike behavior [4], a precise asymptotic formula for pair-connectedness functions in the subcritical regime that incorporates subleading corrections to the exponential decay. The OZE also appears as a by-product of lace expansions [10, Proposition 5.2].

The expansions for connectedness functions appearing in [6] are derived as a means of discussing the following question: is it possible to choose the notion of connectivity in such a way that the percolation transition, if it occurs at all, coincides with the phase transition in the sense of non-uniqueness of Gibbs measures? We remind the reader that the relationship between the two phenomena is rather subtle, and in general the corresponding critical parameters do not match; see [12] and references therein. To the best of our knowledge, the question above has not been fully answered for continuum systems, although Betsch and Last [1] were recently able to show that uniqueness of the Gibbs measure follows from the non-percolation of an associated RCM driven by a Poisson point process.

Moreover, the convergence of the expansions for connectedness functions has not been treated in a mathematically rigorous way, in stark contrast with the rich theory of cluster expansions. Even in the simplest case of the RCM driven by a Poisson point process that we consider in this paper, where activity and density coincide and are called the intensity, rigorous results for the expansion of connectedness functions barely exist: the first ones were obtained by Last and Ziesche in [15]. However Last and Ziesche do not prove that their expansions coincide with the physicists' expansion, and they do not prove quantitative bounds for the domain of convergence of the small-intensity expansion.

Our main result addresses graphical expansions of the direct-connectedness function in infinite volume. The results by Last and Ziesche [15], combined with our combinatorial considerations from Section 6.2, imply that the physicists' expansions have a positive radius of convergence; however, it is not our purpose to provide a quantitative bound for the latter. Instead, we perform first a re-summation, in finite volume, of the physicists' expansion. Although the re-summed expansion is no longer a power series in the intensity of the underlying Poisson point process, it has the (conjectured) advantage of converging in a bigger domain than the physicists' expansion. We provide quantitative bounds on the intensity that allow us to pass to the infinite-volume limit in the re-summed expansion of the direct-connectedness function. The proof uses the continuum BK inequality proved in [10].

In addition, we discuss the relationship of the physicists' and our expansion to the lace expansion for the continuum random connection model [10]. Roughly, the lace expansion could in theory be rederived from the graphical expansion by yet another re-summation step. In fact a notion of laces similar to the laces for the self-avoiding random walk [2, 22] already enters the proof of our main result on graphical expansions (see Section 4.3). Thus, contrary to what is stated in [9, Chapter 6.1], the denomination 'lace expansion' for percolation is not a

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

misnomer, at least for continuum systems. It is unclear, however, whether the discussion offers a new angle of attack on the intricate convergence problems in the theory of lace expansions.

Let us properly introduce the RCM and state our results. The RCM depends on two parameters, namely its *intensity* $\lambda \geq 0$ and the (measurable) *connection function* $\varphi : \mathbb{R}^d \to [0, 1]$, satisfying

$$0 < \int \varphi(x) \mathrm{d}x < \infty$$

as well as radial symmetry $\varphi(x) = \varphi(-x)$ for all $x \in \mathbb{R}^d$. The model is described informally as follows: the vertex set is taken to be a homogeneous Poisson point process (PPP) in $\mathbb{R}^d$ of intensity λ, denoted by η. For any pair $x, y \in \eta$, we add the edge $\{x, y\}$ with probability $\varphi(x - y)$ and independently of all other pairs. We refer to [10, 18] for a formal construction.

The RCM is an undirected simple random spatial graph and a standard model of continuum percolation. We denote it by ξ and we use $\mathbb{P}_\lambda$ to denote the corresponding probability measure. Its vertex set is $V(\xi) = \eta$, and we let $E(\xi)$ denote its edge set.

For $x \in \mathbb{R}^d$, we let ξ^x be the RCM augmented by the point x. In other words, the vertex set of ξ^x is $\eta \cup \{x\}$ and the edges are formed as described above. In particular, edges between x and points of η are drawn independently and according to φ. More generally, for a set of points $x_1, \ldots, x_k$, we let $\xi^{x_1,\ldots,x_k}$ be the RCM with vertex set $\eta \cup \{x_1, \ldots, x_k\}$ (also here, edges between deterministic points x_1, x_2 are drawn independently and according to φ).

We say that $x, y \in \eta$ are connected (and write $x \longleftrightarrow y$ in ξ) if there is a path from x to y in ξ. For $x \in \mathbb{R}^d$, we let $\mathscr{C}(x) = \mathscr{C}(x, \xi^x) = \{y \in \eta^x : x \longleftrightarrow y \text{ in } \xi^x\}$ be the cluster of x and define the *pair-connectedness* (or *two-point*) function $\tau_\lambda : \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ to be

$$\tau_\lambda(x, y) := \mathbb{P}_\lambda\big(x \longleftrightarrow y \text{ in } \xi^{x,y}\big). \tag{1.1}$$

Thanks to the translation-invariance of the model, we have $\tau_\lambda(x, y) = \tau_\lambda(\mathbf{0}, x - y)$ $\big($where $\mathbf{0}$ denotes the origin in $\mathbb{R}^d\big)$, and we can also define τ_λ as a function $\tau_\lambda : \mathbb{R}^d \to [0, 1]$ with $\tau_\lambda(x) = \mathbb{P}_\lambda\big(\mathbf{0} \longleftrightarrow x \text{ in } \xi^{\mathbf{0},x}\big)$.

We say that $x, y \in \eta$ are *2-connected* (or *doubly connected*) and write $x \Longleftrightarrow y$ in ξ if there are two paths from x to y that have only their endpoints in common (or if x and y are directly connected by an edge or if $x = y$). We define

$$\sigma_\lambda(x) := \mathbb{P}_\lambda\big(\mathbf{0} \Longleftrightarrow x \text{ in } \xi^{\mathbf{0},x}\big).$$

Recall that the critical intensity for percolation is defined by

$$\lambda_c = \sup\{\lambda \geq 0 : \mathbb{P}_\lambda(|\mathscr{C}(\mathbf{0})| = \infty) = 0\}$$

and that the identity

$$\sup\{\lambda \geq 0 : \mathbb{P}_\lambda(|\mathscr{C}(\mathbf{0})| = \infty) = 0\} = \sup\{\lambda \geq 0 : \int \tau_\lambda(x) \mathrm{d}x < \infty\}$$

has been shown to hold true for connection functions ϕ that are nonincreasing in the Euclidean distance (see [17]). It is proved in [15] that for $\lambda < \lambda_c$, there exists a uniquely defined integrable and essentially bounded function $g_\lambda : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\tau_\lambda(x, y) = g_\lambda(x, y) + \lambda \int_{\mathbb{R}^d} g_\lambda(x, z) \tau_\lambda(z, y) \, \mathrm{d}(z), \qquad x, y \in \mathbb{R}^d. \tag{1.2}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

This equation is known as the Ornstein–Zernike equation (OZE), and g_λ is called the *direct-connectedness function*.

For two integrable functions $f, g : \mathbb{R}^d \to \mathbb{R}$, we recall the convolution $f * g$ to be given by

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x)g(x-y)\mathrm{d}y.$$

We let $f^{*1} = f$ and $f^{*m} = f^{*(m-1)} * f$. Notice that we can interpret both the pair-connectedness function τ_λ and the direct-connectedness function g_λ as functions on $\mathbb{R}^d$, thanks to translation-invariance. The OZE then can be formulated as

$$\tau_\lambda = g_\lambda + \lambda(g_\lambda * \tau_\lambda). \tag{1.3}$$

Naturally, the question arises whether one can provide an explicit form for the direct-connectedness function g_λ. Unfortunately, an immediate probabilistic interpretation of g_λ is not known. One classical approach from the physics literature is to obtain explicit approximations for the solution g_λ of (1.2) by introducing complementary equations, known as closure relations, the choice of which depends on the specifics of the model considered. Different closure relations provide different explicit approximations for g_λ and thus also for the pair-connectedness function τ_λ, e.g., via a reformulation of the OZE (1.2) for the Fourier transforms of the connectedness functions. Most prominent are the Percus–Yevick closure relations [5, 25]; other examples can be found in [7]. Another approach [6] is to directly provide an independent definition of g_λ in terms of a graphical expansion and then argue that this expansion satisfies the OZE (1.2). We follow the spirit of the latter approach: our main result is a graphical expansion for the direct-connectedness function, with quantitative bounds on the domain of convergence.

Let

$$\lambda_* := \sup\left\{\lambda \geq 0 : \sup_{x\in\mathbb{R}^d} \sum_{k\geq 1} \lambda^{k-1}\sigma_\lambda^{*k}(x) < \infty\right\}, \qquad \tilde{\lambda}_* := \sup\left\{\lambda \geq 0 : \lambda \int \sigma_\lambda(x)\,\mathrm{d}x < 1\right\}. \tag{1.4}$$

It is not hard to see that $\tilde{\lambda}_* \leq \lambda_* \leq \lambda_c$ using (1.5) below.

We can now state our main theorem. It provides (in general dimension) the first rigorous quantitative bounds on λ under which the direct-connectedness function admits a convergent graphical expansion.

Theorem 1.1. (Graphical expansion of the direct-connectedness function.) *For $\lambda < \lambda_*$, the direct-connectedness function $g_\lambda(x_1, x_2)$ is given by the expansion* (4.24)*, which is absolutely convergent pointwise for all $(x_1, x_2) \in \mathbb{R}^{2d}$. Moreover, for $\lambda < \tilde{\lambda}_*$, the expansion* (4.24) *converges in the $L^1(\mathbb{R}^d, \mathrm{d}x_2)$-norm for all $x_1 \in \mathbb{R}^d$.*

The convergence results for the expansion (4.24) are proved in Theorem 4.1 and Theorem 4.2; the equality with the direct-connectedness function is proved in Section 5.

Last and Ziesche show that there is some $\lambda_0 > 0$ such that g_λ is given by a power series for $\lambda \in [0, \lambda_0)$. No quantitative bounds for λ_0 are provided, however. In Section 6.2, we discuss how to relate this expansion to our expression for g_λ. We now make several remarks on Theorem 1.1 and the quantitative nature of the bounds provided there.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

- Since $0 \le \sigma_\lambda \le 1$, we can bound

$$\sum_{k\ge 1} \lambda^{k-1}\sigma_\lambda^{*k}(x) \le \sum_{k\ge 0}\left(\lambda \int \sigma_\lambda(x)\,\mathrm{d}x\right)^k = \sum_{k\ge 0}\left(\mathbb{E}_\lambda\left[\left|\left\{x\in\eta : \mathbf{0} \Longleftrightarrow x \text{ in } \xi^{\mathbf{0}}\right\}\right|\right]\right)^k, \quad (1.5)$$

  where the identity is due to the Mecke equation (2.1). This shows that $\tilde{\lambda}_* \le \lambda_*$ and that $\tilde{\lambda}_*$ is the point where the expected number of points in η that are 2-connected to the origin passes 1 (i.e., we have $\mathbb{E}_\lambda\left[\left|\left\{x\in\eta : \mathbf{0} \Longleftrightarrow x \text{ in } \xi^{\mathbf{0}}\right\}\right|\right] \ge 1$ for all $\lambda > \tilde{\lambda}_*$).

- The argument of the geometric series in (1.5) can be further bounded from above by

$$\lambda \int \tau_\lambda(x)\mathrm{d}x = \mathbb{E}_\lambda\left[\left|\left\{x\in\eta : \mathbf{0} \longleftrightarrow x \text{ in } \xi^{\mathbf{0}}\right\}\right|\right],$$

  the expected cluster size (minus 1). A classical branching-process argument gives that $\tilde{\lambda}_* \ge 1/2$ (see, for example, [21, Theorem 3]).

- In high dimension, we have the following result, proven in [10]: under some additional assumptions on φ (see [10, Section 1.2]), there is an absolute constant c_0 such that

$$\lambda_c \int \sigma_{\lambda_c}(x)\,\mathrm{d}x \le 1 + c_0/d$$

  in sufficiently high dimension, or, for a class of *spread-out* models (closely related to Kac potentials in statistical mechanics; see [8]) with a parameter L,

$$\lambda_c \int \sigma_{\lambda_c}(x)\,\mathrm{d}x \le 1 + c_0 L^{-d}$$

  for all dimensions $d > 6$ (in the spread-out case, c_0 is independent of L but may depend on d). As σ_λ is nondecreasing in λ, this provides a bound for the whole subcritical regime. This also implies that for every $\varepsilon > 0$, there is d_0 (respectively, L_0) such that $\tilde{\lambda}_* \ge 1-\varepsilon$ for all $d \ge d_0$ (respectively, $L \ge L_0$ and $d > 6$). As we also know that $\lambda_c \searrow 1$ as the dimension becomes large, this shows that in high dimension, $\tilde{\lambda}_*$ (and thus also λ_*) gets arbitrarily close to λ_c.

Outline of the paper. The paper proceeds as follows. We introduce most of our important notation in Section 2. This allows us to demonstrate some basic (and mostly well-known) central ideas in Section 3, where the two-point function is discussed in finite volume. Section 4 contains the main body of work for the proof of Theorem 1.1 (the convergence results). The remainder of Theorem 1.1 regarding the OZE is then proved in Section 5.

We discuss our results in Section 6. In particular, we point out where many of the formulas can be found in the physics literature (not rigorously proven) and allude to generalizations to Gibbs point processes. Moreover, we highlight the connection to two other expressions for the pair-connectedness function; in particular, we show how our expansions relate to the lace expansion. Lastly, we address other percolation models very briefly in Section 6.4.

2. Fixing notation

2.1. General notation

We let $[n] := \{1, \ldots, n\}$ and $[n]_0 := [n] \cup \{0\}$. For a set V, we write $\binom{V}{2} := \{E \subseteq V : |E| = 2\}$. For $I = \{i_1, i_2, \ldots, i_\kappa\} \subset \mathbb{N}$, let $\vec{x}_I = \left(x_{i_1}, \ldots, x_{i_\kappa}\right)$. For compact intervals $[a, b] \subset \mathbb{R}$, we write

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

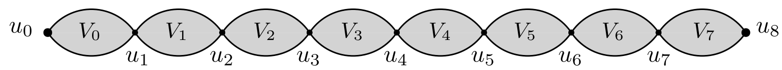

FIGURE 1. A schematic sketch of the pivot decomposition $(u_0, V_0, \dots, V_7, u_8)$ of G, setting $x = u_0$ and $y = u_{k+1}$.

$\vec{x}_{[a,b]} = \vec{x}_I$ with $I = [a, b] \cap \mathbb{N}$. If $a = 1$, we write $\vec{x}_{[b]} = \vec{x}_{[1,b]}$. By some abuse of notation, we are going to interpret $\vec{x}_{[a,b]}$ both as an ordered vector and as a set.

If not specified otherwise, Λ denotes a bounded, measurable subset of $\mathbb{R}^d$.

2.2. Graph theory

We recall that a (simple) graph $G = (V, E) = (V(G), E(G))$ is a tuple with *vertex* set (or set of *points, sites, nodes*) V and *edge* set (or set of *bonds*) $E \subseteq \binom{V}{2}$. In this paper, we will always consider graphs with $V \subset \mathbb{R}^d$, and for $x, y \in \mathbb{R}^d$, an edge $\{x, y\}$ will sometimes be abbreviated xy.

If $xy \in E$, we write $x \sim y$ (and say that x and y are *adjacent*). We extend this notation and write $x \sim W$ for $x \in V$ and $W \subseteq V$ if there is $y \in W$ such that $x \sim y$; also, we write $A \sim B$ if there is $x \in A$ such that $x \sim B$. For $W \subseteq V$, we define the *W-neighborhood* $N_W(x) = \{y \in W : x \sim y\}$ and the *W-degree* of a vertex $x \in V$ as $\deg_W(x) = |N_W(x)|$, and we write $N(x) = N_V(x)$ as well as $\deg(x) = \deg_V(x)$. For two sets $A, B \subseteq V$, we write $E(A, B) = \{xy \in E(G) : x \in A, y \in B\}$.

Given a graph $G = (V, E)$ and $W \subseteq V$, we denote by $G[W] := (W, \{e \in E : e \subseteq W\})$ the subgraph of G *induced* by W. Given two simple graphs G, H, we let $G \oplus H := (V(G) \cup V(H), E(G) \cup E(H))$.

Connectivity. Given a graph G and two of its vertices $x, y \in V(G)$, we say that x and y are *connected* if there is a *path* between x and y—that is, a sequence of vertices $x = v_0, v_1, \dots, v_k = y$ for some $k \in \mathbb{N}_0$ such that $v_{i-1}v_i \in E(G)$ for $i \in [k]$. We write $x \longleftrightarrow y$ in G or simply $x \longleftrightarrow y$. We call $\mathscr{C}(x) = \mathscr{C}(x; G) = \{y \in V(G) : x \longleftrightarrow y\}$ the *cluster* (or *connected component)* of x in G. If there is only one cluster in G, we say that G is connected.

For $x \longleftrightarrow y$ in G, we let $\mathsf{Piv}(x, y; G)$ denote the set of *pivotal* vertices for the connection between x and y. That is, $v \notin \{x, y\}$ is in $\mathsf{Piv}(x, y; G)$ if every path from x to y in G passes through v. We say that x is doubly connected to y in G (and write $x \Longleftrightarrow y$ in G) if $\mathsf{Piv}(x, y; G) = \varnothing$. We remark that in the physics literature, pivotal points are usually known as *nodal* points.

In the pathological case $x = y$, we use the convention $x \longleftrightarrow x$ in G and set $\mathsf{Piv}(x, x; G) = \varnothing$ for any graph G with $x \in V(G)$ (equivalently, $x \Longleftrightarrow x$ in G).

We observe that the pivotal points $\{u_1, \dots, u_k\}$ can be ordered in a way such that every path from x to y passes through the pivotal points in the order $(u_1, \dots, u_k)$. We define $\mathsf{PD}(x, y, G) = \mathsf{PD}(G)$ to be the *pivot decomposition* of G, that is, a partition of the vertex set V into a sequence, $(x, V_0, u_1, V_1, \dots, u_k, V_k, y)$, where $(u_1, \dots, u_k)$ are the ordered pivotal points and V_i is the (possibly empty) set of vertices that can be reached only by passing through u_i and that is still connected to x after the removal of u_{i+1}. See Figure 1.

Classes of graphs. Given a (locally finite) set $X \subset \mathbb{R}^d$, we let $\mathcal{G}(X)$ be the set of graphs with vertex set X. We let $\mathcal{C}(X)$ be the set of connected graphs on X. Moreover, for $x, y \in X$, we let $\mathcal{D}_{x,y}(X) \subseteq \mathcal{C}(X)$ be the set of non-pivotal graphs, i.e., the set of connected graphs such that $\mathsf{Piv}(x, y; G) = \varnothing$.

Given m *bags* $X_1, \dots, X_m \subset \mathbb{R}^d$ with $|X_i \cap X_j| \le 1$ for all $1 \le i < j \le m$, we let $\mathcal{G}(X_1, \dots, X_m)$ denote the set of *m-partite* graphs on $X_1, \dots, X_m$, i.e., the set of graphs G

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

with $V(G) = \cup_{i=1}^{m} X_i$ and $E(G[X_i]) = \varnothing$ for $i \in [m]$. Note that we allow bags to have (at most) one vertex in common, which is a slight abuse of the notation in graph theory, where m-partite graphs have disjoint bags.

The notion of (±)-graphs. We introduce a (±)-graph as a triple

$$G^{\pm} = (V(G), E^{+}(G), E^{-}(G)) = (V, E^{+}, E^{-}),$$

where V is the vertex set and $E^{+}, E^{-} \subseteq \binom{V}{2}$ are disjoint. In other words $G^{\pm}$ is a graph where every edge is of exactly one of two types (plus or minus). We set $E := E^{+} \cup E^{-}$ and associate to $G^{\pm}$ the two simple graphs $G^{|\pm|} := (V, E)$ and $G^{+} := (V^{+}, E^{+})$, where $V^{+} := \{x \in V : \exists e \in E^{+} : x \in e\}$ are the vertices incident to at least one (+)-edge.

We extend all the notions for simple graphs to (±)-graphs. In particular, given $X \subset \mathbb{R}^d$, we let $\mathcal{G}^{\pm}(X)$ be the set of (±)-graphs on X. Moreover, $\mathcal{C}^{\pm}(X)$ are the (±)-connected graphs on X, that is, the graphs such that $G^{|\pm|}$ is connected. Similarly, $\mathcal{C}^{+}(X) \subset \mathcal{C}^{\pm}(X)$ are the (+)-connected graphs, that is, those where G^{+} is connected and $V(G) = V^{+}$. For $x, y \in X$, we denote by $\mathcal{D}^{\pm}_{x,y}(X)$ the set of those (±)-connected graphs on X where $\mathsf{Piv}(x, y; G^{|\pm|}) = \varnothing$, and by $\mathcal{D}^{+}_{x,y}(X) \subset \mathcal{D}^{\pm}_{x,y}(X)$ the set of those (±)-connected graphs on X where $\mathsf{Piv}\big(x, y; G^{+}\big) = \varnothing$. We also define the (±)-pivot decomposition $\mathsf{PD}^{\pm}\big(x, y, G^{\pm}\big) = \mathsf{PD}^{\pm}\big(G^{\pm}\big) = \mathsf{PD}(G^{|\pm|})$ and the (+)-pivot decomposition $\mathsf{PD}^{+}\big(x, y, G^{\pm}\big) = \mathsf{PD}^{+}\big(G^{\pm}\big) = \mathsf{PD}\big(G^{+}\big)$. Lastly, we write $x \overset{+}{\longleftrightarrow} y$ if there is a path from x to y in E^{+}.

Given a (±)-graph G and a simple graph H, we define

$$G \oplus H := (V(G) \cup V(H), E^{+}(G), E^{-}(G) \cup E(H)).$$

Weights. Given a simple graph G, a (±)-graph H on $X \subset \mathbb{R}^d$, and the connection function φ, we define the weights

$$\mathbf{w}(G) := (-1)^{|E(G)|} \prod_{\{x,y\} \in E(G)} \varphi(x-y), \qquad \mathbf{w}^{\pm}(H) := (-1)^{|E^{-}(H)|} \prod_{\{x,y\} \in E(H)} \varphi(x-y).$$

2.3. The random connection model

The RCM ξ can be formally constructed as a point process, that is, a random variable taking values in the space of locally finite counting measures $(\mathbf{N}, \mathcal{N})$ on some underlying metric space $\mathbb{X}$. There are various ways to choose $\mathbb{X}$. One option is to let $\mathbb{X} = \mathbb{R}^d \times \mathbb{M}$ for an appropriate mark space $\mathbb{M}$ (see [18]); another way can be found in [10, 15]. In any case, one can reconstruct from ξ the point process η on $\mathbb{R}^d$ which makes up the vertex set of ξ. We treat η both as a counting measure and as a set, giving meaning to statements of the form $x \in \eta$.

If $e = \{x, y\}$ is an edge, then we write $\varphi(e) = \varphi(x-y)$. For a bounded set $\Lambda \subset \mathbb{R}^d$, we write $\eta_\Lambda = \eta \cap \Lambda$ and let ξ_Λ denote the RCM restricted to Λ, that is, $\xi[\eta_\Lambda]$. The two-point function restricted to Λ is defined as $\tau^{\Lambda}_{\lambda}(x, y) = \mathbb{P}_\lambda\big(x \longleftrightarrow y \text{ in } \xi^{x,y}_{\Lambda}\big)$ for $x, y \in \Lambda$ and zero otherwise.

For $V \subset W$, there is a natural way to couple the models ξ^{V} and ξ^{W}, which is by deleting from ξ^{W} all points in $W \setminus V$ along with their incident edges. We implicitly assume throughout this paper that this coupling for different sets of added points is used.

The Mecke equation. Since it is used repeatedly throughout this paper, we state the Mecke equation, a standard tool in point process theory, in its version for the RCM (see

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

[15]). For $m \in \mathbb{N}$ and a measurable function $f : \mathbf{N} \times \mathbb{R}^{dm} \to \mathbb{R}_{\geq 0}$, the Mecke equation states that

$$\mathbb{E}_\lambda\left[\sum_{\vec{x}_{[m]} \in \eta^{(m)}} f\big(\xi, \vec{x}_{[m]}\big)\right] = \lambda^m \int \mathbb{E}_\lambda\big[f\big(\xi^{x_1,\dots,x_m}, \vec{x}_{[m]}\big)\big]\, \mathrm{d}\vec{x}_{[m]}, \tag{2.1}$$

where $\eta^{(m)} = \big\{\vec{x}_{[m]} \in \eta^m : x_i \neq x_j \text{ for } i \neq j\big\}$ are the pairwise distinct tuples.

Rescaling. It is a standard trick in continuum percolation to rescale space in order to normalize a quantity of interest, which is $\int \varphi(x)\, \mathrm{d}x$ in our case. We refer to [18, Section 2.2]. As a consequence, we may without loss of generality assume that $\int \varphi(x)\, \mathrm{d}x = 1$.

The BK inequality. We say that $A \in \mathcal{N}$ *lives on* Λ if $\mathbb{1}_A(\mu) = \mathbb{1}_A(\mu_\Lambda)$ for every $\mu \in \mathbf{N}$. We call an event $A \in \mathcal{N}$ *increasing* if $\mu \in A$ implies $\nu \in A$ for each $\nu \in \mathbf{N}$ with $\mu \subseteq \nu$. Let $\mathcal{R}$ denote the ring of all finite unions of half-open rectangles with rational coordinates. For two increasing events $A, B \in \mathcal{N}$ we define

$$A \circ B := \{\mu \in \mathbf{N} : \exists K, L \in \mathcal{R} \text{ s.t. } K \cap L = \varnothing \text{ and } \mu_K \in A,\ \ \mu_L \in B\}. \tag{2.2}$$

Informally, this is the event that A and B take place in spatially disjoint regions. It is proved in [10, Theorem 2.1] that for two increasing events A and B living on Λ, we have

$$\mathbb{P}_\lambda(A \circ B) \leq \mathbb{P}_\lambda(A)\mathbb{P}_\lambda(B).$$

The RCM on a fixed vertex set. Given some (finite) set $X \subset \mathbb{R}^d$ and a function $\varphi : \mathbb{R}^d \to [0, 1]$, we will often have to deal with the following random graph: its vertex set is X, and two vertices $x, y \in X$ are adjacent with probability $\varphi(x - y)$, independently of other pairs of vertices. This is simply the RCM conditioned to have the vertex set X. To highlight the difference from ξ, which depends on the PPP η, we denote this random graph by $\Gamma_\varphi(X)$. If $Y \subset X$, then we write $\Gamma_\varphi(Y)$ for $\Gamma_\varphi(X)[Y]$. Since there is no dependence on λ, we write $\mathbb{P}$ for the probability measure of the RCM with fixed vertex set.

3. Fixing ideas: the two-point function in finite volume

We use this section to put the definitions of Section 2 into action and to derive a power series expansion for τ_λ in finite volume. We start by motivating the introduction of $(\pm)$-graphs by linking them to the RCM Γ_φ.

Observation 3.1. (*Connection between* $(\pm)$*-graphs and probabilities.*) Let $X \subset \mathbb{R}^d$ be finite. Let $\mathfrak{P} \subseteq \mathcal{G}(X)$ be a graph property. Then

$$\sum_{\substack{G \in \mathcal{G}^{\pm}(X): \\ (V(G), E^+(G)) \in \mathfrak{P}}} \mathbf{w}^{\pm}(G) = \mathbb{P}\big(\Gamma_\varphi(X) \in \mathfrak{P}\big).$$

Proof. Note that

$$\mathbb{P}\big(\Gamma_\varphi(X) \in \mathfrak{P}\big) = \sum_{\substack{G \in \mathcal{G}(X): \\ G \in \mathfrak{P}}} \prod_{e \in E(G)} \varphi(e) \prod_{e \in \binom{X}{2} \setminus E(G)} (1 - \varphi(e)).$$

Expanding the factor $\prod_{e \in \binom{X}{2} \setminus E(G)} (1 - \varphi(e))$ into a sum proves the claim. □

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Note that the weight of a $(\pm)$-graph may also be calculated by taking the product over all its edges, with factors $\varphi(\cdot)$ and $-\varphi(\cdot)$ for edges in E^+ and E^-, respectively. Observation 3.1 motivates that the edges in E^+ correspond to the edges in the random graph Γ_φ.

Next we prove a power series expansion for τ_λ in terms of the intensity λ. The expansion (3.1) has already been given by Coniglio, De Angelis and Forlani [6, Equation (12)], who work in the more general context of Gibbs point processes but do not prove convergence. The proposition enters the proof of Proposition 5.1.

Notice that the coefficients of power series expansions like (3.1) are given by integrals with respect to the Lebesgue measure, and it is sufficient that the integrands be defined up to Lebesgue null sets for those integrals to be well-defined. Since vectors $\vec{x}_{[3,n+2]} \in \mathbb{R}^{dn}$ with fewer than n distinct entries constitute a Lebesgue null set, we can assume that for $x_1 \neq x_2$ only graphs with vertex sets of cardinality $n+2$ contribute to the nth coefficient in (3.1). The same considerations apply to all graphical expansions appearing from here on, including our main definition (4.6).

Proposition 3.1. (Graphical expansion for the two-point function.) *Consider the RCM restricted to a bounded measurable set $\Lambda \subset \mathbb{R}^d$, and let $x_1, x_2 \in \Lambda$. Then*

$$\tau_\lambda^\Lambda(x_1, x_2) = \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{G \in \mathcal{C}^\pm\left(\vec{x}_{[n+2]}\right): \\ x_1 \xleftrightarrow{+} x_2}} \mathbf{w}^\pm(G) \, \mathrm{d}\vec{x}_{[3,n+2]} \tag{3.1}$$

with

$$\sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \Bigg| \sum_{\substack{G \in \mathcal{C}^\pm\left(\vec{x}_{[n+2]}\right): \\ x_1 \xleftrightarrow{+} x_2}} \mathbf{w}^\pm(G) \Bigg| \, \mathrm{d}\vec{x}_{[3,n+2]} \leq \exp\left\{2\lambda + \lambda|\Lambda|\mathrm{e}^\lambda\right\} < \infty.$$

Note that Proposition 3.1 is valid for all intensities $\lambda \geq 0$. This situation is completely different from familiar cluster expansions [2], where the radius of convergence of relevant expansions is finite in finite volume as well.

The expansion (3.1) amounts to the physicists' expansion in powers of the activity. The expansion in powers of the density instead involves sums over a smaller class of graphs. For PPPs, activity and density are the same and the two expansions must coincide. In our context, we point out that the sum over graphs in (3.1) can be reduced to the sum over the subset of graphs in $\mathcal{C}^\pm$ that contain a $(+)$-path from x_1 to x_2 and that have no *articulation points* (with respect to x_1, x_2). To define articulation points, recall that a cut vertex leaves a connected graph disconnected upon its deletion. Now, an articulation point is a cut vertex that is not pivotal for the x_1–x_2 connection. It is not difficult to see that for fixed points $x_{[n+2]}$, the graphs with articulation points in the sum over graphs G in (3.1) exactly cancel out. This cancellation happens at fixed n and does not require any re-summations between graphs with different numbers of vertices.

The proof of Proposition 3.1 builds on yet another equivalent representation: in Equation (3.1) we can discard those graphs G for which G^+ is not connected and those for which not every $(-)$-edge has at least one endpoint in $V\big(G^+\big)$; see Equation (3.6) below for a precise statement. To the best of our knowledge, Equation (3.6) is new.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Proof of Proposition 3.1. We write $\tau_\lambda = \tau_\lambda^\Lambda$ and $\eta = \eta_\Lambda$. Given $x_1, x_2 \in \Lambda$, we can partition

$$\begin{aligned}\tau_\lambda(x_1, x_2) &= \sum_{n\geq 0} \mathbb{P}_\lambda\Big(x_1 \longleftrightarrow x_2 \text{ in } \xi_\Lambda^{x_1,x_2}, \big|\mathscr{C}\big(x_1, \xi_\Lambda^{x_1,x_2}\big)\big| = n+2\Big)\\ &= \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \mathbb{P}\big(\Gamma_\varphi\big(\vec{x}_{[n+2]}\big) \in \mathcal{C}\big(\vec{x}_{[n+2]}\big)\big)\\ &\qquad \times \exp\left\{ -\lambda \int_\Lambda \left(1 - \prod_{i=1}^{n+2} (1 - \varphi(x_i - y))\right) \mathrm{d}y\right\} \mathrm{d}\vec{x}_{[3,n+2]} \end{aligned} \tag{3.2}$$

The second identity can be found, for example, in [15, Proposition 3.1]. Set

$$f\big(\vec{x}_{[n+2]}, \vec{y}_{[m]}\big) = \mathbb{P}\big(\Gamma_\varphi\big(\vec{x}_{[n+2]}\big) \in \mathcal{C}\big(\vec{x}_{[n+2]}\big)\big) \prod_{j=1}^m \left(\prod_{i=1}^{n+2} (1 - \varphi(x_i - y_j)) - 1\right).$$

Expanding the exponential in (3.2), we find

$$\tau_\lambda(x_1, x_2) = \sum_{n,m\geq 0} \frac{\lambda^{n+m}}{m!n!} \int_{\Lambda^n} \int_{\Lambda^m} f\big(\vec{x}_{[n+2]}, \vec{y}_{[m]}\big)\, \mathrm{d}\vec{y}_{[m]}\, \mathrm{d}\vec{x}_{[3,n+2]}, \tag{3.3}$$

with

$$\begin{aligned}&\sum_{n,m\geq 0} \frac{\lambda^{n+m}}{m!n!} \int_{\Lambda^n} \int_{\Lambda^m} \big|f\big(\vec{x}_{[n+2]}, \vec{y}_{[m]}\big)\big|\, \mathrm{d}\vec{y}_{[m]}\, \mathrm{d}\vec{x}_{[3,n+2]}\\ &= \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \mathbb{P}\big(\Gamma_\varphi\big(\vec{x}_{[n+2]}\big) \in \mathcal{C}\big(\vec{x}_{[n+2]}\big)\big) \exp\left\{\lambda \int_\Lambda \left(1 - \prod_{i=1}^{n+2} (1 - \varphi(x_i - y))\right) \mathrm{d}y\right\} \mathrm{d}\vec{x}_{[3,n+2]}\\ &\leq \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \mathrm{e}^{\lambda(n+2)}\, \mathrm{d}\vec{x}_{[3,n+2]}\\ &= \exp\big\{2\lambda + \lambda|\Lambda|\mathrm{e}^\lambda\big\} < \infty. \end{aligned} \tag{3.4}$$

In the third line, we have used the inequality

$$\int_\Lambda \left(1 - \prod_{i=1}^{n+2} (1 - \varphi(x_i - y))\right) \mathrm{d}y \leq \int_\Lambda \sum_{i=1}^{n+2} \varphi(x_i - y)\, \mathrm{d}y \leq n+2, \tag{3.5}$$

which can be shown as follows. Let $n \in \mathbb{N}$ and let $0 \leq a_1, \ldots, a_n \leq 1$. Notice that the identity $1 - \prod_{i=1}^n (1 - a_i) = (1 - a_n)\big(1 - \prod_{i=1}^{n-1} (1 - a_i)\big) + a_n$ and the estimate $(1 - a_n) \leq 1$ hold for all $n \in \mathbb{N}$. The inequality between the integrands in (3.5) now follows by induction with the choice $a_i = \varphi(x_i - y)$. The rescaling introduced in Section 2.3 ensures that $\int_\Lambda \varphi(x_i - y)\, \mathrm{d}y \leq 1$, $i \in [n+2]$, yielding the second inequality.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Next we turn to a combinatorial representation of f as a sum over $(\pm)$-graphs. Recall that $\mathcal{C}^+$ denotes sets of $(\pm)$-graphs that are $(+)$-connected. The definition of f and Observation 3.1 yield

$$f\big(\vec{x}_{[n+2]}, \vec{y}_{[m]}\big) = \left(\sum_{G\in\mathcal{C}^+\left(\vec{x}_{[n+2]}\right)} \mathbf{w}^{\pm}(G)\right)\left(\sum_{\substack{H\in\mathcal{G}\left(\vec{x}_{[n+2]},\vec{y}_{[m]}\right):\\ y_i\sim\vec{x}_{[n+2]}\ \forall i\in[m]}} \mathbf{w}(H)\right) = \sum_{G\oplus H} \mathbf{w}^{\pm}(G\oplus H),$$

where the last sum is over all $(\pm)$-graphs $G' = G\oplus H$ in $\mathcal{C}^{\pm}\big(\vec{x}_{[n+2]}\cup\vec{y}_{[m]}\big)$ such that, first, there are no edges between points of $\vec{y}$; second, $(G\oplus H)^+$ is connected; and third, the vertices of $(G\oplus H)^+$ are precisely $\vec{x}_{[n+2]}$.

We rearrange the double sum (3.3) over m, n into one sum, indexed by the value of $m+n$, and obtain

$$\tau_\lambda(x_1, x_2) = \sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^n} \sum_{\substack{G\in\mathcal{C}^{\pm}\left(\vec{x}_{[n+2]}\right):\\ \{x_1,x_2\}\subseteq V\left(G^+\right),\,G^+\text{ connected},\\ E(G^{|\pm|}[V\setminus V^+])=\varnothing}} \mathbf{w}^{\pm}(G)\,\mathrm{d}\vec{x}_{[3,n+2]} \tag{3.6}$$

$$= \sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^n} \sum_{\substack{G\in\mathcal{C}^{\pm}\left(\vec{x}_{[n+2]}\right):\\ x_1\overset{+}{\longleftrightarrow}x_2}} \mathbf{w}^{\pm}(G)\,\mathrm{d}\vec{x}_{[3,n+2]}. \tag{3.7}$$

In the second identity, we have added some graphs to the sum, namely those in which G^+ is not connected or where there exist edges between vertices of $V\setminus V^+$.

We claim that the weights of these added graphs sum up to zero. To see this, first identify $[n+2]$ with the vertices $\vec{x}_{[n+2]}$ and fix a graph $G\in\mathcal{C}([n+2])$. Now, let $C\subseteq[n+2]$ with $\{1, 2\}\subseteq C$ and consider the set $\mathcal{G}_G(C)$ of all $(\pm)$-connected graphs $G^{\pm}$ on $[n+2]$ such that $G^{|\pm|}=G$ and C is the vertex set of the $(+)$-component of 1 in $G^{\pm}$. If there is at least one edge e in G that has both endpoints outside of C, we partition $\mathcal{G}_G(C)$ into those graphs where e is in E^+ and those where e is in E^-. This induces a pairing between the graphs of $\mathcal{G}_G(C)$, and they cancel out. What remain are precisely the graphs in (3.6). □

4. The direct-connectedness function

4.1. Motivation and rough outline

The expansion of the direct-connectedness function in powers of the activity given by [6], without proofs and convergence bounds, is

$$g_\lambda^{\Lambda}(x_1, x_2) = \sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^n} \sum_{\substack{G\in\mathcal{D}^{\pm}_{x_1,x_2}\left(\vec{x}_{[n+2]}\right):\\ x_1\overset{+}{\longleftrightarrow}x_2}} \mathbf{w}^{\pm}(G)\,\mathrm{d}\vec{x}_{[3,n+2]}. \tag{4.1}$$

It is obtained from the expansion of the pair-connectedness function in Proposition 3.1 by discarding graphs that have pivotal points (i.e., graphs G where $\mathsf{Piv}^{\pm}(G)$ is nonempty). Before we

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

pass to the thermodynamic limit, we perform a re-summation and find another representation of g_λ^Λ which has the conjectured advantage of increasing the domain of convergence.

Let $G=(V,E^+,E^-)\in\mathcal{C}^\pm(\vec{x}_{[n+2]})$ be a $(\pm)$-graph appearing in the expansion (3.6). Thus $V=\{x_i : 1\le i\le n+2\}$, the graph G^+ is connected, x_1 and x_2 belong to $V^+=V(G^+)$, every vertex $y\in V(G)\setminus V(G^+)$ is linked by at least one $(-)$-edge to V^+, and there are no edges between two vertices in $V\setminus V^+$. We impose the additional constraint that $G^{|\pm|}=(\vec{x}_{[n+2]}, E^+\cup E^-)$ has no pivotal points for paths from x_1 to x_2.

Since x_1 and x_2 are connected by a path of $(+)$-edges, G admits a $(+)$-pivot decomposition $\vec{W}=(u_0,V_0,\dots,u_k,V_k,u_{k+1})$ (with $u_0=x_1$ and $u_{k+1}=x_2$), where $k\in\mathbb{N}_0$ is the number of pivotal points in $\mathsf{Piv}^+(x_1,x_2;G)$. Then, G decomposes into a *core* graph $G_{\mathrm{core}}=(V(G^+),E^+,E^-_{\mathrm{core}})$, with E^-_{core} the set of $(-)$-edges of G with both endpoints in $V_i\cup\{u_i,u_{i+1}\}$ for some $i\in[k]_0$, and a *shell* graph $H=(V,\varnothing,E^-\setminus E^-_{\mathrm{core}})$. By our choice of E^-_{core}, we have $\mathsf{PD}^\pm(G_{\mathrm{core}})=\mathsf{PD}^+(G_{\mathrm{core}})=\vec{W}$. Clearly

$$\mathbf{w}^\pm(G)=\mathbf{w}^\pm(G_{\mathrm{core}})\mathbf{w}^\pm(H).$$

In the right-hand side of (4.1), we restrict to graphs that also appear in (3.6) and rewrite the resulting sum as a double sum over core graphs and shell graphs. This gives rise to the series

$$\sum_{r=0}^\infty\frac{\lambda^r}{r!}\int_{\Lambda^r}\sum_{\vec{W}}\sum_{G_{\mathrm{core}}}\mathbf{w}^\pm(G_{\mathrm{core}})\left(\sum_{m=0}^\infty\frac{\lambda^m}{m!}\int_{\Lambda^m}\sum_H\mathbf{w}^\pm(H)\,\mathrm{d}\vec{y}_{[m]}\right)\mathrm{d}\vec{x}_{[3,r+2]}.$$

The outer sum is over potential pivot decompositions $\vec{W}$ of core vertices $\vec{x}_{[r+2]}$, the second sum over $(\pm)$-graphs $G_{\mathrm{core}}=(\vec{x}_{[r+2]},E^+,E^-_{\mathrm{core}})$ that are $(+)$-connected and for which $\vec{W}$ is both the $(\pm)$-pivot decomposition and the $(+)$-pivot decomposition (in other words, the simple graph $(\vec{x}_{[r+2]},E^+)$ is connected and $\mathsf{PD}^\pm(x_1,x_2,G)=\mathsf{PD}^+(x_1,x_2,G)=\vec{W}$). The inner sum is over $(\pm)$-graphs $H=(V(H),\varnothing,E^-(H))$ with vertex set $\vec{x}_{[r+2]}\cup\vec{y}_{[m]}$ and $(-)$-edges $\{y_i,x_j\}$ such that every vertex y_i is linked to at least one vertex x_j, under the additional constraint that $(\vec{x}_{[r+2]}\cup\vec{y}_{[m]},E^+,E^-_{\mathrm{core}}\cup E^-(H))$ has no $(\pm)$-pivotal points for paths from x_1 to x_2. Let us denote the series associated to such graphs H by $h_\lambda^\Lambda(G_{\mathrm{core}})$:

$$h_\lambda^\Lambda(G_{\mathrm{core}})=\sum_{m=0}^\infty\frac{\lambda^m}{m!}\int_{\Lambda^m}\sum_H\mathbf{w}^\pm(H)\,\mathrm{d}\vec{y}_{[m]}. \tag{4.2}$$

The right-hand side of (4.2) depends on G_{core} only through the pivot decomposition $\vec{W}$. We obtain the representation

$$g_\lambda^\Lambda(x_1,x_2)=\sum_{r=1}^\infty\frac{\lambda^r}{r!}\int_{\Lambda^r}\sum_{\vec{W}}\sum_{G_{\mathrm{core}}}\mathbf{w}^\pm(G_{\mathrm{core}})h_\lambda^\Lambda(G_{\mathrm{core}})\,\mathrm{d}\vec{x}_{[3,r+3]}. \tag{4.3}$$

This expression, written in a slightly different form (see Definition 4.2), forms the starting point of this section. The main results of this section are the following:

1. Let G_{core} be a $(\pm)$-graph as above. Then the corresponding power series $h_\lambda^\Lambda(G_{\mathrm{core}})$ is absolutely convergent for all intensities $\lambda\ge0$ (Proposition 4.1). In addition, $h_\lambda^\Lambda(G_{\mathrm{core}})$ can be expressed in terms of probabilities involving the random connection model on

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

the fixed vertex set $V\big(G_{\mathrm{core}}\big)$ and of Poisson processes in Λ. This alternative expression is used to show that the (pointwise) limit

$$h_\lambda\big(G_{\mathrm{core}}\big) = \lim_{\Lambda \nearrow \mathbb{R}^d} h_\lambda^\Lambda\big(G_{\mathrm{core}}\big)$$

exists for all $\lambda > 0$ (Lemma 4.5).

2. Then we show in Theorem 4.1 that

$$\sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{(\mathbb{R}^d)^r} \sum_{\vec{W}} \sum_{G_{\mathrm{core}}} \mathbf{w}^{\pm}\big(G_{\mathrm{core}}\big) \Big| h_\lambda\big(G_{\mathrm{core}}\big)\Big| \,\mathrm{d}\vec{x}_{[3,r+3]} < \infty.$$

This allows us to define

$$g_\lambda(x_1, x_2) := \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{(\mathbb{R}^d)^r} \sum_{G_{\mathrm{core}}} \mathbf{w}^{\pm}\big(G_{\mathrm{core}}\big) h_\lambda\big(G_{\mathrm{core}}\big) \,\mathrm{d}\vec{x}_{[3,r+3]}$$

and to pass to the limit in (4.3), showing that

$$\lim_{\Lambda \nearrow \mathbb{R}^d} g_\lambda^\Lambda(x_1, x_2) = g_\lambda(x_1, x_2)$$

as part of Theorem 4.1.

4.2. Definition

Here we introduce the precise definitions of core graphs and shell graphs as well as of the functions h_λ^Λ and g_λ^Λ. We follow the ideas outlined in the previous section but make two small changes. First, shell graphs H are defined not as $(\pm)$-graphs with minus edges only but right away as standard graphs. Second, a close look reveals that the shell function $h_\lambda^\Lambda\big(G_{\mathrm{core}}\big)$ defined in (4.2) depends on the core graph only via $\vec{W}$; accordingly we view h_λ^Λ as a function of a sequence of sets. In addition we drop the index from the core graph; thus the graph G in Definition 4.1 below corresponds to G_{core} in the previous section (see Figure 2).

Definition 4.1. (*Core graphs and shell graphs.*)

1. Let $x_1, x_2 \in \mathbb{R}^d$ and let $\{x_1, x_2\} \subset W \subset \mathbb{R}^d$ be a finite set of vertices. We call a graph $G \in \mathcal{C}^+(W)$ with $\mathsf{PD}^{\pm}(x_1, x_2, G) = \mathsf{PD}^{+}(x_1, x_2, G) = \vec{W}$ a *core* graph with pivot decomposition $\vec{W}$ and denote the set of such graphs by $\mathcal{G}_{\mathrm{core}}^{\vec{W}}$.
2. Let $G \in \mathcal{C}^+(W)$ be a core graph with pivot decomposition $\vec{W} = (u_0, V_0, \dots, V_k, u_{k+1})$, $k \in \mathbb{N}_0$, where we set $u_0 := x_1$ and $u_{k+1} := x_2$. Moreover, let $\overline{V}_i := V_i \cup \{u_i, u_{i+1}\}$ and let Y be a finite subset of $\mathbb{R}^d$. A *shell* graph on $W \cup Y$ associated to $\vec{W}$ is a $(k+1)$-partite graph $H \in \mathcal{G}\big(\overline{V}_1, \dots, \overline{V}_k, Y\big)$ such that $G \oplus H \in \mathcal{D}^{\pm}_{x_1,x_2}(W \cup Y)$. We call the vertices $Y \subset V(H)$ satellite vertices and write $S(H) = Y$. Notice that the set of all shell graphs on $W \cup Y$ associated to $\vec{W}$ does not depend on the choice of the core graph G. We denote it by $\mathcal{G}_{\mathrm{shell}}^{Y,\vec{W}}$.

We define h_λ^Λ and g_λ^Λ by expansions similar to (4.2) and (4.3) and postpone the proof of convergence to Proposition 4.3 and Theorem 4.4. By some abuse of language, we refer to the series (4.6) as the direct-connectedness function, and we use the same letter g_λ as in (1.2). This

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

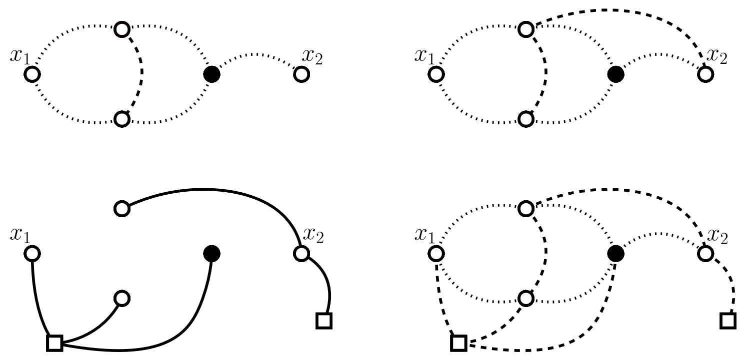

FIGURE 2. In the first line, we see an example of two $(\pm)$-graphs; the $(+)$-edges are depicted by dotted lines and the (minus;)-edges by dashed lines. Notice that both graphs are $(+)$-connected. However, the solid black vertex—which is $(+)$-pivotal for the x_1–x_2 connection in both graphs—is $(\pm)$-pivotal for the x_1–x_2 connection in the graph on the left but not in the graph on the right. Hence, the graph on the left is a core graph according to Definition 4.2 but the graph on the right is not. In the second line, the simple graph on the left is a shell graph for the core graph above, since the $(\pm)$-graph given by their sum (depicted on the right) is $(\pm)$-doubly connected; in particular there are no $(\pm)$-pivotal points for the x_1–x_2 connection.

is justified a posteriori by the proof of Theorem 1.1, where we show that the series is indeed the expansion for the direct-connectedness function g_λ defined as the unique solution of the OZE (1.2).

Definition 4.2. (*Shell functions and direct-connectedness function.*)

1. Let $W \subset \mathbb{R}^d$ be finite and let $\vec{W}$ be given as in Definition 4.2. For $m \in \mathbb{N}_0$, define the *m-shell function* $h^{(m)}$ by

$$h^{(m)}(\vec{W}, Y) := \sum_{H \in \mathcal{G}_{\text{shell}}^{Y,\vec{W}}} \mathbf{w}(H), \qquad Y = \{y_1, \ldots, y_m\} \subset \mathbb{R}^d, \tag{4.4}$$

and the *shell function* h_λ^Λ in finite volume $\Lambda \subset \mathbb{R}^d$ by

$$h_\lambda^\Lambda\big(\vec{W}\big) := \sum_{m \geq 0} \frac{\lambda^m}{m!} \int_{\Lambda^m} h^{(m)}\big(\vec{W}, \vec{y}_{[m]}\big)\, \mathrm{d}\vec{y}_{[m]}. \tag{4.5}$$

2. Let $\lambda < \lambda_*$. We define the *direct-connectedness function* as $g_\lambda : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$,

$$g_\lambda^\Lambda(x_1, x_2) := \sum_{r \geq 0} \frac{\lambda^r}{r!} \int_{\Lambda^r} \sum_{\vec{W}} \left(\sum_{G \in \mathcal{G}_{\text{core}}^{\vec{W}}} \mathbf{w}^{\pm}(G) \right) h_\lambda^\Lambda\big(\vec{W}\big)\, \mathrm{d}\vec{x}_{[3,r+2]}, \tag{4.6}$$

where $W := \{x_1, \ldots, x_{r+2}\}$ and we sum over decompositions $\vec{W}$ of W given as in Definition 4.1. In the pathological case $x_1 = x_2$, (4.6) is to be read as $g_\lambda^\Lambda(x_1, x_2) := 1$. Let $g_\lambda^\Lambda : \mathbb{R}^d \to \mathbb{R}$ be defined by $g_\lambda^\Lambda(x) = g_\lambda^\Lambda(\mathbf{0}, x)$.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

The 0-shell function $h^{(0)}$ is understood to be given in terms of shell graphs without satellite vertices, i.e.,

$$h^{(0)}\big(\vec{W}\big) = \sum_{H \in \mathcal{G}_{\text{shell}}^{\varnothing, \vec{W}}} \mathbf{w}(H).$$

Note that because of translation-invariance, $g_\lambda^\Lambda(x_1, x_2) = g_\lambda^\Lambda(\mathbf{0}, x_2 - x_1) = g_\lambda^\Lambda(x_2 - x_1)$.

4.3. Analysis of the shell functions: laces

If we take a look at the graphs that are summed over in the shell function, we note that the associated minimal structures have a form which is very reminiscent of graphs that are known as *laces* and famously appear in the analysis of, for example, self-avoiding walks [3, 22]. They are also the namesake of the lace-expansion technique.

Proposition 4.1 is the central result of this section. It allows us to bound the shell function by the probability that the points in a PPP η are not connected to the core vertices W. Moreover, we introduce laces and partition the shell graphs with respect to them. For every lace, we obtain a precise expression for its contribution to the shell function.

To prove Proposition 4.1, we will need quite a few definitions (see Definitions 4.3, 4.4, and 4.5) and some intermediate results thereon.

Proposition 4.1. (Bounds on the shell functions) *Let $\lambda \geq 0$ and let $\Lambda \subset \mathbb{R}^d$ be bounded. Let $u_0, \dots, u_{k+1} \in \Lambda$ for $k \in \mathbb{N}_0$, let $V_0, \dots, V_k \subset \Lambda$ be finite sets, and set $\vec{W} = (u_0, V_0, \dots, V_k, u_{k+1})$. Then*

$$\big|h_\lambda^\Lambda\big(\vec{W}\big)\big| \leq \mathbb{P}_\lambda\big(\eta_\Lambda \not\longleftrightarrow W \text{ in } \xi^W\big). \tag{4.7}$$

Moreover,

$$\sum_{m \geq 0} \frac{\lambda^m}{m!} \int_{\Lambda^m} \big|h^{(m)}\big(\vec{W}, \vec{y}_{[m]}\big)\big| \, \mathrm{d}\vec{y}_{[m]} \leq \frac{1}{\sqrt{5}} \mathrm{e}^{3\lambda|W|}\big(3 + \sqrt{5}\big)^{|W|}. \tag{4.8}$$

Proposition 4.1 consists of two parts, and it is (4.8) that guarantees the well-definedness of the shell function h_λ^Λ of Definition 4.2.

Proposition 4.1 is easy to prove for $k = 0$, and we mostly focus on $k \geq 1$. Throughout the remainder of this section, we fix a pivot decomposition $\vec{W} = (u_0, V_0, \dots, V_k, u_{k+1})$ and recall that $\overline{V}_i = V_i \cup \{u_i, u_{i+1}\}$.

We now work towards a deeper understanding of the shell graphs H summed over in (4.4).

Definition 4.3. (*Skeletons*) Let $W \subset \mathbb{R}^d$ and let $\vec{W} = (u_0, V_0, \dots, u_{k+1})$ be a pivot decomposition of some core graph on W. Furthermore, let $Y \subset \mathbb{R}^d$ be finite and let H be a shell graph associated to $\vec{W}$ with satellite vertices $S(H) = Y$. Then we define the skeleton $\hat{H}$ of H as the following graph: its vertex set is $V(\hat{H}) = \{0, \dots, k+1\}$. A bond $\alpha\beta$ is in $E(\hat{H})$ if and only if $|\alpha - \beta| \geq 2$ and there exist $s \in \{u_\alpha\} \cup V_\alpha$, $t \in V_{\beta-1} \cup \{u_\beta\}$ such that

- $st \in E(H)$, or
- $sy, yt \in E(H)$ for some $y \in S(H)$.

In the first case we call $\{s, t\}$ a *direct stitch*, and in the second case we call it an *indirect stitch*. We call an edge $\alpha\beta$ in $E(\hat{H})$ a *bond* to distinguish it from the edge of the underlying graph H.

Thus, the graph $\hat{H}$ has no nearest-neighbor bonds, and $\alpha\beta$ with $|\alpha - \beta| \geq 2$ is a bond in $E(\hat{H})$ if and only if $\{u_\alpha\} \cup V_\alpha$ and $V_{\beta-1} \cup \{u_\beta\}$ are connected by a direct or indirect stitch. See

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

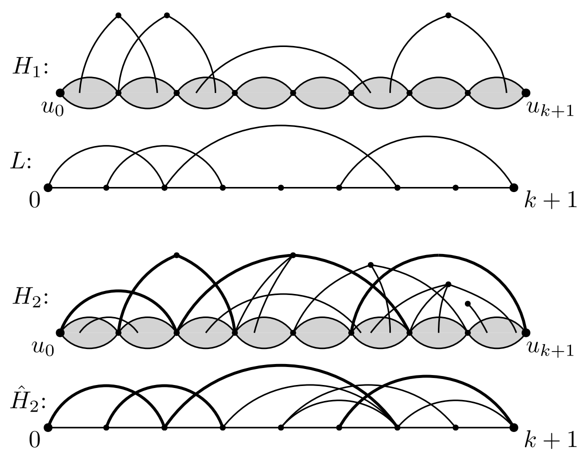

FIGURE 3. In the first line, we see a schematic shell graph H_1. Its skeleton $\hat{H}_1$ is already a lace, namely L. The skeleton of the graph H_2 in the second line is not a lace, but $H_2 \in \langle\langle L \rangle\rangle$. The structure of L is indicated in H_2 and in $\hat{H}_2$ by the thicker edges.

Figure 3 for an illustration. We may now apply the standard vocabulary of lace expansion (for self-avoiding walks) to the graph $\hat{H}$ [22, Section 3.3].

Definition 4.4. (*Laces*)

- The graph $\hat{H}$ with vertex set $\{0, \ldots, k+1\}$ is *irreducible* if 0 and $k+1$ are endpoints of edges in $E(\hat{H})$ and for every $i \in [k]$ there exists $\alpha\beta \in E(\hat{H})$ with $\alpha < i < \beta$.
- The graph $\hat{H}$ is a *lace* if it is irreducible and, for every bond $\alpha\beta \in E(\hat{H})$, removal of the bond destroys the irreducibility.
- We denote by $\mathcal{L}_k$ the set of all laces on $\{0, \ldots, k+1\}$.

In the context of lace expansions, usually the word 'connected' is used instead of 'irreducible', but 'connected' is clearly misleading in our setup; Brydges and Spencer originally called those graphs 'primitive' [3]. We observe that the skeleton graphs $\hat{H}$ arising from our shell graphs H are precisely the irreducible graphs (and so $G \oplus H$ being 2-connected corresponds to the skeleton $\hat{H}$ being irreducible).

We map irreducible graphs to laces by following a standard procedure [22, Section 3.3], performed backwards. That is, we define bonds $\alpha'_j\beta'_j$ with $\beta'_1 > \beta'_2 > \cdots$ inductively as follows: we set

$$\beta'_1 := k+1, \qquad \alpha'_1 := \min\big\{\alpha : \alpha\beta'_1 \in E(\hat{H})\big\},$$

and

$$\alpha'_{j+1} = \min\big\{\alpha : \exists\beta > \alpha'_j \text{ with } \alpha\beta \in E(\hat{H})\big\}, \qquad \beta'_{j+1} = \max\big\{\beta : \alpha'_{j+1}\beta \in E(\hat{H})\big\}.$$

The procedure terminates when $\alpha'_j = 0$. At the end, we let $\alpha_j\beta_j$ be a relabeling of the bonds $\alpha'_j\beta'_j$ from left to right.

It is well known that the algorithm maps irreducible graphs to laces; moreover, the set of irreducible graphs that are mapped to a given lace L can be characterized as follows.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Definition 4.5. (*Compatible bonds and the span of a lace.*)

1. Let L be a lace with vertex set $\{0, \ldots, k+1\}$. A bond is *compatible* with a lace L if the algorithm described above maps the graph $(V(L), E(L) \cup \{\alpha\beta\})$ to the lace L.
2. Let $W \subset \mathbb{R}^d$ and let $\vec{W} = (u_0, V_0, \ldots, u_{k+1})$ be a pivot decomposition of some core graph on W. Further let $Y \subset \mathbb{R}^d$ be finite and let H be a shell graph associated to $\vec{W}$ with $S(H) = Y$. Then we say that H belongs to the span of the lace L, and write $H \in \langle\!\langle L \rangle\!\rangle$, if $E(L) \subseteq E(\hat{H})$ and every bond $\alpha\beta \in E(\hat{H}) \setminus E(L)$ is compatible with L.

In other words, H is in the span of L if the above algorithm maps $\hat{H}$ to L. See Figure 3. Given $\vec{W}$ and a lace L, we define

$$h_\lambda^\Lambda(\vec{W}; L) := \sum_{m \geq 0} \frac{\lambda^m}{m!} \int_{\Lambda^m} \sum_{H \in \langle\!\langle L \rangle\!\rangle \,:\, \mathcal{S}(H) = \vec{y}_{[m]}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[m]}. \tag{4.9}$$

The series $h_\lambda^\Lambda\big(\vec{W}; L\big)$ converges absolutely for every fixed λ. This is shown as part of the proof of (4.8) in Proposition 4.1. Now,

$$h_\lambda^\Lambda\big(\vec{W}\big) = \sum_{L \in \mathcal{L}_k} h_\lambda^\Lambda(\vec{W}; L).$$

The following characterization of compatible bonds will be useful. We recall that the bonds of a lace with m bonds can be labeled as $\alpha_j \beta_j$ with

$$0 = \alpha_1 < \alpha_2 < \beta_1 \leq \alpha_3 < \beta_2 \leq \cdots \leq \alpha_m < \beta_{m-1} < \beta_m = k+1;$$

see [22, Equations (3.15) and (3.16)].

Lemma 4.1. (Characterization of compatible bonds.) *Let L be a lace with vertex set $V(L) = \{0, \ldots, k+1\}$ and bonds $\alpha_j \beta_j$, $j = 1, \ldots, m$, labeled from left to right (i.e., $\alpha_j < \alpha_{j+1}$). Then a bond $\alpha\beta \notin E(L)$ with $\alpha < \beta - 1$ is compatible with L if and only if either*

(a) $\alpha_i \leq \alpha < \beta \leq \beta_i$ *for* $i \in [m]$ *or*

(b) $\alpha_i < \alpha < \beta \leq \alpha_{i+2}$ *for* $i \in [m-1]$ *(where we set* $\alpha_{m+1} := k$*).*

Proof. Let $\alpha\beta \notin E(L)$ be compatible with L; that is, the algorithm below Definition 4.4 maps $E(L) \cup \{\alpha\beta\}$ to $E(L)$, which in turn means that $\alpha\beta$ is not selected to be part of the output lace. We show that then either (a) or (b) is satisfied. Assume the algorithm has already constructed the partial lace up to some $j < m$, producing the bonds $\big(\alpha_i', \beta_i'\big)_{i=1}^j$ (note that they are in reverse order and make up the last j bonds of the lace). Assume moreover that $\alpha_j' < \beta \leq \alpha_{j-1}'$; that is, $\alpha\beta$ is a potential candidate to be chosen as the next bond of the lace. Since it is *not* chosen, there is $\alpha_{j+1}' \beta_{j+1}'$ with $\beta_{j+1}' \in \big(\alpha_j', \alpha_{j-1}'\big]$ such that either

- $\alpha_{j+1}' < \alpha$, or
- $\alpha_{j+1}' = \alpha$ and $\beta_{j+1}' > \beta$.

Both the second case and the first case under the additional assumption $\beta_{j+1}' \geq \beta$ imply that $\alpha\beta$ satisfies (a). Let us thus focus on the case where $\alpha_{j+1}' < \alpha$ and $\beta_{j+1}' < \beta$. Remembering the stage of the algorithm, we have $\beta \leq \alpha_{j-1}'$, implying (b).

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

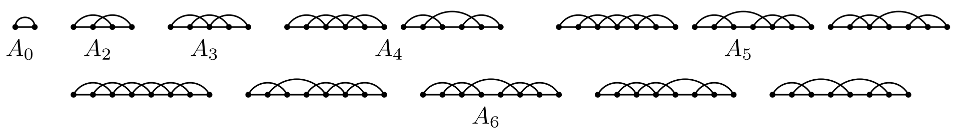

FIGURE 4. Schematic illustration of A_i from the proof of Lemma 4.2 for $i = 0, 2, 3, 4, 5, 6$.

Now let $\alpha\beta \notin E(L)$ be a bond that satisfies (a) or (b). We claim that $\alpha\beta$ is compatible with L. Let i be the index such that $\alpha_i\beta_i$ satisfies (a) or (b). Note that in the execution of the algorithm below Definition 4.3, $\alpha\beta$ does not appear as a candidate to be added to the constructed lace up until the point where $\alpha_m\beta_m, \alpha_{m-1}\beta_{m-1}, \ldots, \alpha_{i+1}\beta_{i+1}$ have already been added to the partial lace. At this stage of the algorithm, if $\alpha\beta$ satisfies (b), then it is not picked, because the left endpoint of the bond $\alpha_i\beta_i$ has a smaller value (i.e., $\alpha_i < \alpha$). If $\alpha\beta$ satisfies (a), however, then either also $\alpha_i < \alpha$, or $\alpha_i = \alpha$, but $\alpha_i\beta_i$ has its right endpoint further to the right (i.e., $\beta < \beta_i$, since the two bonds cannot be equal), and so again, $\alpha_i\beta_i$ is picked by the algorithm. $\square$

To prove the second result of Proposition 4.1, we need the following counting lemma, which may be of independent interest.

Lemma 4.2. (On the number of laces.) *Let f_i be the ith Fibonacci number with $f_1 = 0$, $f_2 = 1$. Then*

$$|\mathcal{L}_k| = 1 + \sum_{i=1}^{k} \binom{k}{i} f_i \qquad \textit{and, as } k \to \infty, \qquad |\mathcal{L}_k| \sim \frac{1}{\sqrt{5}} \left(\frac{3+\sqrt{5}}{2} \right)^k.$$

Proof. We first choose i vertices in $\{1, \ldots, k\}$ and then count the laces that use exactly those vertices. To this end, let A_i be the set of laces L with $V(L) = \{0, \ldots, i+1\}$ so that every vertex is the endpoint of at least one stitch. We claim that $|A_i| = f_i$ for $i \geq 1$. Clearly, $|A_0| = 1$, $|A_1| = 0$, $|A_2| = 1$. See Figure 4 for an illustration.

Let $i \geq 3$. We now establish the Fibonacci recursion. First, note that the bond incident to 0 (the 'first' bond) must always have 2 as the second endpoint. Now, depending on whether or not the third bond is incident to 2, the remaining lace lives on $\{1, 2, \ldots, i+1\}$ or on $\{1, 3, 4, \ldots, i+1\}$, and so $|A_i| = |A_{i-1}| + |A_{i-2}|$.

The asymptotic behavior follows from the fact that $f_n \sim \Phi^n/\sqrt{5}$, where $\Phi = \frac{1}{2}(1+\sqrt{5})$ is the golden ratio. $\square$

We can now work towards finding an explicit expression for $h_\lambda^\Lambda(\vec{W}; L)$ for a fixed lace. The next lemma is in the spirit of Observation 3.1 and will help us find probabilistic factors in the shell function.

Lemma 4.3. (Bipartite graphs and probabilities.) *Let $Y, A, B, C \subset \mathbb{R}^d$ be finite, disjoint sets.*

1. *Then*

$$\sum_{\substack{H \in \mathcal{G}(A \cup C, Y): \\ \forall y \in Y: y \sim A}} \mathbf{w}(H) = \prod_{y \in Y} \big(-\mathbb{P}(A \sim y \nsim C) \big) = (-1)^{|Y|} \mathbb{P}(\forall y \in Y : A \sim y \nsim C).$$

2. *Moreover,*

$$\sum_{\substack{H \in \mathcal{G}(A \cup B \cup C, Y): \\ \forall y \in Y: A \sim y \sim B}} \mathbf{w}(H) = \prod_{y \in Y} \mathbb{P}\big(A \sim y \sim B, y \nsim C \big).$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

3. *Lastly,*

$$\sum_{\substack{H\in\mathcal{G}(A,Y):\\ E(H)\neq\varnothing}} \mathbf{w}(H) = -\mathbb{P}(A\sim Y).$$

Proof. The first part of the statement is rather straightforward. If $Y=\{y\}$, then $\mathcal{G}(A\cup C, \{y\})$ is the set of star graphs (with center y). Observe first that

$$\sum_{H\in\mathcal{G}(A\cup C,\{y\}):\, y\sim A} \mathbf{w}(H) = \Bigg(\sum_{H'\in\mathcal{G}(A,\{y\}):\, y\sim A} \mathbf{w}(H')\Bigg)\Bigg(\sum_{H''\in\mathcal{G}(C,\{y\})} \mathbf{w}\big(H''\big)\Bigg).$$

The first sum is over all star graphs in $\mathcal{G}(A, \{y\})$ except the empty one, the second is over all star graphs in $\mathcal{G}(C, \{y\})$, and so

$$\sum_{H\in\mathcal{G}(A\cup C,\{y\}):\, y\sim A} \mathbf{w}(H) = -\Bigg(1-\prod_{x\in A}(1-\varphi(y, x))\Bigg)\prod_{x\in C}(1-\varphi(y, x)) = -\mathbb{P}(A\sim y\nsim C).$$

It is an easy induction to prove that for general Y, the sum factors into a product over sums over star graphs. For the second statement, assume again that $Y=\{y\}$ and observe that

$$\begin{aligned}\sum_{\substack{H\in\mathcal{G}(A\cup B\cup C,\{y\}):\\ A\sim y\sim B}} \mathbf{w}(H) &= \Bigg(\sum_{H\in\mathcal{G}(A\cup C,\{y\}):\, y\sim A} \mathbf{w}(H)\Bigg)\Bigg(\sum_{H\in\mathcal{G}(B,\{y\}):\, y\sim B} \mathbf{w}(H)\Bigg)\\ &= \mathbb{P}(A\sim y\sim B,\, y\nsim C),\end{aligned}$$

where the last identity is due to independence. The statement easily extends to general Y (again, the sum factors).

For the third statement, note that we sum over every graph except the empty one. □

Since the explicit expression for $h_\lambda^\Lambda\big(\vec{W}; L\big)$ is a lengthy product of probabilities, we first introduce some notation to represent the factors of this product compactly. Let A, B be two subsets of $[k+1]_0$. We define the set of all possible direct stitches in H leading to bonds $\alpha\beta\in E(\hat{H})$ with $\alpha\in A$, $\beta\in B$ as

$$\Upsilon(A, B) := \big\{xy\subset W:\ \exists\alpha\in A,\, \beta\in B \text{ with } \alpha<\beta-1 \text{ and } x\in\{u_\alpha\}\cup V_\alpha,\ \ y\in V_{\beta-1}\cup\{u_\beta\}\big\},$$

and we write $\Upsilon(A)=\Upsilon(A, A)$. We define

$$q_{\alpha,\beta} := \prod_{xy\in\Upsilon([\alpha,\beta))\cup\Upsilon((\alpha,\beta])} (1-\varphi(x-y))$$

and, for $\alpha_1<\alpha_2<\alpha_3$,

$$q_{\alpha_1,\alpha_2,\alpha_3} := \prod_{xy\in\Upsilon([\alpha_1+1,\alpha_2),[\alpha_2,\alpha_3))} (1-\varphi(x-y)).$$

Note that these products encode the sum over all $\mathbf{w}$-weighted graphs on the set of edges multiplied over.

To lighten notation, for $0\le\alpha\le\beta\le k+1$, set

$$\begin{aligned}[u_\alpha]\!] &:= \{u_\alpha\}\cup V_\alpha, \qquad [\![u_\beta] := V_{\beta-1}\cup\{u_\beta\},\\ [u_\alpha, u_\beta] &:= \{u_\alpha\}\cup V_\alpha\cup\cdots\cup V_{\beta-1}\cup\{u_\beta\}.\end{aligned}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

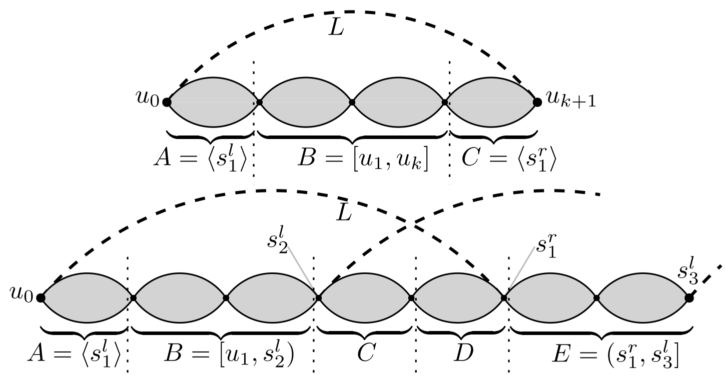

FIGURE 5. Illustration of the induction proof of Lemma 4.4. The lace L is sketched using dashed lines. The left picture shows the base case $m = 1$, where $u_0 = s_1$ and $u_{k+1} = t_1$. To the right, the first three stitches of L are (partially) sketched. The sets C, D are defined as $C = [s_2, t_1\rangle\!\rangle$ and $D = [\![t_1]\!]$.

We extend this notation further: for $a, b \in \{u_0, \ldots, u_{k+1}\}$, let $(a, b) := [a, b] \setminus \{a, b\}$, let $[a, b) := [a, b] \setminus \{b\}$, and let $(a, b] := [a, b] \setminus \{a\}$. We set $\langle\!\langle a, b\rangle\!\rangle := [a, b] \setminus ([a]\!] \cup [\![b])$ and define sets $\langle\!\langle a, b]$ etc. accordingly.

Moreover, define

$$Q_{\alpha,\beta} = \mathbb{P}_\lambda\big(\nexists y \in \eta_\Lambda \text{ s.t. } [u_\alpha]\!] \sim y \sim [\![u_\beta],\, y \nsim [u_{\alpha+1}, u_{\beta-1}]\big)$$

for $\beta \geq \alpha + 2$. We extend this notation by writing

$$Q_{A,B} = \prod_{\alpha \in A} \prod_{\beta \in B} Q_{\alpha,\beta}$$

for sets of pivotal points A, B; we abbreviate $Q_{a,[b,c]} = Q_{\{a\},[b,c]}$.

We are now ready to state Lemma 4.4, for which we recall the definition of $h_\lambda^\Lambda\big(\vec{W}; L\big)$ in (4.9).

Lemma 4.4. (The shell function of a lace.) *Let $\lambda \geq 0$ and let $\Lambda \subset \mathbb{R}^d$ be bounded. Let $W \subset \mathbb{R}^d$ be a core vertex set with pivot decomposition $\vec{W} = (u_0, V_0, \ldots, u_{k+1})$. Let L be a lace with vertex set $[k+1]_0$ and m bonds $\alpha_i\beta_i$, $i \in [m]$. Then, setting $\alpha_{m+1} = k$, we have*

$$\begin{aligned} h_\lambda^\Lambda\big(\vec{W}; L\big) = \mathbb{P}_\lambda(\eta_\Lambda \not\longleftrightarrow W) \prod_{i=1}^{m} q_{\alpha_i,\beta_i}\Big[1 - Q_{\alpha_i,\beta_i} - \mathbb{P}([u_{\alpha_i}]\!] \sim [\![u_{\beta_i}])\Big] \\ \times \prod_{i=1}^{m-1} q_{\alpha_i,\alpha_{i+1},\alpha_{i+2}} Q_{\alpha_i,(\beta_i,k+1]} Q_{(\alpha_i,\alpha_{i+1}),(\alpha_{i+2},k+1]}. \end{aligned} \tag{4.10}$$

Moreover,

$$\sum_{n \geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \Big| \sum_{H \in \langle\!\langle L \rangle\!\rangle : \mathcal{S}(H) = \vec{y}_{[n]}} \mathbf{w}(H) \Big| \,\mathrm{d}\vec{y}_{[n]} \leq 2^m \mathrm{e}^{3\lambda|W|}. \tag{4.11}$$

Proof. We abbreviate $\eta = \eta_\Lambda$, $h = h_\lambda^\Lambda$, and prove the statement by induction on m.

Base case. Let $m = 1$, which means that $\alpha_1 = 0$ and $\beta_1 = k + 1$. Set $A = [u_0]\!]$, $B = [u_1, u_k]$, and $C = [\![u_{k+1}]$. See Figure 5 for an illustration of A, B, C.

Note first that the edge set $\Upsilon([k+1]_0) \setminus E(A, C)$, that is, the possible direct stitches between points of W except the direct ones between A and C, do not determine membership of H in

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

$\langle\langle L\rangle\rangle$. Any such edge xy may or may not be present, resulting in a factor $(1-\varphi(x-y))$ that can be extracted. In total, this produces the factor $q_{0,k+1}$, and we can restrict to considering graphs $H \in \langle\langle L\rangle\rangle$ that do not possess any such edge. The remaining graphs H only have edges that are incident to $A \cup C \cup \mathcal{S}(H)$.

We split this set of remaining graphs H into those that have a direct stitch between A and C and those that do not. Among the former, the sum over graphs factors into graphs $H' \in \mathcal{G}(A, C)$ (the direct stitches) and graphs $H'' \in \mathcal{G}(W, \mathcal{S}(H))$. With Lemma 4.3,

$$\begin{aligned} h(\vec{W}; L) = q_{0,k+1} \Bigg[& \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \Bigg(\sum_{H'\in\mathcal{G}(A,C):E(H')\neq\varnothing} \mathbf{w}(H') \Bigg) \Bigg(\sum_{\substack{H''\in\mathcal{G}(W,\vec{y}_{[n]}):\\ y_i\sim W \forall i\in[n]}} \mathbf{w}\big(H''\big) \Bigg) \,\mathrm{d}\vec{y}_{[n]} \\ & + \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(W,\vec{y}_{[n]}):\\ \deg(y_i)\geq 1 \forall i\in[n],\\ \exists i: A\sim y_i\sim C}} \mathbf{w}(H)\,\mathrm{d}\vec{y}_{[n]} \Bigg] \\ = q_{0,k+1} \Bigg[& -\mathbb{P}(A\sim C)\mathbb{P}_\lambda(\eta \not\longleftrightarrow W) + \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(W,\vec{y}_{[n]}):\\ \deg(y_i)\geq 1 \forall i\in[n],\\ \exists i: A\sim y_i\sim C}} \mathbf{w}(H)\,\mathrm{d}\vec{y}_{[n]} \Bigg]. \end{aligned} \tag{4.12}$$

For now the power series are treated as formal power series; convergence is proven later. To treat the sum in (4.12), we define

$$\mathcal{S}_1 := \{y : A\sim y\sim C\}, \quad \mathcal{S}_2 := \{y : C \not\sim y \sim (A\cup B)\}, \quad \text{and } \mathcal{S}_3 := \{y : C\sim y \not\sim A\}.$$

With these definitions, we can partition $\vec{y} = \mathcal{S}(H) = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$. Moreover, we know that $\mathcal{S}_1 \neq \varnothing$. Re-summing and then applying Lemma 4.3, the sum over n in (4.12) becomes

$$\begin{aligned} & \sum_{n_1,n_2,n_3\geq 0} \frac{\lambda^{n_1+n_2+n_3}}{n_1!n_2!n_3!} \int_{\Lambda^{n_1+n_2+n_3}} \sum_{\substack{H\in\langle\langle L\rangle\rangle:\\ \mathcal{S}_i(H)=\vec{y}_{i,[n_i]}\forall i\in[3]}} \mathbf{w}(H)\,\mathrm{d}\big(\vec{y}_{1,[n_1]}, \vec{y}_{2,[n_2]}, \vec{y}_{3,[n_3]}\big) \\ &= \Bigg(\sum_{n\geq 1} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(A\cup B\cup C,\vec{y}_{[n]}):\\ \forall i\in[n]: A\sim y_i\sim C}} \mathbf{w}(H)\,\mathrm{d}\vec{y}_{[n]} \Bigg) \Bigg(\sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(A\cup B,\vec{y}_{[n]}):\\ \forall i\in[n]: y_i\sim(A\cup B)}} \mathbf{w}(H)\,\mathrm{d}\vec{y}_{[n]} \Bigg) \\ &\quad \times \Bigg(\sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(B\cup C,\vec{y}_{[n]}):\\ \forall i\in[n]: y_i\sim C}} \mathbf{w}(H)\,\mathrm{d}\vec{y}_{[n]} \Bigg) \\ &= \Bigg(\sum_{n\geq 1} \frac{\lambda^n}{n!} \Big(\int_{\Lambda} \mathbb{P}(A\sim y\sim C, y \not\sim B)\mathrm{d}y \Big)^n \Bigg) \Bigg(\sum_{n\geq 0} \frac{\lambda^n}{n!} \Big(-\int_{\Lambda} \mathbb{P}(y\sim (A\cup B))\,\mathrm{d}y \Big)^n \Bigg) \\ &\quad \times \Bigg(\sum_{n\geq 0} \frac{\lambda^n}{n!} \Big(-\int_{\Lambda} \mathbb{P}(C\sim y\not\sim B)\,\mathrm{d}y \Big)^n \Bigg). \end{aligned} \tag{4.13}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Recognizing the exponential series in the expression above, we can rewrite the probabilities with respect to $\mathbb{P}$ (appearing in the exponents) as probabilities with respect to $\mathbb{P}_\lambda$ associated to ξ^y, e.g., $\mathbb{P}(y \sim (A \cup B)) = \mathbb{P}_\lambda(y \sim (A \cup B) \text{ in } \xi^y)$. Then we can apply the univariate Mecke formula (see (2.1) for $m = 1$) to rewrite (4.13) as

$$\begin{aligned}&\Big(\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: A\sim y\sim C, y\nsim B\}|]} - 1\Big)\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B)\}|]}\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: C\sim y\nsim B\}|]}\\ =&\Big(1 - \mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: A\sim y\sim C, y\nsim B\}|]}\Big)\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B)\}|]}\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: C\sim y\nsim (A\cup B)\}|]}\\ =&(1 - Q_{0,k+1})\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B\cup C)\}|]}.\end{aligned}$$

Since $\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B\cup C)\}|]} = \mathbb{P}_\lambda(\eta \not\leftrightarrow W)$, we can plug this back into (4.12) and obtain

$$h\big(\vec{W}; L\big) = \mathbb{P}_\lambda(\eta \not\leftrightarrow W) q_{0,k+1}\Big(1 - Q_{0,k+1} - \mathbb{P}(A \sim C)\Big)$$

on the level of formal power series. Now we prove convergence and check that the previous computational steps are justified not only on the level of formal power series. We first revisit Equation (4.13). On the left-hand side, let us put absolute values inside the integral (but outside the sum over shell graphs H). The resulting expression is bounded by the middle part of (4.13), again with absolute values inside the integral. Each integrand is bounded in absolute value by a probability; hence it is smaller than or equal to 1. The resulting series are exponential series and, in particular, absolutely convergent. As a consequence, Equation (4.13) is justified and the last sum in (4.12) is bounded as

$$\begin{aligned}\sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^n}&\left|\sum_{\substack{H\in\mathcal{G}(W,\vec{y}_{[n]}):\\ \deg(y_i)\geq 1\,\forall i\in[n],\\ \exists i: A\sim y_i\sim C}} \mathbf{w}(H)\right| \mathrm{d}\vec{y}_{[n]}\\ &\leq \mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: A\sim y\sim C, y\nsim B\}|]}\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B)\}|]}\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: C\sim y\nsim B\}|]}\\ &\leq \mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: y\sim A\}|]}\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: y\sim(A\cup B)\}|]}\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: y\sim C\}|]}\\ &\leq \mathrm{e}^{2\lambda|W|},\end{aligned} \tag{4.14}$$

where for the last inequality we use the fact that the expected number of direct neighbors of any fixed element of W with respect to η is given by $\lambda\int\varphi(x)\,\mathrm{d}x$, as well as the rescaling introduced in Section 2.3 ensuring that $\int\varphi(x)\,\mathrm{d}x = 1$; compare this bound to the one used in (3.4). For the other contribution to $h\big(\vec{W}; L\big)$, we notice that

$$\begin{aligned}\sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^n}&\left|\left(\sum_{H'\in\mathcal{G}(A,C): E(H)\neq\varnothing} \mathbf{w}(H')\right)\left(\sum_{\substack{H''\in\mathcal{G}(W,\vec{y}_{[n]}):\\ y_i\sim W\,\forall i\in[n]}} \mathbf{w}\big(H''\big)\right)\right| \mathrm{d}\vec{y}_{[n]}\\ &\leq \mathbb{P}(A\sim C)\,\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta: y\sim W\}|]} \leq \mathrm{e}^{\lambda|W|},\end{aligned} \tag{4.15}$$

by the same argument as in (4.14).

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Combining (4.14) and (4.15) with (4.12) and $0 \le q_{0,k+1} \le 1$, we deduce

$$\sum_{n\ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \left| \sum_{H\in\langle\!\langle L\rangle\!\rangle : \mathcal{S}(H)=\vec{y}_{[n]}} \mathbf{w}(H) \right| \, \mathrm{d}\vec{y}_{[n]} \le \mathrm{e}^{\lambda|W|} + \mathrm{e}^{2\lambda|W|} \le 2\mathrm{e}^{2\lambda|W|} < \infty.$$

Inductive step. For the inductive step, let $m > 1$. We write the lace L in terms of its vertices (s_i, t_i) in W (that is $s_i = u_{\alpha_i}$ and $t_i = u_{\beta_i}$) and let L' be the lace on $W' := W \setminus [s_1, s_2)$ obtained from L by deleting the first stitch. We note that if $H \in \langle\!\langle L\rangle\!\rangle$, then $H[[s_2, u_{k+1}]] \in \langle\!\langle L'\rangle\!\rangle$. Observe that

$$h\big(\vec{W}; L\big) = h\big(\vec{W}'; L'\big) \sum_{n\ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(\overline{V}_0,\dots,\overline{V}_{\alpha_3-1},\vec{y}_{[n]}):\\ H\oplus L'\in\langle\!\langle L\rangle\!\rangle}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[n]}. \tag{4.16}$$

Again we first prove (4.10) and carry out computations on the level of formal power series; we prove convergence (and thus (4.11)) at the end. We can apply the induction hypothesis to $h(\vec{W}'; L)$; it remains to deal with the second factor. We partition the vertices in $[s_1, s_3]$ as $A = [s_1]\!]$, $B = (\!(s_1, s_2)$, $C = [s_2, t_1)\!)$, $D = [\![t_1]$, and $E = (t_1, s_3]$ (see Figure 5). If $m = 2$, we let $E = (t_1, u_k]$.

The graphs summed over in (4.16) must satisfy the following restraints: there must be at least one direct or indirect stitch between A and D, and there cannot be any (direct or indirect) edge between A and E. In particular, the remaining direct stitches may or may not be there, and thus can be extracted as the factor $q_{\alpha_1,\alpha_2,\alpha_3}$.

We partition $\mathcal{S}(H) = \cup_{i=1}^4 \mathcal{S}_i$, where

$$\mathcal{S}_1 = \{y : A \sim y \sim D, N(y) \subseteq [s_1, t_1]\}, \qquad \mathcal{S}_2 = \{y : A \sim y \sim C, N(y) \subseteq [s_1, t_1)\!)\},$$
$$\mathcal{S}_3 = \{y : \varnothing \ne N(y) \subseteq [s_1, s_2)\}, \qquad \mathcal{S}_4 = \{y : B \sim y \sim (C \cup D \cup E), N(y) \subseteq (\!(s_1, s_3]\}.$$

Again, we intend to split the sum over graphs into those that have at least one direct stitch between A and D, and those that do not. We can thus rewrite the second factor in (4.16) as

$$\begin{aligned}
&q_{\alpha_1,\alpha_2,\alpha_3} \sum_{n\ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}(\overline{V}_0,\dots,\overline{V}_{\alpha_3-1},\vec{y}_{[n]}):\\ H\oplus L'\in\langle\!\langle L\rangle\!\rangle,\\ \forall e\in E(H): e\cap(A\cup D\cup\vec{y}_{[n]})\ne\varnothing}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[n]} \\
&= q_{\alpha_1,\alpha_2,\alpha_3} \prod_{i=2}^{4} \left(\sum_{n_i\ge 0} \frac{\lambda^{n_i}}{n_i!} \int_{\Lambda^{n_i}} \sum_{\substack{H\in\mathcal{G}(W,\vec{y}_{i,[n_i]}):\\ \mathcal{S}(H)=\mathcal{S}_i}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[i,[n_i]} \right) \\
&\quad \times \Bigg[-\mathbb{P}(A\sim D) \sum_{n\ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}([s_1,s_3],\vec{y}_{[n]}):\\ y_i\sim A\cup B \forall i\in[n]}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[n]} \\
&\qquad + \sum_{n\ge 1} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\substack{H\in\mathcal{G}([s_1,s_3],\vec{y}_{[n]}):\\ y_i\sim A\cup B \forall i\in[n]}} \mathbf{w}(H) \, \mathrm{d}\vec{y}_{[n]} \Bigg] \\
&= q_{\alpha_1,\alpha_2,\alpha_3} \Big(-\mathbb{P}(A\sim D) \mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta : y\in\mathcal{S}_1\}|]} + \mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta : y\in\mathcal{S}_1\}|]} - 1 \Big) \\
&\quad \times \exp\Big\{ \mathbb{E}_\lambda[|\{y\in\eta : y\in\mathcal{S}_2\}|] - \mathbb{E}_\lambda[|\{y\in\eta : y\in\mathcal{S}_3\}|] + \mathbb{E}_\lambda[|\{y\in\eta : y\in\mathcal{S}_4\}|] \Big\},
\end{aligned} \tag{4.17}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

where the last identity was obtained using Lemma 4.3. Note that the factor $h\big(\vec{W}'; L'\big)$ contains the factor $\mathbb{P}(\eta \not\leftrightarrow [s_2, u_{k+1}]) = \mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta: y\sim[s_2,u_{k+1}]\}|]}$. Together with this factor, (4.17) equals

$$\begin{aligned}\exp\big\{\mathbb{E}_\lambda\big[&-|\{y\in\eta:(A\cup B\sim y\sim[s_2,u_{k+1}]\}|+|\{y\in\eta:A\sim y\sim D,\,y\nsim(B\cup C)\}|\\&+|\{y\in\eta:A\sim y\sim C,\,y\nsim B\}|+|\{y\in\eta:B\sim y\sim(C\cup D\cup E)\}|\big]\big\}\\&\times\mathbb{P}_\lambda(\eta\not\leftrightarrow W)\big(1-Q_{A,D}-\mathbb{P}(A\sim D)\big).\end{aligned}\tag{4.18}$$

It remains to rewrite the argument in the expectation of the exponent in (4.18). Note that

$$\begin{aligned}&-|\{y\in\eta:A\sim y\sim[s_2,u_{k+1}],\,y\nsim B\}|-|\{y\in\eta:B\sim y\sim[s_2,u_{k+1}]\}|\\&+|\{y\in\eta:A\sim y\sim(C\cup D),\,y\nsim B\}|+|\{y\in\eta:B\sim y\sim(C\cup D\cup E)\}|\\=&-|\{y\in\eta:A\sim y\sim(t_1,u_{k+1}],\,y\nsim(B\cup C\cup D)\}|\\&-|\{y\in\eta:B\sim y\sim(s_3,u_{k+1}],\,y\nsim(C\cup D\cup E)\}|.\end{aligned}$$

This gives two exponential terms. The first is

$$\begin{aligned}&\exp\Big\{-\mathbb{E}_\lambda\Big[\Big|\Big\{y\in\eta:[u_{\alpha_1}]\!]\sim y\sim(u_{\beta_1},u_{k+1}],\,y\nsim\big(\!\big(u_{\alpha_1},u_{\beta_1}\big]\!\big]\Big\}\Big|\Big]\Big\}\\&=\prod_{j=\beta_1+1}^{k+1}\exp\Big\{-\mathbb{E}_\lambda\Big[\Big|\Big\{y\in\eta:[u_{\alpha_1}]\!]\sim y\sim[\![u_j],\,y\nsim\big(\!\big(u_{\alpha_1},u_j\big)\!\big)\Big\}\Big|\Big]\Big\}\\&=Q_{\alpha_1,(\beta_1,k+1]}.\end{aligned}$$

Similarly, the second exponential term equals $Q_{(\alpha_1,\alpha_2),(\alpha_3,k+1]}$.

Again, we prove convergence and justify the previous computational steps. Revisiting the left-hand side of (4.17), we insert absolute values inside the integral (and outside the sum over graphs H). As in the base case, this is bounded by the middle part of (4.17) with absolute values in the integrals, and each integrand is a probability. With the Mecke equation, we obtain

$$\begin{aligned}&\sum_{n\ge0}\frac{\lambda^n}{n!}\int_{\Lambda^n}\Bigg|\sum_{\substack{H\in\mathcal{G}(\overline{V}_0,\dots,\overline{V}_{\alpha_3-1},\vec{y}_{[n]}):\\ H\oplus L'\in\langle\!\langle L\rangle\!\rangle,\\ \forall e\in E(H):e\cap(A\cup D\cup\vec{y}_{[n]})\ne\varnothing}}\mathbf{w}(H)\Bigg|\,\mathrm{d}\vec{y}_{[n]}\\&\le2\exp\Big\{\mathbb{E}_\lambda[|\{y\in\eta:y\in\mathcal{S}_1\}|]+\mathbb{E}_\lambda[|\{y\in\eta:y\in\mathcal{S}_2\}|]\\&\qquad\qquad+\mathbb{E}_\lambda[|\{y\in\eta:y\in\mathcal{S}_3\}|]+\mathbb{E}_\lambda[|\{y\in\eta:y\in\mathcal{S}_4\}|]\Big\}\\&\le2\mathrm{e}^{3\lambda|A\cup B|},\end{aligned}$$

arguing as in (4.14) for the last inequality.

Note that by the induction hypothesis, the term $h\big(\vec{W}'; L'\big)$ with absolute values in the respective integrals is bounded by $2^{m-1}\mathrm{e}^{3\lambda|W'|}$. Since $A\cup B$ and W' are disjoint, this proves (4.11). □

Proof of Proposition 4.1. Again, we abbreviate $\eta=\eta_\Lambda$ and $h=h_\lambda^\Lambda$. First, consider $k=0$, i.e., pivot decompositions with no pivotal points. Then there are no direct stitches, and we have

$$h^{(m)}\big(\vec{W},\vec{y}_{[m]}\big)=(-1)^m\prod_{i=1}^m\mathbb{P}(y_i\sim W),\qquad h\big(\vec{W}\big)=\mathbb{P}_\lambda(\eta\not\leftrightarrow W).$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Moreover,

$$\sum_{m\geq 0}\frac{\lambda^m}{m!}\int_{\Lambda^m}|h^{(m)}(\vec{W},\vec{y}_{[m]})|\,\mathrm{d}\vec{y}_{[m]}=\mathrm{e}^{\mathbb{E}_\lambda[|\{y\in\eta:y\sim W\}|]}\leq \mathrm{e}^{\lambda|W|},$$

using the same bound as in (4.15). Since this proves the proposition for $k=0$, we turn to $k\geq 1$ and we first prove (4.7).

We rewrite $h(\vec{W})$ by explicitly writing out the sum over laces L in terms of the endpoints of their stitches in W (note that any lace can have at most k stitches). We first exhibit this for $k=2$, where $\vec{W}=(u_0,V_0,u_1,V_1,u_2,V_2,u_3)$ and there are two different laces. With the abbreviation $\tilde{Q}_{i,j}=Q_{i,j}+\mathbb{P}([u_i]\!] \sim [\![u_j])$,

$$\begin{aligned}h(\vec{W})&=h(\vec{W};L_1)+h(\vec{W};L_2)=\mathbb{P}_\lambda(\eta \not\leftrightarrow W)\big(q_{0,3}(1-\tilde{Q}_{0,3})+Q_{0,3}(1-\tilde{Q}_{0,2})(1-\tilde{Q}_{1,3})\big)\\&=\mathbb{P}_\lambda(\eta \not\leftrightarrow W)\sum_{\beta_1=2}^{3}q_{0,\beta_1}(1-\tilde{Q}_{0,\beta_1})\Bigg[\mathbb{1}_{\{\beta_1=3\}}+\mathbb{1}_{\{\beta_1<3\}}\sum_{\alpha_2=1}^{\beta_1-1}Q_{0,3}(1-\tilde{Q}_{\alpha_2,3})\Bigg].\end{aligned}\tag{4.19}$$

Clearly, this is unnecessarily complicated for $k=2$, as the sum over α_2 contains only one term and $q_{0,2}=1$. However, this turns out to be convenient for general k. We use the convention that $Q_{[a,b],\varnothing}=Q_{\varnothing,[a,b]}=1$. Carefully rearranging the sum over all laces yields

$$\begin{aligned}h(\vec{W})=\sum_{L\in\mathcal{L}(\vec{W})}h(\vec{W};L)&=\mathbb{P}_\lambda(\eta \not\leftrightarrow W)\sum_{\beta_1=2}^{k+1}q_{0,\beta_1}(1-\tilde{Q}_{0,\beta_1})Q_{0,(\beta_1,k+1]}\\&\times\Bigg[\mathbb{1}_{\{\beta_1=k+1\}}+\sum_{\alpha_2=1}^{\beta_1-1}\sum_{\beta_2=\beta_1+1}^{k+1}q_{\alpha_2,\beta_2}(1-\tilde{Q}_{\alpha_2,\beta_2})Q_{(0,\alpha_2],(\beta_2,k+1]}Q_{(0,\alpha_2),\beta_2}\\&\times\Bigg[\mathbb{1}_{\{\beta_2=k+1\}}+\sum_{\alpha_3=\beta_1}^{\beta_2-1}\sum_{\beta_3=\beta_2+1}^{k+1}q_{\alpha_3,\beta_3}(1-\tilde{Q}_{\alpha_3,\beta_3})q_{\alpha_1,\alpha_2,\alpha_3}\\&\qquad Q_{(\alpha_2,\alpha_3],(\beta_3,k+1]}Q_{(\alpha_2,\alpha_3),\beta_3}Q_{(0,\alpha_2),(\alpha_3,\beta_2)}\\&\times\Bigg[\mathbb{1}_{\{\beta_3=k+1\}}+\sum_{\alpha_4=\beta_2}^{\beta_3-1}\sum_{\beta_4=\beta_3+1}^{k+1}q_{\alpha_4,\beta_4}(1-\tilde{Q}_{\alpha_4,\beta_4})q_{\alpha_2,\alpha_3,\alpha_4}\\&\qquad Q_{(\alpha_3,\alpha_4],(\beta_4,k+1]}Q_{(\alpha_3,\alpha_4),\beta_4}Q_{(\alpha_2,\alpha_3),(\alpha_4,\beta_3)}\times\cdots\\&\times\Bigg[\mathbb{1}_{\{\beta_{k-1}=k+1\}}+\mathbb{1}_{\{\beta_{k-1}<k+1\}}\sum_{\alpha_k=\beta_{k-2}}^{\beta_{k-1}-1}q_{\alpha_k,k+1}(1-\tilde{Q}_{\alpha_k,k+1})\prod_{j=k,k+1}q_{\alpha_{j-2},\alpha_{j-1},\alpha_j}\\&\qquad Q_{(\alpha_{k-1},\alpha_k),k+1}Q_{(\alpha_{k-2},\alpha_{k-1}),(\alpha_k,\beta_{k-1})}\Bigg]\cdots\Bigg]\Bigg]\Bigg].\end{aligned}$$

Note that if $\beta=k+1$ for some i, then the double sum following the corresponding indicator breaks down to 0. Also, only the innermost bracketed term contains two factors of $q_{a,b,c}$.

We now show that, starting with the innermost square brackets, the bracketed terms are bounded by 1 in absolute value.

To lighten notation, we write the innermost sum as $\sum_{\alpha=b_1}^{b_2-1}R(\alpha)$. We split the factor

$$1-\tilde{Q}_{\alpha_k,k+1}=(1-Q_{\alpha_k,k+1})-\mathbb{P}([u_{\alpha_k}]\!]\sim[\![u_{k+1}]).$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

This yields two sums

$$\sum_{\alpha=b_1}^{b_2-1} R(\alpha) = \sum_{\alpha=b_1}^{b_2-1} R'(\alpha) - \sum_{\alpha=b_1}^{b_2-1} R''(\alpha),$$

where R' and R'' are both nonnegative. Now, with the estimate $Q_{(\alpha_{k-1},\alpha),k+1} \le Q_{[\beta_{k-2},\alpha),k+1} = Q_{[b_1,\alpha),k+1}$, we can bound

$$\begin{aligned}\sum_{\alpha=b_1}^{b_2-1} R'(\alpha) &\le \sum_{\alpha=b_1}^{b_2-1} (1-Q_{\alpha,k+1})Q_{[b_1,\alpha),k+1}\\ &= (1-Q_{b_1,k+1}) + Q_{b_1,k+1} \sum_{\alpha=b_1+1}^{b_2-1} (1-Q_{\alpha,k+1})Q_{[b_1+1,\alpha),k+1},\end{aligned} \tag{4.20}$$

which is readily proven to be at most 1 by induction. Moreover,

$$\sum_{\alpha=b_1}^{b_2-1} R''(\alpha) \le \sum_{\alpha=b_1}^{b_2-1} q_{\alpha,k+1}\mathbb{P}([u_\alpha]\!] \sim [\![u_{k+1}]). \tag{4.21}$$

The above summands can be rewritten as the probability of the event that the direct stitch $(\alpha, k+1)$ is present, while all direct stitches $(j, k+1)$ for $j \in (\alpha, k+1]$ are not. Hence, these are disjoint events for different values of α, and so the sum is at most 1.

In total, we rewrote $\sum_{\alpha=b_1}^{b_2-1} R(\alpha)$ as the difference of two nonnegative values, both at most 1, proving our claim.

To deal with the summands for $2 \le i < k$, we write the double sum as

$$\sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} R(\alpha, \beta)$$

and split the term $1 - \tilde{Q}_{\alpha_i,\beta_i} = (1 - Q_{\alpha_i,\beta_i}) - \mathbb{P}([u_{\alpha_i}]\!] \sim [\![u_{\beta_i}])$ so that

$$\sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} R(\alpha, \beta) = \sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} R'(\alpha, \beta) - \sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} R''(\alpha, \beta) \tag{4.22}$$

for nonnegative summands R', R''. We prove a bound on the sum over $R'(\alpha, \beta)$ by induction on $k - b_2$. If $b_2 = k$, then the bound is the same as for the bound (4.20). For $b_2 < k$, we first bound $Q_{(\alpha_i,\alpha],(\beta,k+1]} \le Q_{[b_1,\alpha],(\beta,k+1]}$ and then extract the summand for $\beta = k + 1$, yielding

$$\begin{aligned}\sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} R'(\alpha, \beta) &\le \sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k+1} (1-Q_{\alpha,\beta})Q_{[b_1,\alpha],(\beta,k+1]}Q_{[b_1,\alpha),\beta}\\ &\le \sum_{\alpha=b_1}^{b_2-1} (1-Q_{\alpha,k+1})Q_{[b_1,\alpha),k+1}\\ &\quad + Q_{[b_1,b_2-1],k+1} \sum_{\alpha=b_1}^{b_2-1} \sum_{\beta=b_2+1}^{k} (1-Q_{\alpha,\beta})Q_{[b_1,\alpha],(\beta,k]}Q_{[b_1,\alpha),\beta}.\end{aligned} \tag{4.23}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

By the induction hypothesis, the double sum in (4.23) is at most 1. Therefore,

$$\begin{aligned}\sum_{\alpha=b_1}^{b_2-1}\sum_{\beta=b_2+1}^{k+1} R'(\alpha,\beta) &\le \sum_{\alpha=b_1}^{b_2-2}(1-Q_{\alpha,k+1})Q_{[b_1,\alpha),k+1}\\ &\quad + (1-Q_{b_2-1,k+1})Q_{[b_1,b_2-1),k+1} + Q_{b_2-1,k+1}Q_{[b_1,b_2-1),k+1}\\ &= \sum_{\alpha=b_1}^{b_2-2}(1-Q_{\alpha,k+1})Q_{[b_1,\alpha),k+1} + Q_{[b_1,b_2-2],k+1}\\ &= 1,\end{aligned}$$

where the last identity is now an easy induction.

Turning to the second summand in (4.22), by the same argument used to treat (4.21), the summands $R''(\alpha,\beta)$ are probabilities of events which are disjoint for different values of (α,β), and so they sum to at most 1.

The observation that the bracket term for $i=1$ is handled analogously finishes the proof of (4.7).

We proceed to prove (4.8) for $k>1$. By combining Lemma 4.2 with Lemma 4.4, we obtain

$$\begin{aligned}\sum_{m\ge 0}\frac{\lambda^m}{m!}\int_{\Lambda^m}\left|h^{(m)}(\vec{W},\vec{y}_{[m]})\right|\,\mathrm{d}\vec{y}_{[m]} &\le \sum_{L\in\mathcal{L}_k}\sum_{m\ge 0}\frac{\lambda^m}{m!}\int_{\Lambda^m}\Bigg|\sum_{H\in\langle\!\langle L\rangle\!\rangle:\mathcal{S}(H)=\vec{y}_{[m]}}\mathbf{w}(H)\Bigg|\,\mathrm{d}\vec{y}_{[m]}\\ &\le \frac{1}{\sqrt{5}}\left(\frac{3+\sqrt{5}}{2}\right)^k 2^k \mathrm{e}^{3\lambda|W|}.\end{aligned}$$

Using the bound $k\le|W|$ finishes the proof. □

Lemma 4.5. (Thermodynamic limit of the shell function.) *For every $\lambda\ge 0$, the pointwise limit*

$$\lim_{\Lambda\nearrow\mathbb{R}^d} h_\lambda^\Lambda(\vec{W}) = h_\lambda(\vec{W})$$

along $\mathbb{R}^d$-exhausting sequences exists.

Proof. Let $(\Lambda_n)_{n\in\mathbb{N}}$ be an $\mathbb{R}^d$-exhausting sequence. For fixed $\vec{W}=(u_0,V_0,\ldots,u_{k+1})$, note that

$$h_\lambda^{\Lambda_n}(\vec{W}) = \sum_{L\in\mathcal{L}_k} h_\lambda^{\Lambda_n}(\vec{W};L).$$

For each lace L, the limit

$$h_\lambda(\vec{W};L) = \lim_{n\to\infty} h_\lambda^{\Lambda_n}(\vec{W};L)$$

exists and does not depend on the precise choice of $\mathbb{R}^d$-exhausting sequence. This is clear from the representation for $h_\lambda^\Lambda(\vec{W};L)$ proven in Lemma 4.4. In particular, $h_\lambda^\Lambda(\vec{W};L)$ is given as the finite product of Λ-independent factors and factors that describe the probability of certain point processes containing no points (namely, $\mathbb{P}_\lambda(\eta_\Lambda \not\longleftrightarrow W)$ and the factors $Q_{i,j}$). As probabilities that are decreasing in the volume, the latter admit a $\Lambda\nearrow\mathbb{R}^d$ limit. It follows that the limit of the shell function exists as well and is given by

$$h_\lambda(\vec{W}) = \sum_{L\in\mathcal{L}_k} h_\lambda(\vec{W};L).$$

□

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

4.4. The direct-connectedness function in infinite volume

In this section, we consider the limit $\lim_{\Lambda\nearrow\mathbb{R}^d} g_\lambda^\Lambda$ with g_λ^Λ as in (4.6) and give sufficient conditions under which it exists, thereby proving the two convergence statements from Theorem 1.1.

The candidate limit is given by the analogue of (4.6) with Λ replaced by $\mathbb{R}^d$; the existence of $h_\lambda^{\mathbb{R}^d} \equiv h_\lambda$ has been checked in Lemma 4.5. Thus,

$$g_\lambda(x_1, x_2) = \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{(\mathbb{R}^d)^r} \sum_{\vec{W}} \left(\sum_{G\in\mathcal{G}_{\mathrm{core}}^{\vec{W}}} \mathbf{w}^{\pm}(G) \right) h_\lambda\big(\vec{W}\big)\, \mathrm{d}\vec{x}_{[3,r+2]}, \tag{4.24}$$

where the inner sum is over core graphs G on $\vec{x}_{[r+2]}$ with pivot decomposition $\vec{W}$, i.e., over $(+)$-connected graphs G on $\vec{x}_{[r+2]}$ with $\mathsf{PD}^+(x_1, x_2, G) = \mathsf{PD}^{\pm}(x_1, x_2, G) = \vec{W}$. Remember the quantities $0 < \tilde{\lambda}_* \le \lambda_*$ introduced before Theorem 1.1. We will see in (4.25) that the sum over core graphs for a given pivot decomposition is a probability, hence in particular nonnegative.

Theorem 4.1. (The thermodynamic limit of g_λ^Λ: pointwise convergence.) *If $\lambda < \lambda_*$, then*

$$\sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{(\mathbb{R}^d)^r} \sum_{\vec{W}} \left(\sum_{G\in\mathcal{G}_{\mathrm{core}}^{\vec{W}}} \mathbf{w}^{\pm}(G) \right) \big|h_\lambda\big(\vec{W}\big)\big|\, \mathrm{d}\vec{x}_{[3,r+2]} < \infty$$

for all $x_1, x_2 \in \mathbb{R}^d$. Moreover, for every $\mathbb{R}^d$-exhausting sequence $(\Lambda_n)_{n\in\mathbb{N}}$, we have the pointwise convergence

$$\lim_{n\to\infty} g_\lambda^{\Lambda_n}(x_1, x_2) = g_\lambda(x_1, x_2)$$

with g_λ given in (4.24) (equivalently, Equation (4.6) with Λ replaced by $\mathbb{R}^d$).

Theorem 4.2. (Integrability and convergence in the L^1-norm.) *If $\lambda < \tilde{\lambda}_*$, then for all $x_1 \in \mathbb{R}^d$,*

$$\int_{\mathbb{R}^d} |g_\lambda(x_1, x_2)|\, \mathrm{d}x_2 \le \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{\mathbb{R}^d} \left(\int_{(\mathbb{R}^d)^r} \sum_{\vec{W}} \left(\sum_{G\in\mathcal{G}_{\mathrm{core}}^{\vec{W}}} \mathbf{w}^{\pm}(G) \right) \big|h_\lambda\big(\vec{W}\big)\big|\, \mathrm{d}\vec{x}_{[3,r+2]} \right) \mathrm{d}x_2 < \infty.$$

Proof of Theorem 4.1. We consider a summand in (4.24) for fixed $\vec{W}$ and set $x_1 = u_0$ as well as $x_2 = u_{k+1}$. Let $\vec{W} = (u_0, V_0, \ldots, V_k, u_{k+1})$. Remember $\overline{V}_i = \{u_i\} \cup V_i \cup \{u_{i+1}\}$. A first important observation is the fact that the weight of a core graph with pivot decomposition $\vec{W}$ factors into the product over the $k(\pm)$-subgraphs induced by the vertex sets $\overline{V}_i$. The sum over core graphs thus factors as

$$\begin{aligned}
\sum_{\substack{G\in\mathcal{C}^+(W):\\ \mathsf{PD}^+(G)=\mathsf{PD}^{\pm}(G)=\vec{W}}} \mathbf{w}^{\pm}(G) &= \prod_{i=0}^{k} \left(\sum_{H\in\mathcal{D}^+_{u_i,u_{i+1}}(\overline{V}_i)} \mathbf{w}^{\pm}(H) \right) \\
&= \prod_{i=0}^{k} \mathbb{P}\Big(\Gamma_\varphi(\overline{V}_i) \in \mathcal{D}_{u_i,u_{i+1}}\Big) \\
&= \mathbb{P}\left(\bigcap_{i=0}^{k} \{\Gamma_\varphi(\overline{V}_i) \in \mathcal{D}_{u_i,u_{i+1}}\} \right).
\end{aligned} \tag{4.25}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Hence, the core can be written as a probability. Combining this with Proposition 4.1, we get

$$\left(\sum_{\substack{G \in \mathcal{C}^+(W): \\ \mathsf{PD}^+(G) = \mathsf{PD}^\pm(G) = \vec{W}}} \mathbf{w}^\pm(G) \right) \left| h_\lambda^\Lambda \big(\vec{W}\big) \right| \le \mathbb{P}_\lambda(\eta_\Lambda \not\longleftrightarrow W) \mathbb{P}\left(\bigcap_{i=0}^{k} \{\Gamma_\varphi(\overline{V}_i) \in \mathcal{D}_{u_i, u_{i+1}}\} \right)$$

$$= \mathbb{P}_\lambda\left(\left\{ \mathscr{C}\big(u_0, \xi_\Lambda^W\big) = W \right\} \cap \bigcap_{i=0}^{k} \left\{ \xi_\Lambda^W \big[\overline{V}_i\big] \in \mathcal{D}_{u_i, u_{i+1}} \right\} \right).$$

Above, we used independence as well as the fact that for $V \subseteq W$, the two random graphs $\Gamma_\varphi(V)$ and $\xi^W[V]$ are identical in distribution. The inequality holds true for bounded Λ as well as $\Lambda = \mathbb{R}^d$.

We now go back to (4.6) and rearrange the sum by first summing over the number of pivotal points k, giving

$$\sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \int_{\Lambda^r} \sum_{\vec{W}} \left(\sum_{G \in \mathcal{G}_{\text{core}}^{\vec{W}}} \mathbf{w}^\pm(G) \right) \left| h_\lambda^\Lambda \big(\vec{W}\big) \right| \,\mathrm{d}\vec{x}_{[3, r+2]}$$

$$= \sum_{k \ge 0} \lambda^k \sum_{n \ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^{k+n}} \sum_{\vec{W}} \left(\sum_{G \in \mathcal{G}_{\text{core}}^{\vec{W}}} \mathbf{w}^\pm(G) \right) \left| h_\lambda^\Lambda \big(\vec{W}\big) \right| \,\mathrm{d}\vec{v}_{[n]} \,\mathrm{d}\vec{u}_{[k]}. \tag{4.26}$$

In the second term, the sum is over pivot decompositions $\vec{W} = (u_0, V_0, \ldots, V_k, u_{k+1})$ where $u_0 = x_1$, $u_{k+1} = x_2$, and $\cup_{i=0}^k V_i = \{v_1, \ldots, v_n\}$.

When rewriting the integrand of (4.26) as a probability, the event that u_i and u_{i+1} are 2-connected for $i \in [k]_0$ in disjoint vertex sets V_i becomes the event that these connection events occur disjointly within W; see Section 2 and recall the definition (2.2). The inner series can thus be bounded as

$$\sum_{n \ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \sum_{\vec{W}} \left(\sum_{G \in \mathcal{G}_{\text{core}}^{\vec{W}}} \mathbf{w}^\pm(G) \right) |h_\lambda^\Lambda \big(\vec{W}\big)| \,\mathrm{d}\vec{v}_{[n]}$$

$$\le \sum_{n \ge 0} \frac{\lambda^n}{n!} \int_{\Lambda^n} \mathbb{P}_\lambda\Bigg(\left\{ \mathscr{C}\Big(u_0, \xi_\Lambda^{\vec{u}_{[k]}, \vec{v}_{[n]}}\Big) = \vec{u}_{[k]} \cup \vec{v}_{[n]} \right\}$$

$$\cap \left(\left\{ u_0 \Longleftrightarrow u_1 \text{ in } \xi^{u_0, u_1, \vec{v}_{[n]}} \right\} \circ \cdots \circ \left\{ u_k \Longleftrightarrow u_{k+1} \text{ in } \xi^{u_k, u_{k+1}, \vec{v}_{[n]}} \right\} \right) \Bigg) \,\mathrm{d}\vec{v}_{[n]}$$

$$= \mathbb{P}_\lambda\left(\left\{ u_0 \Longleftrightarrow u_1 \text{ in } \xi^{u_0, u_1} \right\} \circ \cdots \circ \left\{ u_k \Longleftrightarrow u_{k+1} \text{ in } \xi^{u_k, u_{k+1}} \right\} \right), \tag{4.27}$$

where the identity is due to the Mecke equation and the fact that by summing over $\vec{v}$, we were partitioning over what the joint cluster of $\vec{u}_{[0, k+1]}$ is. We can now use the BK inequality ([10, Theorem 2.1]) to bound (4.27) by

$$\prod_{i=0}^{k} \mathbb{P}_\lambda\Big(u_i \Longleftrightarrow u_{i+1} \text{ in } \xi_\Lambda^{u_i, u_{i+1}} \Big) \le \prod_{i=0}^{k} \sigma_\lambda(u_{i+1} - u_i). \tag{4.28}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Inserting this back into (4.26),

$$\sum_{k\geq 0}\lambda^k\sum_{n\geq 0}\frac{\lambda^n}{n!}\int_{\Lambda^{k+n}}\sum_{\vec{W}}\left(\sum_{G\in\mathcal{G}^{\vec{W}}_{\mathrm{core}}}\mathbf{w}^{\pm}(G)\right)\left|h^{\Lambda}_{\lambda}(\vec{W})\right|\,\mathrm{d}\vec{v}_{[n]}\,\mathrm{d}\vec{u}_{[k]}$$
$$\leq\sum_{k\geq 0}\lambda^k\sigma_{\lambda}^{*(k+1)}(x_2-x_1). \tag{4.29}$$

The last expression is finite for $\lambda<\lambda_*$, by the definition of λ_*. The pointwise convergence of $g^{\Lambda_n}_{\lambda}$ to g_λ follows by dominated convergence. □

Proof of Theorem 4.2. If we integrate over x_2 in (4.29), this yields the upper bound

$$\lambda^{-1}\sum_{k\geq 1}\left(\lambda\int\sigma_\lambda(x)\,\mathrm{d}x\right)^k,$$

which is finite for $\lambda<\tilde{\lambda}_*$, by definition of $\tilde{\lambda}_*$. The theorem follows by Fubini–Tonelli and the triangle inequality. □

5. The Ornstein–Zernike equation

Here we complete the proof of Theorem 1.1. In view of Theorems 4.1 and 4.2, it remains to prove that the expansion (4.24) is indeed equal to the direct-connectedness function given by the OZE (1.2). This is proven by showing first that g^{Λ}_{λ} from Definition 4.2 fulfills the OZE in finite volume and then passing to the limit $\Lambda\nearrow\mathbb{R}^d$.

The idea of the proof in finite volume is basically well known; the same proof works for the OZE for the total correlation function.

Proposition 5.1. (The Ornstein–Zernike equation in finite volume.) *Let $\Lambda\subset\mathbb{R}^d$ be bounded and let $x_1, x_2\in\Lambda$. Then*

$$\tau^{\Lambda}_{\lambda}(x_1,x_2)=g^{\Lambda}_{\lambda}(x_1,x_2)+\lambda\int_{\Lambda}g^{\Lambda}_{\lambda}(x_1,x_3)\tau^{\Lambda}_{\lambda}(x_3,x_2)\,\mathrm{d}x_3.$$

Proof. We drop the Λ-dependence in the superscript of τ^{Λ}_{λ} and g^{Λ}_{λ}. Thanks to Proposition 3.1, we can re-sum the series expansion for τ_λ at will. Given a pivot decomposition $\vec{W}=(u_0,V_0,\ldots,u_{k+1})$ of an arbitrary core graph G with the vertex set W, define

$$\bar{h}^{(m)}_{\lambda}(\vec{W},\vec{y}_{[m]}):=\sum_{\substack{H\in\mathcal{G}(\overline{V}_1,\ldots,\overline{V}_k,\vec{y}_{[m]}):\\ G\oplus H\in\mathcal{C}^{\pm}_{u_0,u_{k+1}}(W\cup\vec{y}_{[m]})}}\mathbf{w}(H),\qquad \bar{h}_\lambda(\vec{W}):=\sum_{m\geq 0}\frac{\lambda^m}{m!}\int_{\Lambda^m}\bar{h}^{(m)}_{\lambda}(\vec{W},\vec{y}_{[m]})\,\mathrm{d}\vec{y}_{[m]}, \tag{5.1}$$

in analogy to the shell function h_λ in (4.5) (just like the latter, $\bar{h}_\lambda$ only depends on G through its pivot decomposition $\vec{W}$). To be more precise, the shell function h_λ is recovered from $\bar{h}_\lambda$ by summing over a smaller subset of graphs H in (5.1), adding the restriction that $G\oplus H$ shall not contain $(\pm)$-pivot points for the u_0–u_{k+1} connection. Note that

$$0\leq\bar{h}_\lambda(\vec{W})=\mathrm{e}^{-\mathbb{E}_\lambda[|\{y\in\eta:y\sim W\}|]}\prod_{x,y\in W:\nexists i\in[k]_0:\{x,y\}\subseteq\bar{V}_i}(1-\varphi(x-y))$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

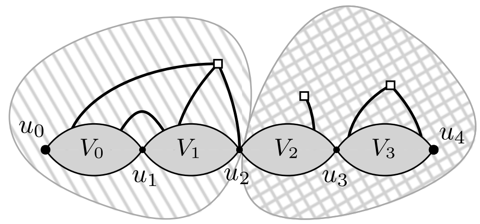

FIGURE 6. The schematic representation of a ($\pm$)-graph $G \oplus H$ in $\mathcal{C}^{\pm}_{u_0,u_4}(W \cup \vec{y}_{[3]})$ illustrates the factorization of the graph weight from Equation (5.2): the edges of H are explicitly depicted in the picture, while the core graph G is represented by its pivot decomposition $(u_0, V_0, \dots, u_4)$. The vertices $\vec{y}_{[3]}$ are depicted by squares, ordered from left to right. The first ($\pm$)-pivot point for the u_0–u_4 connection in $G \oplus H$ is u_2. Thus, the weight of the simple graph H factors into the weight of its subgraph induced by $\vec{W}_2' \cup \vec{y}_{[1]} = \{u_0, u_1, u_2\} \cup V_0 \cup V_1 \cup \{y_1\}$ (hatched on the left) and the weight of the subgraph induced by $\vec{W}_2'' \cup \vec{y}_{[3]\setminus[1]} = \{u_3, u_4\} \cup V_2 \cup V_3 \cup \{y_2, y_3\}$ (crosshatched on the right).

and that when replacing h_λ with $\bar{h}_\lambda$ in the right-hand side of (4.6), we get τ_λ instead of g_λ. We can split the sum $\bar{h}^{(m)}_\lambda\big(\vec{W}, \vec{y}_{[m]}\big) = h^{(m)}_\lambda\big(\vec{W}, \vec{y}_{[m]}\big) + f^{(m)}_\lambda\big(\vec{W}, \vec{y}_{[m]}\big)$, where $f^{(m)}_\lambda$ contains the sum over those graphs H such that $G \oplus H$ *does* have ($\pm$)-pivotal points with respect to the u_0–u_{k+1} connection. We set

$$f_\lambda\big(\vec{W}\big) := \sum_{m \geq 0} \frac{\lambda^m}{m!} \int_{\Lambda^m} f^{(m)}_\lambda\big(\vec{W}, \vec{y}_{[m]}\big)\, \mathrm{d}\vec{y}_{[m]}.$$

Assume now that u_j for $j \in [k]$ is the first pivotal point of $G \oplus H \in \mathcal{C}^{\pm}_{x_1,x_2}\big(W \cup \vec{y}_{[m]}\big)$. Furthermore, let $\vec{W}_j' := \big(u_0, V_0, \dots, u_j\big)$, let $\vec{W}_j'' := (u_j, V_j, \dots, u_{k+1})$, and let $y_{[s]}$ for $s \leq m$ be the points adjacent to $\vec{W}_j'$ (possibly after reordering the vertices). The weight of such a graph H then factors into the product of the weights of two graphs, namely the subgraphs of H induced by $\vec{W}_j' \cup \vec{y}_{[s]} \subset V(H)$ and by $\vec{W}_j'' \cup \vec{y}_{[m]\setminus[s]} \subset V(H)$; see Figure 6. That is,

$$\mathbf{w}(H) = \mathbf{w}\big(H\big[\vec{W}_j' \cup \vec{y}_{[s]}\big]\big)\mathbf{w}\big(H\big[\vec{W}_j'' \cup \vec{y}_{[m]\setminus[s]}\big]\big). \tag{5.2}$$

Moreover, we see that $H\big[\vec{W}_j' \cup \vec{y}_{[s]}\big] \oplus G\big[W_j'\big]$ does not contain ($\pm$)-pivot points (for the u_0–u_j connection) and $H\big[\vec{W}_j'' \cup \vec{y}_{[m]\setminus[s]}\big] \oplus G\big[W_j''\big]$ is in general just ($\pm$)-connected.

By partitioning over j, we thus obtain the decomposition

$$f_\lambda\big(\vec{W}\big) = \sum_{j=1}^{k} h_\lambda\big(\vec{W}_j'\big)\bar{h}_\lambda\big(\vec{W}_j''\big).$$

Since both h_λ and $\bar{h}_\lambda$ converge absolutely, so does f_λ, justifying all re-summations. Letting $x_1 = u_0$ and $x_2 = u_{k+1}$,

$$(\tau_\lambda - g_\lambda)(x_1, x_2) = \sum_{k \geq 1} \lambda^k \sum_{n_0,\dots,n_k \geq 0} \frac{\lambda^{\sum_{i=0}^k n_i}}{\prod_{i=0}^k n_i!} \int_{\Lambda^{k+\sum_{i=0}^k n_i}} \left(\prod_{i=0}^{k} \mathbb{P}\big(\Gamma_\varphi(\bar{V}_i) \in \mathcal{D}_{u_i,u_{i+1}}(\bar{V}_i)\big) \right)$$
$$\times \left(\sum_{j=1}^{k} h_\lambda\big(\vec{W}_j'\big)\bar{h}_\lambda\big(\vec{W}_j''\big) \right) \prod_{i=0}^{k} \mathrm{d}\vec{v}_{i,[n_i]}\, \mathrm{d}\vec{u}_{[k]}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

$$
\begin{aligned}
&= \sum_{j\geq 1, k\geq 0} \lambda^{j+k} \int_\Lambda \sum_{n_0,\ldots,n_{j+k}\geq 0} \frac{\lambda^{\sum_{i=0}^{j+k} n_i}}{\prod_{i=0}^{j+k} n_i!} \Bigg[\int_{\Lambda^{j-1+\sum_{i=0}^{j-1} n_i}} \prod_{i=0}^{j-1} \mathbb{P}\Big(\Gamma_\varphi(\bar{V}_i) \in \mathcal{D}_{u_i,u_{i+1}}(\bar{V}_i)\Big)\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \times h_\lambda\big(\vec{W}'_j\big) \prod_{i=0}^{j-1} \mathrm{d}\vec{v}_{i,[n_i]}\, \mathrm{d}\vec{u}_{[j-1]} \Bigg] \\
&\qquad\qquad \times \Bigg[\int_{\Lambda^{k+\sum_{i=j}^{j+k} n_i}} \prod_{i=j}^{j+k} \mathbb{P}\Big(\Gamma_\varphi(\bar{V}_i) \in \mathcal{D}_{u_i,u_{i+1}}(\bar{V}_i)\Big)\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \bar{h}_\lambda\big(\vec{W}''_j\big) \prod_{i=j}^{j+k} \mathrm{d}\vec{v}_{i,[n_i]}\, \mathrm{d}\vec{u}_{[j+1,j+k]} \Bigg] \mathrm{d}u_j \\
&= \lambda \int_\Lambda g_\lambda(x_1, u)\tau_\lambda(u, x_2)\,\mathrm{d}u.
\end{aligned}
$$

The re-summation with respect to j and k is justified as the resulting series converges for $\lambda < \lambda_*$ even when we put h_λ in absolute values. □

We can now extend the result of Proposition 5.1 to $\Lambda \nearrow \mathbb{R}^d$ and thus prove that the expansion (4.24) is indeed equal to the direct-connectedness function for $\lambda < \lambda_*$, finalizing the proof of our main result.

Proof of Theorem 1.1. We have

$$
\begin{aligned}
\tau_\lambda(x_1, x_2) &= \lim_{\Lambda\nearrow\mathbb{R}^d} \tau_\lambda^\Lambda(x_1, x_2) \\
&= \lim_{\Lambda\nearrow\mathbb{R}^d} g_\lambda^\Lambda(x_1, x_2) + \lambda \lim_{\Lambda\nearrow\mathbb{R}^d} \int_{\mathbb{R}^d} g_\lambda^\Lambda(x_1, x_3)\mathbb{1}_\Lambda(x_3)\tau_\lambda^\Lambda(x_3, x_2)\,\mathrm{d}x_3, \qquad (5.3)
\end{aligned}
$$

where the first equality holds by the continuity of probability measures along sequences of increasing events and the second one by Proposition 5.1.

Note that the integrand in (5.3) is bounded uniformly in Λ by

$$
C\tau_\lambda(x_3, x_2),
$$

where $C = \sup_{y\in\mathbb{R}^d} \sum_k \lambda^k \sigma_\lambda^{*(k+1)}(y)$ is a constant obtained in (4.29). Since τ_λ is integrable for all $\lambda < \lambda_c$, the theorem follows by dominated convergence. □

6. Discussion

6.1. Connections to percolation on Gibbs point processes

The Ornstein–Zernike equation gets its name from the seminal paper [20] and has since been a well-known formalism in liquid-state statistical mechanics. It relates the total correlation function to the direct correlation function and it naturally connects to power series expansions of these correlation functions (see [6, 23, 24]; the terminology is not the same in all of these references).

The correlation functions admit graphical expansions that also consist of connected graphs. It was observed [11] that a similar formalism can be formulated for the pair-connectedness function, and a key reference for this is [6]. The pair-connectedness function is deemed part

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

of the pair-correlation function. The connected graphs appearing in the expansion of the latter are referred to as 'mathematical clusters', and they correspond to our $(\pm)$-connected graphs. Isolating the $(+)$-connected components within these graphs yields the 'physical clusters', and the graphs in which x_1 and x_2 lie in the same physical cluster make up the expansion for $\tau_\lambda(x_1, x_2)$. In the following, we elaborate on this.

The percolation models considered in the physics literature are mostly based not on a PPP (Stell calls the Poisson setup *random percolation* [24]), but on a Gibbs point process (called *correlated percolation* in the language of Stell). (The denomination 'random percolation' for the Poisson setup feels quite misleading for probabilists; but it reflects language commonly adopted across physics, with 'random' understood as 'completely random' in the sense of *completely random measures* [13], a class comprising the PPP.)

To define the latter, consider a nonnegative pair potential $v : \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ and some finite volume Λ. Let $\mathbf{N}(\Lambda)$ be the set of finite counting measures on Λ and let $\mu \in \mathbf{N}(\Lambda)$. Then the energy of $\{x_1, \dots, x_n\}$ under the boundary condition μ is

$$H(\{x_1, \dots, x_n\} \mid \mu) = \sum_{1 \leq i < j \leq n} v\big(x_i - x_j\big) + \sum_{i=1}^{n} \sum_{y \in \mu} v(x_i - y).$$

Let $f : \mathbf{N}(\Lambda) \to \mathbb{R}$ be bounded. We define a probability measure as

$$\mathbb{E}_z[f] := \frac{1}{\Xi(z)} \sum_{n \geq 0} \frac{z^n}{n!} \int_{\Lambda^n} f(\{x_1, \dots, x_n\}) \mathrm{e}^{-H(\{x_1, \dots, x_n\})} \, \mathrm{d}\vec{x}_{[n]},$$

where the partition function $\Xi(z)$ is such that $\mathbb{E}_z[1] = 1$ and $z \in \mathbb{R}_{\geq 0}$ is called the activity. If we denote by η a random variable with law $\mathbb{E}_z$, then η is a point process. Note that we recover the homogeneous PPP with intensity $\lambda = z$ by setting $v \equiv 0$.

We can define the RCM ξ on this general point process, and we denote its probability measure by $\mathbb{P}_{z,\varphi}$. We furthermore define the (one-particle) density as

$$\rho_1(x) = z \mathbb{E}_z\big[\mathrm{e}^{-H(\{x\} \mid \eta)}\big] = \rho,$$

and we define the pair-correlation function as

$$\rho_2(x, y) = z^2 \mathbb{E}_z\big[\mathrm{e}^{-H(\{x,y\} \mid \eta)}\big].$$

Again, in the case of a homogeneous PPP with intensity $\lambda = z$, we have $\rho = z$ and $\rho_2 = z^2$. Defining the pair-connectedness function as

$$\tau_{z,\varphi}(x, y) := \mathbb{E}_{z,\varphi}\big[\mathrm{e}^{-H(\{x,y\} \mid \eta)} \mathbb{1}_{\{x \longleftrightarrow y \text{ in } \xi^{x,y}\}}\big],$$

we can decompose

$$\rho_2(x, y) = z^2 \tau_{z,\varphi}(x, y) + z^2 \mathbb{E}_{z,\varphi}\big[\mathrm{e}^{-H(\{x,y\} \mid \eta)} \mathbb{1}_{\{x \not\longleftrightarrow y \text{ in } \xi^{x,y}\}}\big].$$

In [6], Coniglio *et al.* define the pair-connectedness function as $\tilde{\tau}_{z,\varphi} = \big(z^2/\rho^2\big) \tau_{z,\varphi}$.

The function $\tilde{\tau}_{z,\varphi}$ has a density expansion (note that $\tau_{z,\varphi}$ is better suited for activity expansions) that can be found in [6, Equation (12)], which can be obtained from the density expansion of the pair-correlation function: the latter is obtained by expanding the Mayer f-functions $f(x, y) = \mathrm{e}^{-v(x,y)} - 1$ in the partition function, which is the starting point of a cluster

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

expansion. Splitting the Mayer f-function as $f = f^+ + f^*$ with $f^+ = \mathrm{e}^{-v(x,y)}\varphi(x-y)$ and executing the same expansion for the correlation function 'doubles' every edge into a $(+)$-edge and a $(*)$-edge. Only summing over graphs in which x and y are connected by $(+)$-edges yields the pair-connectedness function.

In general, the graphs appearing in the density expansion are a subset of those in the activity expansion, namely the ones without articulation points (articulation points were defined after Proposition 3.1). In the case of a homogeneous PPP, we have $\lambda = z = \rho$, and so activity and density expansion coincide (and the graphs with articulation points cancel out). Moreover, $f^+(x, y) = -f^*(x, y) = \varphi(x-y)$, and the graphs summed over in the expansion become the $(\pm)$-graphs, yielding the expansion (3.1) for τ_λ.

It is an interesting question which ideas of this paper can be generalized to RCMs based on Gibbs point processes. While some aspects generalize without much effort, the crucial difference lies in the fact that the weight of graphs showing up in expansions for Gibbs point processes also encodes the pair interaction induced by the potential v. To recover probabilistic interpretations for terms after performing re-summations and bounds is therefore much more delicate.

6.2. Connections to Last and Ziesche

In [15], Last and Ziesche use a Margulis–Russo-type formula to prove analyticity of τ_λ in presumably the whole subcritical regime. Moreover, they show the existence of some $\lambda_0 > 0$ (which is not quantified) such that both τ_λ and g_λ have an absolutely convergent graphical expansion in $[0, \lambda_0)$ that seems closely related to the ones discussed here. We want to illustrate how to relate the respective expressions.

The two-point function. Last and Ziesche show that $\tau_\lambda(x_1, x_2)$ is equal to

$$\sum_{n\geq 0}\frac{\lambda^n}{n!}\int\sum_{J\subset[3,n+2]}(-1)^{n-|J|}\mathbb{P}\big(x_1 \longleftrightarrow x_2 \text{ in } \Gamma_\varphi\big(\vec{x}_{J\cup\{1,2\}}\big),\ \Gamma_\varphi\big(\vec{x}_{[n+2]}\big)\text{ is connected}\big)\,\mathrm{d}\vec{x}_{[3,n+2]}. \tag{6.1}$$

We show that the above integrand is the same as the one in (3.1). We can rewrite the one in (6.1) as

$$\mathbb{E}\left[\mathbb{1}_{\left\{\Gamma_\varphi\left(\vec{x}_{[n+2]}\right)\text{ is connected}\right\}}\sum_{J\subset[n+2]}(-1)^{n-|J|}\mathbb{1}_{\left\{x_1\longleftrightarrow x_2\text{ in }\Gamma_\varphi(\vec{x}_J)\right\}}\right]. \tag{6.2}$$

Note that now, any nonvanishing J needs to contain $\{1, 2\}$. We are now going to observe some cancellations. For a fixed graph $G \in \mathcal{C}\big(\vec{x}_{[n+2]}\big)$,

$$\begin{aligned}\sum_{J\subseteq[n+2]}(-1)^{n-|J|}\mathbb{1}_{\{x_1\longleftrightarrow x_2\text{ in }G[\vec{x}_J]\}} &= \sum_{I,J\subseteq[n+2]}(-1)^{n-|J|}\mathbb{1}_{\{\{1,2\}\subseteq I\subseteq J\}}\mathbb{1}_{\{\mathscr{C}(x_1,G[\vec{x}_J])=\vec{x}_I\}}\\ &= \sum_{I,J\subseteq[n+2]}(-1)^{n-|J|}\mathbb{1}_{\{\{1,2\}\subseteq I\subseteq J\}}\mathbb{1}_{\{G[\vec{x}_I]\text{ connected}\}}\mathbb{1}_{\{\forall j\in J\setminus I: x_j\nsim\vec{x}_I\}}\\ &= \sum_{\{1,2\}\subseteq I\subseteq[n+2]}(-1)^{n-|I|}\mathbb{1}_{\{G[\vec{x}_I]\text{ connected}\}}\\ &\qquad\times\sum_{J\subseteq[n+2]\setminus I}(-1)^{|J|}\mathbb{1}_{\{\forall j\in J: x_j\nsim\vec{x}_I\}}.\end{aligned} \tag{6.3}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Note that for given G and I, defining $\mathcal{I}(G,I)=\{j\in[n+2]\setminus I : x_j \nsim \vec{x}_I\}$, we can rewrite

$$\sum_{J\subseteq[n+2]\setminus I}(-1)^{|J|}\mathbb{1}_{\{\forall j\in J: x_j\nsim\vec{x}_I\}}=\sum_{J\subseteq\mathcal{I}(G,I)}(-1)^{|J|}. \tag{6.4}$$

The only case for which (6.4) does not vanish is when $\mathcal{I}(G,I)=\varnothing$. We can therefore rewrite (6.3) as

$$\sum_{J\subseteq[n+2]}(-1)^{n-|J|}\mathbb{1}_{\left\{x_1\longleftrightarrow x_2 \text{ in } G[\vec{x}_J]\right\}}=\sum_{\{1,2\}\subseteq I\subseteq[n+2]}(-1)^{n-|I|}\mathbb{1}_{\{G[\vec{x}_I]\text{ connected}\}}\mathbb{1}_{\left\{\forall j\in[n+2]\setminus I: x_j\sim\vec{x}_I\right\}},$$

and so (6.2) becomes

$$\begin{aligned}&\sum_{\{1,2\}\subseteq I\subseteq[n+2]}(-1)^{n-|I|}\mathbb{P}\big(\Gamma_\varphi(\vec{x}_I)\text{ is connected},\ x_j\sim\vec{x}_I\forall j\in[n+2]\setminus I\big)\\&=\sum_{\{1,2\}\subseteq I\subseteq[n+2]}\mathbb{P}\big(\Gamma_\varphi(\vec{x}_I)\text{ is connected}\big)\prod_{j\in[n+2]\setminus I}\left[\prod_{i\in I}\big(1-\varphi\big(x_i-x_j\big)\big)-1\right]\\&=\sum_{I\subseteq[n+2]}\sum_{G}\mathbf{w}^{\pm}(G).\end{aligned}$$

In the last line, summation is over the same set of graphs as in (3.6), with the additional restriction that $V\big(G^+\big)=I$. Resolving the partition over I gives that (6.1) is equal to (3.6).

The direct-connectedness function. In [15, Theorem 5.1], it is shown that there exists λ_0 such that for $\lambda\in[0,\lambda_0)$,

$$\begin{aligned}g_\lambda(x_1,x_2)=\sum_{n\geq 0}\frac{\lambda^n}{n!}\int&\sum_{G\in\mathcal{D}_{x_1,x_2}\left(\vec{x}_{[n+2]}\right)}\prod_{e\in E(G)}\varphi(e)\\&\times\sum_{\substack{H\in\mathcal{C}\left(\vec{x}_{[n+2]}\right):\\ H\subseteq G}}\ \sum_{\substack{J\subseteq[n+2]:\\ x_1\longleftrightarrow x_2\text{ in }H[\vec{x}_J]}}(-1)^{n-|J|+|E(G)\setminus E(H)|}\,\mathrm{d}\vec{x}_{[3,n+2]}.\end{aligned}\tag{6.5}$$

We show that the integrand in (6.5) is equal to the one in (4.1). With the calculations (6.3) and (6.4) performed for the two-point function, letting $I^c=[n+2]\setminus I$, we have

$$\sum_{J\subseteq[n+2]}(-1)^{n-|J|}\mathbb{1}_{\left\{x_1\longleftrightarrow x_2\text{ in }H[\vec{x}_J]\right\}}=\sum_{\{1,2\}\subseteq I\subseteq[n+2]}(-1)^{n-|I|}\mathbb{1}_{\{H[\vec{x}_I]\text{ is connected}\}}\mathbb{1}_{\left\{\forall j\in I^c: x_j\sim\vec{x}_I\text{ in }H\right\}}.$$

The two indicators imply that H is connected, and so

$$\begin{aligned}&\sum_{\substack{H\in\mathcal{C}\left(\vec{x}_{[n+2]}\right):\\ H\subseteq G}}\ \sum_{\substack{J\subseteq[n+2]:\\ x_1\longleftrightarrow x_2\text{ in }H[\vec{x}_J]}}(-1)^{n-|J|+|E(G)\setminus E(H)|}\\&=\sum_{\{1,2\}\subseteq I\subseteq[n+2]}\sum_{H\subseteq G}(-1)^{n-|I|+|E(G)\setminus E(H)|}\mathbb{1}_{\{H[\vec{x}_I]\text{ is connected}\}}\mathbb{1}_{\left\{\forall j\in I^c: x_j\sim\vec{x}_I\text{ in }H\right\}}\\&=\sum_{\{1,2\}\subseteq I\subseteq[n+2]}\ \sum_{\substack{H'\in\mathcal{C}(\vec{x}_I):\\ H'\subseteq G}}(-1)^{n-|I|+|E(G)\setminus E(H')|}\sum_{\substack{F\subset E(G)\cap\left((I\times I^c)\cup\binom{I^c}{2}\right):\\ \forall j\in I^c: F\cap(I\times\{j\})\neq\varnothing}}(-1)^{|F|}.\end{aligned}\tag{6.6}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

Note that for the second identity in (6.6), we split the edges of H into those contained in H' (the subgraph induced by I) and the remaining ones, called F.

When $E(G)\cap\binom{I^c}{2}\neq\varnothing$, the sum over F vanishes. Hence, the sum over I can be reduced to those I such that $G[I^c]$ contains no edges. For such sets I, we have

$$\sum_{\substack{F\subset E(G)\cap(I\times I^c):\\ \forall j\in I^c: F\cap(I\times\{j\})\neq\varnothing}}(-1)^{|F|}=\prod_{j\in I^c}\sum_{\varnothing\neq F_j\subseteq E(G)\cap(I\times\{j\})}(-1)^{|F_j|}=\prod_{j\in I^c}(-1)=(-1)^{n-|I|}. \tag{6.7}$$

If we insert (6.7) back into (6.6), the two factors $(-1)^{n-|I|}$ cancel out, and so

$$\begin{aligned}
&\sum_{\substack{H\in\mathcal{C}\left(\vec{x}_{[n+2]}\right):\\ H\subseteq G}}\ \sum_{\substack{J\subseteq[n+2]:\\ x_1\longleftrightarrow x_2\text{ in }H[\vec{x}_J]}}(-1)^{n-|J|+|E(G)\setminus E(H)|}\\
&=\sum_{\substack{\{1,2\}\subseteq I\subseteq[n+2]:\\ E(G[I^c])=\varnothing}}\ \sum_{\substack{H\in\mathcal{C}(\vec{x}_I):\\ H\subseteq G}}(-1)^{|E(G)\setminus E(H)|}\\
&=\sum_{H\in\mathcal{G}\left(\vec{x}_{[n+2]}\right):G\rhd H}\mathbb{1}_{\{\{1,2\}\subseteq V(H)\}}(-1)^{|E(G)\setminus E(H)|},
\end{aligned}\tag{6.8}$$

where $G\rhd H$ means that $E(H)\subseteq E(G)$, the subgraph of H induced by the vertices incident to at least one edge (call this set $V_{\geq 1}(H)$) is connected, and the subgraph of G induced by $[n+2]\setminus V_{\geq 1}(H)$ contains no edges.

With the identity (6.8), and letting $X=\vec{x}_{[n+2]}$, the integrand of (6.5) is equal to

$$\begin{aligned}
&\sum_{\substack{H\in\mathcal{G}(X):\\ \{x_1,x_2\}\subseteq V_{\geq 1}(H),\\ H[V_{\geq 1}(H)]\text{ connected}}}\ \prod_{e\in E(H)}\varphi(e)\sum_{\substack{F\subseteq\binom{X}{2}\setminus E(H):\\ \forall e\in F: e\cap V_{\geq 1}(H)\neq\varnothing,\\ (X,F\cup E(H))\in\mathcal{D}_{x_1,x_2}(X)}}(-1)^{|F|}\prod_{e\in F}\varphi(e)\\
&=\sum_{\substack{H\in\mathcal{G}(X):\\ x_1\longleftrightarrow x_2}}\ \prod_{e\in E(H)}\varphi(e)\sum_{\substack{F\subseteq\binom{X}{2}\setminus E(H):\\ (X,F\cup E(H))\in\mathcal{D}_{x_1,x_2}(X)}}(-1)^{|F|}\prod_{e\in F}\varphi(e)\\
&=\sum_{\substack{C\in\mathcal{D}^{\pm}_{x_1,x_2}(X):\\ x_1\overset{+}{\longleftrightarrow}x_2}}\mathbf{w}^{\pm}(G).
\end{aligned}\tag{6.9}$$

The argument for the first identity in (6.9) is the same as for the identity of (3.6) and (3.7).

6.3. Connections to the lace expansion

Both the graphical power series expansions and the lace expansion provide expressions for the direct-connectedness function. In this section, we show how to get from one to the other. Note that the statements to follow hold for sufficiently small intensities and *cannot* replace the lace expansion, which works all the way up to λ_c. The emphasis of this section is on the qualitative nature of the results.

We first summarize some results of [10], where the lace expansion is applied to the RCM. We keep some of the definitions brief and informal, and we refer to [10] for the detailed definitions in these cases.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

On the lace expansion. In [10], among other things, the OZE is proved for τ_λ in *high dimension* (and for certain classes of connection functions φ; see [10, Section 1.2]). In particular, it is shown that

$$g_\lambda(x) = \varphi(x) + \Pi_\lambda(x),$$

with $\Pi_\lambda(x) = \sum_{n\geq 0}(-1)^n \Pi_\lambda^{(n)}(x)$. The functions $\Pi_\lambda^{(n)}$ are called the *lace-expansion coefficients*; they are nonnegative and have a quite involved probabilistic interpretation. To briefly define them, let $\{x \overset{A}{\longleftrightarrow} y \text{ in } \xi^{x,y}\}$ be the event that $x \longleftrightarrow y$ in $\xi^{x,y}$, but x is no longer connected to y in an A-thinning of η^y. Informally, every point $z \in \eta$ survives an A-thinning with probability $\prod_{y\in A}(1-\varphi(z-y))$. See [10, Definition 3.2] for a formal definition. Letting

$$E\big(x, y; A, \xi^{x,y}\big) = \big\{x \overset{A}{\longleftrightarrow} y \text{ in } \xi^{x,y}\big\} \cap \big\{\nexists w \in \mathsf{Piv}\big(x, y; \xi^{x,y}\big) : x \overset{A}{\longleftrightarrow} w \text{ in } \xi^{x}\big\},$$

we introduce a sequence $\xi_0, \ldots, \xi_n$ of independent RCMs and define

$$\begin{aligned}\Pi_\lambda^{(0)}(x) &:= \sigma_\lambda(x) - \varphi(x),\\ \Pi_\lambda^{(n)}(u_n) &:= \lambda^n \int \mathbb{P}_\lambda\left(\{\mathbf{0} \Longleftrightarrow u_0 \text{ in } \xi_0^{\mathbf{0},u_0}\} \cap \bigcap_{i=1}^n E\Big(u_{i-1}, u_i; \mathscr{C}\big(u_{i-2}, \xi_{i-1}^{u_{i-2}}\big), \xi_i^{u_{i-1},u_i}\Big)\right) \mathrm{d}\vec{u}_{[0,n-1]}\end{aligned} \tag{6.10}$$

for $n \geq 1$ (with $u_{-1} = \mathbf{0}$). The method of proof is called the *lace expansion*, a perturbative technique in which one first proves via induction that

$$\tau_\lambda(x) = \varphi(x) + \sum_{m=0}^{n}(-1)^m \Pi_\lambda^{(m)}(x) + \lambda\left(\left(\varphi + \sum_{m=0}^{n}(-1)^m \Pi_\lambda^{(m)}\right) * \tau_\lambda\right)(x) + R_{\lambda,n}(x) \tag{6.11}$$

for $n \in \mathbb{N}_0$ and some remainder term $R_{\lambda,n}$ (see [10, Definition 3.7]), and then shows that the partial sum converges to $\Pi_\lambda = g_\lambda - \varphi$ and that $R_{\lambda,n} \to 0$ as $n \to \infty$.

The lace expansion was first devised for self-avoiding walks by Brydges and Spencer [3] and takes some inspiration from cluster expansions. It was later applied to percolation (specifically, bond percolation on $\mathbb{Z}^d$) by Hara and Slade [8]. While the name stems from laces that appear in the pictorial representation in [3], laces are absent in the representation for percolation models.

We show that we can rewrite $\Pi_\lambda^{(n)}$ in terms of graphs that are associated to a lace of size n. More generally, rewriting $\Pi_\lambda^{(n)}$ should serve as a bridge between the graphical expansions for g_λ that are well known in the physics literature, and the expression for g_λ in terms of lace-expansion coefficients.

The big advantage in the lace expansion lies in the probabilistic nature of all the terms that appear, allowing one to bound most of the integrals that appear by the expected cluster size, which is finite for $\lambda < \lambda_c$. The downside is the absence of a direct expression for g_λ and thus a direct proof of the OZE, which is only obtained after performing the $n \to \infty$ limit in (6.11).

We now show how to re-sum the graphical expansion for τ_λ and how to obtain the lace-expansion coefficients by appropriate grouping of terms.

Building the connection. For $x, y \in X$, let $\tilde{\mathcal{C}}^{\pm}_{x,y}(X) \subset \mathcal{C}^{\pm}(X)$ be the set of graphs in $\mathcal{C}^{\pm}(X)$ such that G^+ is connected and contains $\{x, y\}$, and $E(G[V \setminus V^+]) = \varnothing$. Hence, all $(-)$-edges are

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

incident to at least one vertex in $V\big(G^{+}\big)$. This is exactly the set of graphs summed over in (3.6). Indeed,

$$\tau_\lambda(x_1, x_2) = \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{G\in \tilde{\mathcal{C}}^{\pm}_{x_1,x_2}\left(\vec{x}_{[n+2]}\right)} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec{x}_{[3,n+2]}. \tag{6.12}$$

If we define $\tilde{\mathcal{D}}^{\pm}_{x,y}(X) := \mathcal{D}^{\pm}_{x,y}(X) \cap \tilde{\mathcal{C}}^{\pm}_{x,y}(X)$, we can express $g_\lambda(x_1, x_2)$ by replacing the graphs summed over in (6.12) by $\tilde{\mathcal{D}}^{\pm}_{x,y}\big(\vec{x}_{[n+2]}\big)$.

We are going to recycle some notation from Section 4. We split G into its core G_{core} and its shell H, so that

$$\mathsf{PD}^{+}(x, y, G_{\mathrm{core}}) = \mathsf{PD}^{\pm}(x, y, G_{\mathrm{core}}) = (u_0, V_0, u_1, \ldots, u_k, V_k, u_{k+1})$$

for some k (where $u_0 = x$ and $u_{k+1} = y$). We also recall that G 'contains' a skeleton (see Definition 4.3), a graph on $[k+1]_0$.

Definition 6.1. (*The minimal lace.*) Let G be a graph with core G_{core} and shell H; let $\vec{W} = (u_0, V_0 \ldots, u_{k+1})$ for $k \in \mathbb{N}$ be its $(+)$-pivot decomposition. We define the minimal lace $L_{\min}(x, y; G)$ as the lace with the following properties:

- L (having bonds $\alpha_i\beta_i$ with $i \in [m]$ for some $m \in \mathbb{N}$) is contained as a subgraph in the skeleton $\hat{H}$;
- for every $i \in [m]$, among all the bonds $\alpha\beta$ in $\hat{H}$ satisfying $\alpha < \beta_{i-1}$, the bond $\alpha_i\beta_i$ maximizes the value of β. For $i = 1$, we take $\beta_0 = 1$.

If $\mathsf{Piv}^{+}(x, y; G) = \varnothing$, we say that G has a minimal lace of size 0.

In other words, the first stitch $0\beta_1$ maximizes the value of β_1 among all stitches starting at 0, the second stitch has a maximal value of β_2 among the stitches with $1 \leq \alpha_2 < \beta_1$, and so on.

As a side remark, it is worth noting that the minimal laces offer an alternative way of partitioning the set of all shell graphs by mapping every shell graph H onto its minimal lace. This gives a standard procedure used in lace expansion for self-avoiding walks; performing it 'backwards' yields precisely the mapping described below Definition 4.4.

With the notion of minimal laces, we partition

$$g_\lambda(x_1, x_2) = \sum_{m\geq 0} \pi^{(m)}_\lambda(x_1, x_2), \tag{6.13}$$

where

$$\pi^{(m)}_\lambda(x_1, x_2) := \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{\substack{G\in \tilde{\mathcal{D}}^{\pm}_{x_1,x_2}\left(\vec{x}_{[n+2]}\right):\\ \|L_{\min}\|=m}} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec{x}_{[3,n+2]}. \tag{6.14}$$

We also set $\pi^{(m)}_\lambda(x) = \pi^{(m)}_\lambda(\mathbf{0}, x)$.

We strongly expect that the (pointwise) absolute convergence of the power series on the right-hand side of (6.14) holds (at least) in the domain of absolute convergence of the physicists' expansion (4.1) and thus, as already discussed, for sufficiently small intensities $\lambda > 0$.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

However, a proof would go beyond the scope of the discussion here; therefore we formulate the absolute convergence of $\pi_\lambda^{(m)}$ (in the above sense) as an assumption for the following result (Lemma 6.1).

Assumption 6.1. *There exists $0 < \lambda_\bigstar \le \lambda_c$ such that the right-hand side of (6.14) is (pointwise) absolutely convergent for all $m \in \mathbb{N}$ and $\lambda < \lambda_\bigstar$.*

Under Assumption 6.1, we show that the coefficients defined in (6.14) are basically identical to the lace-expansion coefficients introduced in (6.10).

Lemma 6.1. (Identity for the lace-expansion coefficients.) *Let $m \ge 1$ and let $\lambda < \lambda_\bigstar$. Then*

$$\Pi_\lambda^{(0)}(x) = \pi_\lambda^{(0)}(\mathbf{0}, x) - \varphi(x),$$
$$(-1)^m \Pi_\lambda^{(m)}(x) = \pi_\lambda^{(m)}(\mathbf{0}, x).$$

As a side note, since $\Pi_\lambda^{(m)}$ is nonnegative, Lemma 6.1 shows that the sign of $\pi_\lambda^{(m)}$ alternates, which is far from obvious from the definition in (6.14).

Next, we prove an approximate version of the OZE in analogy to [10, Proposition 3.8]. Clearly, Lemma 6.2 follows immediately from the latter via Lemma ; however, we want to present a short independent proof on the level of formal power series, which we consider instructive for the understanding of the underlying combinatorics. We emphasize that the proof presented here treats the claim of Lemma 6.2 as an identity between formal power series; in particular, we do not concern ourselves with absolute convergence of the power series appearing in (6.20) and in (6.21).

Lemma 6.2. (The lace expansion in terms of ($\pm$)-graph coefficients.) *Let $m \in \mathbb{N}_0$, let $\lambda < \lambda_\bigstar$, and set $\pi_{\lambda,m}(x) := \sum_{i=0}^m \pi_\lambda^{(i)}(\mathbf{0}, x)$. Then*

$$\tau_\lambda(x) = \pi_{\lambda,m}(x) + \big(\pi_{\lambda,m} * \tau_\lambda\big)(x) + R_{\lambda,m}(x),$$

where $R_{\lambda,m}$ is defined in [10, Definition 3.7].

Before carrying out the proof of Lemma 6.1, we define

$$\bar\varphi(A, B) = \prod_{a\in A}\prod_{b\in B}(1 - \varphi(a-b))$$

and $\bar\varphi(a, B) = \bar\varphi(\{a\}, B)$. Now, observe that, given a set $A \subset \mathbb{R}^d$ and an RCM event F,

$$\sum_{n\ge 0}\frac{\lambda^n}{n!}\int_{(\mathbb{R}^d)^n}\bar\varphi\big(A, \vec v_{[n+2]}\big)\sum_{\substack{G\in\tilde{\mathcal{C}}^{\pm}_{v_1,v_2}(\vec v_{[n+2]}):\\ G^+\in F}}\mathbf{w}^{\pm}(G)\,\mathrm{d}\vec v_{[3,n+2]} = \mathbb{P}_\lambda\big(\xi\big(\eta^{v_1,v_2}_{\langle A\rangle}\big)\in F\big), \tag{6.15}$$

where $\xi(\eta)$ is the RCM on the basis of the point process η and $\eta^{v}_{\langle A\rangle}$ is an A-thinning of η^v (the usual PPP of intensity λ and added point v). In particular, v may be thinned out as well. We remark that $\eta_{\langle A\rangle}$ has the same distribution as a PPP of intensity $\lambda\bar\varphi(A, \cdot)$.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

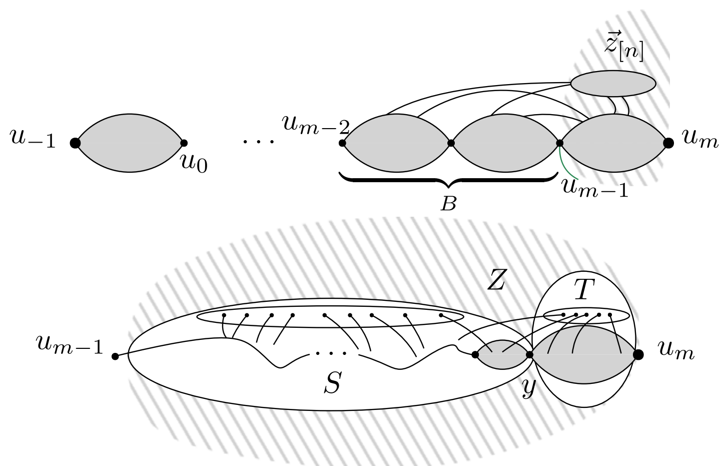

FIGURE 7. Illustration for the proof of Lemma 6.1. On the left, we see an example graph $G \in \mathcal{B}$; the grey bags on the bottom represent $\mathsf{PD}^+(u_{-1}, u_m, G)$ (note that there can be pivotal points within a grey bag). The minimal lace $L_{\min}$ is not depicted; however, note that $p_m \in B$. On the right, we see a schematic zoom into $G[Z]$, where $Z = \vec{z}_{[n]}$, together with the partition $Z = S \cup T$.

Proof of Lemma 6.1. The statement for $m = 0$ is clear. For $m > 0$, we can rewrite $\pi_\lambda^{(m)}$ as

$$\pi_\lambda^{(m)}(u_{-1}, u_m) = \lambda^m \int_{(\mathbb{R}^d)^m} \sum_{k,n \geq 0} \frac{\lambda^{k+n}}{k!n!} \int_{(\mathbb{R}^d)^{k+n}} \sum_{G \in \mathcal{B}} \mathbf{w}^\pm(G) \,\mathrm{d}\big(\vec{u}_{[0,m-1]}, \vec{x}_{[k]}, \vec{z}_{[n]}\big), \tag{6.16}$$

where $\mathcal{B} \subseteq \tilde{\mathcal{D}}^\pm_{u_{-1},u_m}\big(\vec{u}_{[-1,m]} \cup \vec{x}_{[k]} \cup \vec{z}_{[n]}\big)$ are the graphs such that

- u_0 is the first pivotal point in $\mathsf{Piv}^+(u_{-1}, u_m; G)$ (i.e., $\mathsf{ord}(u_0) = 2$);
- $\vec{u}_{[0,m-1]} \subseteq \mathsf{Piv}^+(u_{-1}, u_m; G)$ and $u_{i-1} \prec u_i$;
- there are points $p_2, \dots, p_m$ such that $L_{\min} \mathrel{\hat{=}} \{(u_{-1}, u_1), (p_2, u_2), \dots, (p_m, u_m)\}$;
- $\vec{z}_{[n]}$ are those vertices $z \notin \{u_{m-1}, u_m\}$ in G such that $\{z\} \cup N(z)$ contains at least one vertex y of order $y \succ u_{m-1}$.

Given a graph $G \in \mathcal{B}$, let B denote the set of points x in $V\big(G^+\big)$ with $u_{m-2} \preccurlyeq x \prec u_{m-1}$. See Figure 7 for an illustration of such a graph G. We integrate out the points $\vec{z}$ first and claim that their contribution to (6.16) is

$$\lambda \sum_{n \geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{H \in \mathcal{B}^\star} \mathbf{w}^\pm(H) \,\mathrm{d}\vec{z}_{[n]} = -\lambda \mathbb{P}_\lambda\big(E\big(u_{m-1}, u_m; B, \xi^{u_{m-1},u_m}\big)\big), \tag{6.17}$$

where every $H \in \mathcal{B}^\star$ is the subgraph of some $G \in \mathcal{B}$ and has vertex set $B \cup \{u_{m-1}, u_m\} \cup \vec{z}_{[n]}$ and precisely those edges in G that have at least one endpoint in $\{u_m\} \cup \vec{z}_{[n]}$.

We let y be the last pivotal point in $V\big(G^+\big)$, that is, $\mathsf{ord}(y) = \mathsf{ord}(u_m) - 2$. We write $Z = \vec{z}_{[n]}$ and split Z once more into those vertices 'in front of' and 'behind' y; that is, $Z = S \cup T$, where T are the points in G^+ of order $\mathsf{ord}(u_m) - 1$ together with the points in $V \setminus V\big(G^+\big)$ that are adjacent to the former, and $S = Z \setminus T$. Possibly $y = u_{m-1}$, in which case $S = \varnothing$. See Figure 7 for an illustration of this split of the vertices in Z.

Note that there are no restrictions on the $(-)$-edges between B and $S \cup \{y\}$, whereas there must be at least one $(-)$-edge between B and $T \cup \{u_m\}$. There are no restrictions on the

$(-)$-edges between $\{u_{m-1}\} \cup S \cap V\big(G^+\big)$ and $T \cup \{u_m\}$, whereas there cannot be any $(-)$-edges between $S \setminus V\big(G^+\big)$ and $T \cup \{u_m\}$. By distinguishing whether or not $S = \varnothing$, we find that the left-hand side of (6.17) is equal to

$$\begin{aligned}
&\sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \big(\bar\varphi(B, \vec z_{[n]} \cup \{u_m\}) - 1\big) \sum_{\substack{G \in \tilde{\mathcal{C}}^{\pm}_{u_{m-1},u_m}\left(\{u_{m-1},u_m\}\cup\vec z_{[n]}\right):\\ u_{m-1} \overset{+}{\Longleftrightarrow} u_m}} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec z_{[n]} \\
&+ \lambda \sum_{k\geq 0} \frac{\lambda^k}{k!} \int_{(\mathbb{R}^d)^{k+1}} \bar\varphi\big(B, \vec s_{[k]} \cup \{y\}\big) \sum_{\substack{H \in \tilde{\mathcal{C}}^{\pm}_{u_{m-1},y}\left(\{u_{m-1},y\}\cup\vec s_{[k]}\right):\\ u_{m-1} \overset{+}{\longleftrightarrow} y}} \mathbf{w}^{\pm}(H) \\
&\qquad \times \Bigg(\sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \big(\bar\varphi(B, \vec t_{[n]} \cup \{u_m\}) - 1\big) \bar\varphi\big(V^+(H) \setminus \{y\}, \vec t_{[n]} \cup \{u_m\}\big) \\
&\qquad\qquad \sum_{\substack{G \in \tilde{\mathcal{C}}^{\pm}_{y,u_m}(\{y,u_m\}\cup\vec t_{[n]}):\\ y \overset{+}{\Longleftrightarrow} u_m}} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec t_{[n]} \Bigg) \mathrm{d}\vec s_{[k]}\, \mathrm{d}y \\
&= \mathbb{P}_\lambda\big(u_{m-1} \Longleftrightarrow u_m \text{ in } \xi\big(\{u_{m-1}\} \cup \eta^{u_m}_{\langle B\rangle}\big)\big) - \mathbb{P}_\lambda\big(u_{m-1} \Longleftrightarrow u_m \text{ in } \xi^{u_{m-1},u_m}\big) \\
&\quad + \lambda \int_{\mathbb{R}^d} \mathbb{E}_\lambda\Big[\mathbb{1}_{\left\{u_{m-1} \longleftrightarrow y \text{ in } \xi\left(\{u_{m-1}\}\cup\eta^{y}_{\langle B\rangle}\right)\right\}} \\
&\qquad\qquad \times \Big(\mathbb{P}_\lambda\Big(y \Longleftrightarrow u_m \text{ in } \xi\Big(\{y\} \cup \eta^{u_m}_{\langle B\cup\mathscr{C}\rangle}\Big)\Big) - \mathbb{P}_\lambda\Big(y \Longleftrightarrow u_m \text{ in } \xi\Big(\{y\} \cup \eta^{u_m}_{\langle\mathscr{C}\rangle}\Big)\Big)\Big)\Big]\, \mathrm{d}y,
\end{aligned} \tag{6.18}$$

where we abbreviate $\mathscr{C} = \mathscr{C}(u_{m-1}, \xi^{u_{m-1}})$. Note that the inner probabilities are conditional on the random variable $\mathscr{C}$. We now resolve the integral over y by use of the Mecke equation and incorporate the first two summands as the case $y = u_{m-1}$. With this, (6.18) becomes

$$\begin{aligned}
&\mathbb{E}_\lambda\Bigg[\sum_{y\in\eta^{u_{m-1}}} \mathbb{1}_{\left\{u_{m-1} \longleftrightarrow y \text{ in } \xi\left(\{u_{m-1}\}\cup\eta^{y}_{\langle B\rangle}\right)\right\}} \mathbb{1}_{\left\{y \Longleftrightarrow u_m \text{ in } \xi\left(\{y\}\cup(\eta^{u_m}\setminus\mathscr{C}')_{\langle B\rangle}\right)\right\}}\Bigg] \\
&-\mathbb{E}_\lambda\Bigg[\sum_{y\in\eta^{u_{m-1}}} \mathbb{1}_{\left\{u_{m-1} \longleftrightarrow y \text{ in } \xi^{u_{m-1}}\right\}} \mathbb{1}_{\left\{y \Longleftrightarrow u_m \text{ in } \xi\left(\eta^{u_m}\setminus\mathscr{C}'\right)\right\}}\Bigg],
\end{aligned} \tag{6.19}$$

where $\mathscr{C}' = \mathscr{C}\big(u_{m-1}, \xi\big(\eta^{u_{m-1}} \setminus \{y\}\big)\big)$. But both terms in (6.19) are simply a partition over the last pivotal point for the connection between u_{m-1} and u_m, and so (6.19) equals

$$\mathbb{P}_\lambda\Big(u_{m-1} \longleftrightarrow u_m \text{ in } \xi\Big(\{u_{m-1}\} \cup \eta^{u_m}_{\langle B\rangle}\Big)\Big) - \tau_\lambda(u_m - u_{m-1}) = -\mathbb{P}_\lambda\big(E\big(u_{m-1}, u_m; B, \xi^{u_{m-1},u_m}\big)\big),$$

proving (6.17). Lemma 6.1 can now be proven by iteratively applying (6.17). □

Proof of Lemma 6.2. For $m \in \mathbb{N}_0$, we can write

$$\tau_\lambda(x_1, x_2) = \sum_{l=0}^{m} \pi_\lambda^{(l)}(x_1, x_2) + \sum_{n\geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{G\in\mathcal{A}} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec x_{[3,n+2]}, \tag{6.20}$$

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

where $\mathcal{A}$ is the set of graphs $G \in \tilde{\mathcal{C}}^{\pm}_{x_1,x_2}(\vec{x}_{[n+2]}) \setminus \tilde{\mathcal{D}}^{\pm}_{x_1,x_2}(\vec{x}_{[n+2]})$ together with the graphs $G \in \tilde{\mathcal{D}}^{\pm}_{x_1,x_2}(\vec{x}_{[n+2]})$ where $\|L_{\min}\| > m$. Note that if $G \in \mathcal{A}$, then $\mathsf{Piv}^+(x_1, x_2; G) \neq \varnothing$.

For $G \in \mathcal{A}$ and $u \in \mathsf{Piv}^+(x_1, x_2; G)$, define

$$V^{\preccurlyeq}(u) := \{y \in V(G^+) : y \preccurlyeq u\} \cup \{y \in V(G) \setminus V(G^+) : \exists z \in N(y) \cap V(G^+) \textit{ with } z \prec u\},$$

that is, all the core vertices of order at most that of u together with the shell vertices adjacent to at least one vertex of strictly smaller order than u. Next, let $u^{\text{cut}} = u^{\text{cut}}(x_1, x_2; G)$ be the vertex in $\mathsf{Piv}^+(x_1, x_2; G)$ such that

$$E\big(V^{\preccurlyeq}(u^{\text{cut}}) \setminus \{u^{\text{cut}}\}, V \setminus V^{\preccurlyeq}(u^{\text{cut}})\big) = \varnothing \qquad \text{and} \qquad G\big[V^{\preccurlyeq}(u^{\text{cut}})\big] \in \tilde{\mathcal{D}}^{\pm}_{x_1,u^{\text{cut}}}.$$

If such a point exists, it is unique; if no such point exists, set $u^{\text{cut}} = x_2$. We can now partition $\mathcal{A}$ as

$$\mathcal{A} = \left(\bigcup_{i=1}^{m} \mathcal{A}_i\right) \cup \mathcal{A}_{>m},$$

where

$$\mathcal{A}_i := \big\{G \in \mathcal{A} : u^{\text{cut}} \neq x_2 \text{ and } \big\|L_{\min}\big(x_1, u^{\text{cut}}; G\big[V^{\preccurlyeq}(u^{\text{cut}})\big]\big)\big\| = i\big\},$$
$$\mathcal{A}_{>m} := \big\{G \in \mathcal{A} : \big\|L_{\min}\big(x_1, u^{\text{cut}}; G\big[V^{\preccurlyeq}(u^{\text{cut}})\big]\big)\big\| > m\big\}.$$

Now, if $x_s = u^{\text{cut}}$ and $V' := V^{\preccurlyeq}(u^{\text{cut}})$ as well as $V'' := \{x_s\} \cup \big(\vec{x}_{[n+2]} \setminus V'\big)$, then

$$\mathbf{w}^{\pm}(G) = \mathbf{w}^{\pm}\big(G[V']\big)\mathbf{w}^{\pm}\big(G[V'']\big);$$

that is, the weight factors. Therefore, for every $i \in [m]$,

$$\sum_{n \geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{G \in \mathcal{A}_i} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec{x}_{[3,n+2]} = \lambda \int_{\mathbb{R}^d} \pi^{(i)}_{\lambda}(x_1, u)\tau_{\lambda}(u, x_2)\, \mathrm{d}u. \tag{6.21}$$

Setting

$$\bar{R}_{\lambda,m}(x_2 - x_1) := \sum_{n \geq 0} \frac{\lambda^n}{n!} \int_{(\mathbb{R}^d)^n} \sum_{G \in \mathcal{A}_{>m}} \mathbf{w}^{\pm}(G)\, \mathrm{d}\vec{x}_{[3,n+2]},$$

we can rewrite (6.20) as

$$\tau_{\lambda}(x) = \pi_{\lambda,m}(x) + \lambda\big(\pi_{\lambda,m} * \tau_{\lambda}\big)(x) + \bar{R}_{\lambda,m}(x).$$

One can now prove by hand or by employing Lemma 6.1 that $\bar{R}_{\lambda,m} = R_{\lambda,m}$. □

6.4. Other percolation models

The results of this paper should apply in quite analogous fashion to all other percolation models that enjoy sufficient independence—in particular, to (long-range) bond and site percolation on $\mathbb{Z}^d$. We take bond percolation on $\mathbb{Z}^d$ with edge parameter p as an example. We can adjust our notation by using $\mathcal{C}(x, y, \mathbb{Z}^d)$ to denote the connected subgraphs of $\mathbb{Z}^d$ containing x

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

and y, and we define $\mathcal{D}_{x,y}(\mathbb{Z}^d)$ and the notions for $(\pm)$-graphs analogously. Then one can show that, if we restrict to a finite box $\Lambda \subset \mathbb{Z}^d$, the two-point function satisfies

$$\tau_p^{\Lambda}(x_1, x_2) = \sum_{n \geq 0} p^n \sum_{\substack{G \in \mathcal{C}^{\pm}(x_1, x_2, \Lambda): |E(G)| = n, \\ x_1 \overset{+}{\longleftrightarrow} x_2}} (-1)^{|E^-(G)|}. \tag{6.22}$$

One can easily observe that all graphs summed over in (6.22) that contain more than one $(+)$-cluster cancel out, which is also what happens in the RCM. The direct-connectedness function can be defined analogously to Definition 4.2, providing a suitable setup for an analysis analogous to the one in Section 4.

Acknowledgements

We thank David Brydges, Tyler Helmuth, and Markus Heydenreich for interesting discussions.

Funding information

There are no funding bodies to thank in relation to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

[1] BETSCH, S. AND LAST, G. (2021). On the uniqueness of Gibbs distributions with a non-negative and subcritical pair potential. Preprint. Available at https://arxiv.org/abs/2108.06303.

[2] BRYDGES, D. C. (1986). A short course on cluster expansions. In *Phénomènes Critiques, Systèmes Aléatoires, Théories de Jauge (Les Houches, 1984)*, Parts I–II, North-Holland, Amsterdam, pp. 129–183.

[3] BRYDGES, D. C. AND SPENCER, T. (1985). Self-avoiding walk in 5 or more dimensions. *Commun. Math. Phys.* **97**, 125–148.

[4] CAMPANINO, M. AND IOFFE, D. (2002). Ornstein–Zernike theory for the Bernoulli bond percolation on $\mathbb{Z}^d$. *Ann. Prob.* **30**, 652–682.

[5] CHIEW, Y. C. AND STELL, G. (1989). Connectivity and percolation of randomly centered spheres: correction to the Percus–Yevick approximation. *J. Chemical Phys.* **90**, 4956–4959.

[6] CONIGLIO, A., DEANGELIS, U. AND FORLANI, A. (1977). Pair connectedness and cluster size. *J. Phys. A* **10**, 1123–1139.

[7] HANSEN, J. P. AND MCDONALD, I. R. (2013). *Theory of Simple Liquids: with Applications to Soft Matter*. Academic Press, Amsterdam.

[8] HARA, T. AND SLADE, G. (1990). Mean-field critical behaviour for percolation in high dimensions. *Commun. Math. Phys.* **128**, 333–391.

[9] HEYDENREICH, M. AND VAN DER HOFSTAD, R. (2017). *Progress in High-Dimensional Percolation and Random Graphs*. Springer, Cham.

[10] HEYDENREICH, M., VAN DER HOFSTAD, R., LAST, G. AND MATZKE, K. (2019). Lace expansion and mean-field behavior for the random connection model. Preprint. Available at https://arxiv.org/abs/1908.11356.

[11] HILL, T. L. (1955). Molecular clusters in imperfect gases. *J. Chem. Phys.* **23**, 617–622.

[12] JANSEN, S. (2016). Continuum percolation for Gibbsian point processes with attractive interactions. *Electron. J. Prob.* **21**, article no. 47, 22 pp.

[13] KINGMAN, J. F. C. (1967). Completely random measures. *Pacific J. Math.* **21**, 59–78.

[14] LAST, G. AND PENROSE, M. D. (2018). *Lectures on the Poisson Process*. Cambridge University Press.

[15] LAST, G. AND ZIESCHE, S. (2017). On the Ornstein–Zernike equation for stationary cluster processes and the random connection model. *Adv. Appl. Prob.* **49**, 1260–1287.

[16] MALYSHEV, V. A. AND MINLOS, R. A. (1991). *Gibbs Random Fields: Cluster Expansions* (Mathematics and Its Applications (Soviet Series), Vol. 44). Kluwer Academic Publishers Group, Dordrecht.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press

[17] MEESTER, R. (1995). Equality of critical densities in continuum percolation. *J. Appl. Prob.* **32**, 90–104.
[18] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.
[19] OGATA, Y. AND TANEMURA, M. (1989). Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns. *Ann. Inst. Statist. Math.* **41**, 583–600.
[20] ORNSTEIN, L. S. AND ZERNIKE, F. (1914). Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci. (Amsterdam)* **17**, 793–806.
[21] PENROSE, M. D. (1993). On the spread-out limit for bond and continuum percolation. *Ann. Appl. Prob.* **3**, 253–276.
[22] SLADE, G. (2006). *The Lace Expansion and Its Applications: École d'été de Probabilités de Saint-Flour XXXIV—2004*. Springer, Berlin.
[23] STELL, G. (1976). Correlation functions and their generating functionals: general relations with applications to the theory of fluids. In *Phase Transitions and Critical Phenomena*, Vol. 5b, Academic Press, London, pp. 205–258.
[24] STELL, G. (1996). Continuum theory of percolation and association. *Physica A* **231**, 1–19.
[25] TORQUATO, S. (2002). *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Interdisciplinary *Applied Mathematics*, Vol. 16). Springer, New York.

https://doi.org/10.1017/apr.2022.22 Published online by Cambridge University Press