On the six-vertex model's free energy

Hugo Duminil-Copin*†, Karol Kajetan Kozlowski§, Dmitry Krachun†, Ioan Manolescu‡, Tatiana Tikhonovskaia†

December 21, 2020

Abstract

In this paper, we provide new proofs of the existence and the condensation of Bethe roots for the Bethe Ansatz equation associated with the six-vertex model with periodic boundary conditions and an arbitrary density of up arrows (per line) in the regime $\Delta < 1$. As an application, we provide a short, fully rigorous computation of the free energy of the six-vertex model on the torus, as well as an asymptotic expansion of the six-vertex partition functions when the density of up arrows approaches 1/2. This latter result is at the base of a number of recent results, in particular the rigorous proof of continuity/discontinuity of the phase transition of the random-cluster model, the localization/delocalization behaviour of the six-vertex height function when $a = b = 1$ and $c \geq 1$, and the rotational invariance of the six-vertex model and the Fortuin-Kasteleyn percolation.

1 Introduction

1.1 Definition of the model

The six-vertex model, first proposed by Pauling in 1935 to study the thermodynamic properties of ice, became the archetypical example of a planar integrable model with Lieb's solution of the model in 1967 in its anti-ferroelectric and ferroelectric phases [31, 32, 33, 34] using the Bethe Ansatz. We refer to [3, 35, 41] for detailed expositions and reviews and to [2] for the most general solution. The six-vertex model on the torus is defined as follows. For $N, M > 0$ with N even, let $\mathbb{T}_{N,M} := (\mathbb{Z}/N\mathbb{Z}) \times (\mathbb{Z}/M\mathbb{Z})$ be the N by M torus. An *arrow configuration* ω is a choice of orientation for every edge of $\mathbb{T}_{N,M}$. We say that ω satisfies the *ice rule* (or equivalently that it is a *six-vertex configuration*) if every vertex of $\mathbb{T}_{N,M}$ has two incoming and two outgoing incident edges in ω. These edges can

*Institut des Hautes Études Scientifiques and Université Paris-Saclay
†Université de Genève
§ENS Lyon
‡University of Fribourg

be arranged in six different types around each vertex as depicted in Figure 1, hence the name of the model. One may easily check that the ice-rule guarantees that each horizontal line of vertical edges contains the same number of up arrows. From now on, let $\Omega(\mathbb{T}_{N,M})$ (resp. $\Omega^{(n)}(\mathbb{T}_{N,M})$) be the set of six-vertex configurations (resp. containing exactly n up arrows on each line).

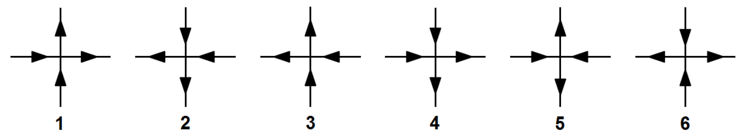


Figure 1: The 6 possibilities, or "types" of vertices in the six-vertex model. Each type comes with a weight $a_1, a_2, b_1, b_2, c_1, c_2$.

For parameters $a_1, a_2, b_1, b_2, c_1, c_2 \geq 0$, define the *weight* of a six-vertex configuration ω to be

$$W_{6\mathrm{V}}(\omega) := a_1^{n_1} a_2^{n_2} b_1^{n_3} b_2^{n_4} c_1^{n_5} c_2^{n_6},$$

where n_i is the number of vertices of $\mathbb{T}_{N,M}$ having type i in ω. In this paper, we choose to focus on $a_1 = a_2 = a$, $b_1 = b_2 = b$ and $c_1 = c_2 = c$. Some of the results of this paper may extend to the asymmetric case and will be the object of a future paper.

Define the *partition functions* of the six-vertex model and of the six-vertex model with n up arrows per line, respectively, by setting

$$Z(\mathbb{T}_{N,M}, a, b, c) := \sum_{\omega \in \Omega(\mathbb{T}_{N,M})} W_{6\mathrm{V}}(\omega),$$
$$Z^{(n)}(\mathbb{T}_{N,M}, a, b, c) := \sum_{\omega \in \Omega^{(n)}(\mathbb{T}_{N,M})} W_{6\mathrm{V}}(\omega).$$

In the analysis of the model, it is customary to introduce the parameter

$$\Delta := \frac{a^2 + b^2 - c^2}{2ab}. \tag{1}$$

Below, we consider the region of parameters (a, b, c) such that $\Delta < 1$; see Figure 2 for the phase diagram of the model.

1.2 Main results for the symmetric six-vertex model

It appears convenient to adopt a parameterisation of the weights which makes transparent the connection with the algebraic Bethe Ansatz construction of the model's transfer matrix. We thus introduce auxiliary parameters $\theta \in (0, \pi)$, $r \in \mathbb{R}_+$ and ζ such that[1]

[1] The existence and uniqueness of r, ζ and θ is proved via basic computation. Notice that the parameter r has no influence on the probabilistic behaviour of the model.

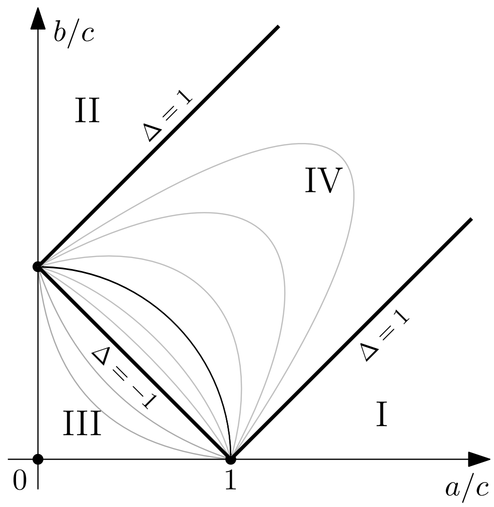


Figure 2: The expected phase diagram of the six-vertex model contains four regions: I and II are called ferroelectric, III is antiferroelectric and IV is disordered. The latter two regions correspond to $\Delta < -1$ and $\Delta \in [-1, 1]$, respectively. The present paper concerns only regions III and IV. The gray lines represent lines of constant Δ; the black quarter-circle corresponds to $\Delta = 0$.

- for $-1 < \Delta < 1$, $\Delta = -\cos\zeta$ with $\zeta \in (0, \pi)$

$$a \sin\tfrac{\zeta}{2} := r \sin\,(1 - \tfrac{\theta}{\pi})\zeta, \quad b \sin\tfrac{\zeta}{2} := r \sin\tfrac{\theta\zeta}{\pi}, \quad c := 2r\cos\tfrac{\zeta}{2}, \tag{2}$$

- for $\Delta = -1$,

$$a := 2r\tfrac{\pi-\theta}{\pi}, \quad b := 2r\tfrac{\theta}{\pi}, \quad c := 2r, \tag{3}$$

- for $\Delta < -1$, $\Delta = -\cosh\zeta$ with $\zeta \in \mathbb{R}_+$

$$a \sinh\tfrac{\zeta}{2} := r \sinh\,(1 - \tfrac{\theta}{\pi})\zeta, \quad b \sinh\tfrac{\zeta}{2} := r \sinh\tfrac{\theta\zeta}{\pi}, \quad c := 2r\cosh\tfrac{\zeta}{2}. \tag{4}$$

The first result goes back to Lieb [32, 33, 34] and Sutherland [40] and deals with the per-site *free energy* defined by

$$f(a, b, c) := \lim_{N\to\infty} \lim_{M\to\infty} \tfrac{1}{MN} \log Z(\mathbb{T}_{M,N}, a, b, c), \tag{5}$$

in which the limits may be taken in any order as established in [35]. The mentioned papers characterised the per-site free energy relying on the same strategy as the original paper [43] which deals with the XXZ quantum spin chain. At the time, the closed expressions for $f(a, b, c)$ were derived under the hypothesis of the so-called condensation of Bethe roots. As will be discussed more precisely later on in the introduction, the condensation property has nowadays been rigorously established. Here, we develop an alternative technique for proving condensation which, on the one hand, turns out to be particularly effective for our goals and, on the other hand, allows one to go beyond what can be rigorously proven within the existing scope of techniques.

Theorem 1. *For every $a \geq b > 0$ and $c \geq 0$ such that $\Delta < 1$ (c.f. (1)), using the parameterisation (2)–(4),*

$$f(a,b,c) = \begin{cases} \log b + \int\limits_{-\infty}^{+\infty} \frac{1}{2t} \frac{\sinh[\frac{2(\pi-\theta)}{\pi}\zeta t]}{\cosh[\zeta t]} \frac{\sinh[(\pi-\zeta)t]}{\sinh[\pi t]} dt & \text{if } -1 < \Delta < 1, \\ \log b + \int\limits_{-\infty}^{+\infty} \frac{\sinh[\frac{2(\pi-\theta)}{\pi} t]}{\cosh[t]} \frac{\mathrm{e}^{-|t|}}{2t} dt & \text{if } \Delta = -1, \\ \log a + \frac{\zeta\theta}{\pi} + \sum\limits_{n=1}^{\infty} \frac{\mathrm{e}^{-|n|\zeta}}{n} \frac{\sinh[2n\zeta\theta/\pi]}{\cosh(n\zeta)} & \text{if } \Delta < -1. \end{cases} \tag{6}$$

In particular $f(1,1,1) = \frac{3}{2}\log(\frac{4}{3})$ and $f(1,1,2) = 2\log[2\Gamma(\frac{5}{4})/\Gamma(\frac{3}{4})]$, with Γ the gamma function.

Our second result deals with the following extension of the per-site free energy to values of n and N:

$$f_N^{(n)}(a,b,c) := \lim_{M\to\infty} \tfrac{1}{NM} \log Z^{(n)}(\mathbb{T}_{N,M}, a, b, c).$$

It provides a characterisation of the subleading corrections to $f_N^{(n)}(a,b,c)$ as $n, N \to +\infty$ in such a way that $n/N \to 1/2$. The condition on n and N appearing in the statement below is technical and takes its origin in the statements of the subsequent theorems in this paper.

Theorem 2. *For $N \geq 2$ even and $a, b, c \geq 0$ leading to $\Delta < 1$ (c.f. (1)), there exist constants $C, C(\zeta), C'(\zeta,\theta) \in (0,\infty)$ such that for every*

$$n \leq \tfrac{1}{2}N - C\min\{\zeta^{-2}, \log(N)^2\}, \tag{7}$$

using the paremetrisation (2)–(4), we have

$$f_N^{(n)}(a,b,c) = f(a,b,c) + O(\tfrac{1}{N}) - (1+o(1)) \begin{cases} C(\zeta)\sin\theta\big(1-\frac{2n}{N}\big)^2 & \text{if } -1 \leq \Delta < 1, \\ C'(\zeta,\theta)(1-\frac{2n}{N}) & \text{if } \Delta < -1, \end{cases} \tag{8}$$

where $o(1)$ means a quantity tending to 0 as n/N tends to $1/2$.

Notice that for $\Delta \in [-1,1)$, (8) only gives meaningful information when $\frac{N}{2} - n$ exceeds $\sqrt{N}$.

This extension has important applications for the six-vertex model and other related models. The six-vertex model lies at the crossroads of a vast family of two-dimensional lattice models; for instance, it has been related to the dimer model, the Ising and Potts models, Fortuin-Kasteleyn (FK) percolation, the loop $O(n)$ models, the Ashkin-Teller models, random permutations, stochastic growth models, quantum spin chains, to cite but a few examples. Among such links, one can use the Baxter-Kelland-Wu mapping between the six-vertex model and FK percolation [4] to deduce from Theorem 2 and the dichotomy result of [13] that the phase transition of FK percolation on the square lattice is continuous if the cluster weight q satisfies $1 \leq q \leq 4$, and is discontinuous for $q > 4$. We refer to the papers where the results were proved (using alternative methods) for additional details [13, 14]. It should be mentioned that the continuity result of [13] may be deduced directly from Theorem 2 using the same procedure as in [18]. In the same spirit, the results can be used to derive dimerisation properties of the anti-ferromagnetic Heisenberg chain [1].

A second application of our results is related to the height function h of the six-vertex model, which can be proved to be localised (meaning that the variance of $h(x) - h(y)$ is bounded uniformly in $|x - y|$) whenever $a = b = 1$ and $c > 2$, and delocalised (meaning that the variance of $h(x) - h(y)$ tends to infinity logarithmically fast in $|x - y|$) when $a = b = 1$ and $1 \leq c \leq 2$. We refer to [16, 18, 22] for more details. It is conjectured more generally that the height-function is localised when $\Delta < -1$, and delocalised when $-1 \leq \Delta < 1$. This property is closely related to the existence of a massive ($\Delta < -1$) and massless ($-1 \leq \Delta < 1$) regime in the XXZ spin-1/2 Heisenberg chain; there, the ground state correlation functions of local operators at distance m decay exponentially fast in $m \to +\infty$ in the massive regime and algebraically in m in the massless regime. Indeed, one can show that the XXZ spin-1/2 Heisenberg Hamiltonian ground state generating function of the longitudinal spin-spin correlations does coincide with the generating function of variances of the height function of the six vertex model. Thus, the power-law decay of the correlators in the XXZ chain translates to the logarithmic growth of the variance of the height function for the six-vertex model.

Finally, perhaps the most important use of our results is in [17], where a refined version of Theorem 2 (see Section 8) is employed to show that the correlations of the height function of the six-vertex model are invariant under rotations in the scaling limit, when $a = b = 1$ and $\sqrt{3} \leq c \leq 2$. This rotation invariance should in fact hold for every $c \in [1, 2]$ (but this has not been proven yet) and be wrong for $c > 2$ (when the height function localises, as discussed above). The argument of [17] involves the FK percolation representation of the six-vertex model, and the rotational invariance result also applies to critical FK percolation on $\mathbb{Z}^2$ with cluster weight $q \in [1, 4]$.

1.3 Transfer matrix of the six-vertex model and the Bethe Ansatz

In order to understand the large scale asymptotics of $Z^{(n)}(\mathbb{T}_{N,M}, a, b, c)$ with $0 \leq n \leq N/2$, one introduces the *transfer matrix* $V_N = V_N(a, b, c)$ (that we do not explicit here; see *e.g.* [7]) defined as an endomorphism of the 2^N-dimensional real vector space spanned by the basis $\{\mathbf{\Psi}_{\vec{x}}\}_{\vec{x}}$, where $\vec{x} = (x_1, \ldots, x_n)$ with $1 \leq x_1 < \cdots < x_n \leq N$, $0 \leq n \leq N$ (below, we use $|\vec{x}| := n$ for the *length* of $\vec{x}$) and $\mathbf{\Psi}_{\vec{x}} = (\Psi_{\vec{x}}(1), \ldots, \Psi_{\vec{x}}(N)) \in \{\pm 1\}^N$ is given by

$$\Psi_{\vec{x}}(i) := \begin{cases} +1 & \text{if } i \in \{x_1, \ldots, x_{|\vec{x}|}\}, \\ -1 & \text{if } i \notin \{x_1, \ldots, x_{|\vec{x}|}\}. \end{cases}$$

In particular, one finds that

$$\begin{aligned} Z(\mathbb{T}_{N,M}, a, b, c) &= \mathrm{Trace}[V_N(a,b,c)^M], \\ Z^{(n)}(\mathbb{T}_{N,M}, a, b, c) &= \mathrm{Trace}[V_N^{(n)}(a,b,c)^M], \end{aligned} \tag{9}$$

where $V_N^{(n)}(a,b,c)$ is the block of the matrix $V_N(a,b,c)$ restricted to the vector space spanned by the $\mathbf{\Psi}_{\vec{x}}$ with $|\vec{x}| = n$. This vector space is indeed stable by $V_N(a,b,c)$ because of the conservation of the number of up arrow per horizontal line. In light of the above displayed equation, we have a clear interest in studying the spectral properties of $V_N(a,b,c)$ and $V_N^{(n)}(a,b,c)$. Standard arguments of rigorous statistical mechanics, see *e.g.* [35], allow one to conclude that

$$f(a,b,c) = \lim_{N\to+\infty} \tfrac{1}{N} \log \Lambda_N(a,b,c) \qquad \text{and} \qquad f_N^{(n)}(a,b,c) = \tfrac{1}{N} \log \Lambda_N^{(n)}(a,b,c) \ ,$$

where $\Lambda_N(a,b,c)$ and $\Lambda_N^{(n)}(a,b,c)$ are the largest eigenvalues of $V_N(a,b,c)$ and $V_N^{(n)}(a,b,c)$, respectively. Note that since $V_N^{(n)}(a,b,c)$ is a Perron-Frobenius matrix, *c.f.* *e.g.* [33], $\Lambda_N^{(n)}(a,b,c)$ is the Perron-Frobenius eigenvalue of $V_N^{(n)}(a,b,c)$. The full transfer matrix $V_N(a,b,c)$ is not Perron-Frobenius, but it may be shown that its single largest eigenvalue is $\Lambda_N^{(0)}(a,b,c)$.

The coordinate Bethe Ansatz, introduced by Bethe [6] in 1931, provides mathematicians and physicists with a powerful way of obtaining eigenvalues of one-dimensional quantum models and of the transfer matrices of certain two-dimensional lattice models. In particular, Orbach [38] put it in a form allowing one to study the eigenvalues of the XXZ spin-1/2 Heisenberg chain, a model sharing the same eigenvectors as the six-vertex transfer matrix, *see* [36] for the explanation of this last fact. Further, since the visionary work of the Leningrad School [19], the coordinate Bethe Ansatz has been put into a fully algebraic framework, called nowadays the algebraic Bethe Ansatz, which is deeply connected with the representation theory of quantum groups. This picture strongly simplified the analysis of integrable models.

We now summarise the program corresponding to the implementation of the Bethe Ansatz to understand the asymptotic of the largest eigenvalue of $V_N^{(n)}(a,b,c)$. The survey [15] contains an elementary derivation of Bethe's Ansatz intended for probabilists, and is a useful reference for most of what is discussed above.

The Bethe Ansatz approach to the dominant eigenvalue

Step 1 Fix distinct integers or half-integers, depending on the parity of n, $(n_1, \ldots, n_n)$ and consider a solution $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ to the following set of equations called the *logarithmic Bethe equations*

$$\frac{N}{2\pi}\mathfrak{p}(\lambda_i) - \frac{1}{2\pi}\sum_{j=1}^{n}\vartheta(\lambda_i - \lambda_j) = n_i, \qquad \forall\, 1 \le i \le n, \tag{10}$$

where $\mathfrak{p}$ and ϑ are defined in Appendix A. While these functions do depend on Δ and have quite different expressions in the regimes $\Delta < -1$, $\Delta = -1$ and $|\Delta| < 1$, we shall keep this dependence implicit. The coordinates of solutions to (10) are called *Bethe roots*.

Step 2 Consider the vector

$$\Psi_N^{(n)}(\boldsymbol{\lambda}) := \sum_{|\vec{x}|=n} \psi(\vec{x}|\boldsymbol{\lambda})\, \boldsymbol{\Psi}_{\vec{x}}\,,$$

for which $\psi(\vec{x}|\boldsymbol{\lambda})$ is defined for every $\vec{x}$ with $|\vec{x}| = n$ by

$$\psi(\vec{x}|\boldsymbol{\lambda}) := \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} \mathrm{e}^{i\mathfrak{p}(\lambda_{\sigma(k)})x_k} \prod_{k<\ell}^{n} \mathfrak{s}(\lambda_{\sigma(\ell)} - \lambda_{\sigma(k)})\,, \tag{11}$$

where $\mathfrak{S}_n$ is the symmetric group on n elements, $\varepsilon(\sigma)$ is the signature of the permutation σ, and

$$\mathfrak{s}(x) := \begin{cases} \sinh(i\zeta + x) & -1 < \Delta < 1, \\ i + x & \Delta = 1, \\ \sin(i\zeta + x) & \Delta < -1. \end{cases} \tag{12}$$

The Bethe Ansatz guarantees that for a solution to (10) which has pairwise distinct coordinates that lie away from the singularities of $\mathfrak{p}$ and ϑ,

$$V_N^{(n)}(a, b, c)\Psi_N^{(n)}(\boldsymbol{\lambda}) = \Lambda_N^{(n)}(\boldsymbol{\lambda})\Psi_N^{(n)}(\boldsymbol{\lambda}), \tag{13}$$

where $\Lambda_N^{(n)}(\boldsymbol{\lambda})$ is given by the formula

$$\Lambda_N^{(n)}(\boldsymbol{\lambda}) := a^N \prod_{j=1}^{n} L(\lambda_j) \,+\, b^N \prod_{j=1}^{n} M(\lambda_j)\,, \tag{14}$$

in which

$$L(\lambda) := \begin{cases} -\frac{\sinh\left(\lambda - \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta\right)}{\sinh\left(\lambda + \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta\right)} \\ -\frac{\lambda - \frac{i}{2} - \frac{i}{\pi}\theta}{\lambda + \frac{i}{2} - \frac{i}{\pi}\theta} \\ -\frac{\sin(\lambda - \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta)}{\sin(\lambda + \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta)} \end{cases} \quad \text{and} \quad M(\lambda) := \begin{cases} -\frac{\sinh(\lambda + \frac{3i}{2}\zeta - \frac{i}{\pi}\theta\zeta)}{\sinh(\lambda + \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta)} & -1 < \Delta < 1, \\ -\frac{\lambda + \frac{3i}{2} - \frac{i}{\pi}\theta}{\lambda + \frac{i}{2} - \frac{i}{\pi}\theta} & \Delta = 1, \\ -\frac{\sin(\lambda + \frac{3i}{2}\zeta - \frac{i}{\pi}\theta\zeta)}{\sin(\lambda + \frac{i}{2}\zeta - \frac{i}{\pi}\theta\zeta)} & \Delta < -1. \end{cases}$$

Step 3 Show that for the specific choice of (half-)integers $n_i \equiv I_i := i - \frac{n+1}{2}$ for $1 \leq i \leq n$, the vector $\Psi_N^{(n)}(\boldsymbol{\lambda})$ produced by Step 2 is the Perron-Frobenius eigenvector of $V_N^{(n)}(a, b, c)$.

Step 4 Perform a large n, N asymptotic expansion of the formula in (14) to conclude.

Note that the Bethe equations ensure that the Bethe roots are not poles of L or M, so that $\Lambda_N^{(n)}(\boldsymbol{\lambda})$ is indeed well defined. Also notice that the Bethe equations and the resulting vector $\Psi_N^{(n)}$ only depend on Δ (or equivalently on ζ); the only dependence on a and b (or equivalently on θ) is in the formula for $\Lambda_N^{(n)}(\boldsymbol{\lambda})$.

At this stage, implementing the above program rigorously requires particular attention at certain points, namely:

In Step 1, for a given choice of $\boldsymbol{n} = (n_1, \ldots, n_n)$, one must prove the existence of solutions to (10). In the regime $\Delta < 1$, Yang and Yang [43] proved the existence of Bethe roots when $n_i = I_i$ as above. Then Griffiths [24] established the existence of solutions to a certain class of (half-)integers $\boldsymbol{n}$. More recently, Kozlowski [30] established the existence of solutions, as well as their uniqueness when N is large enough, for a wide class of (half-)integers $\boldsymbol{n}$ describing the so-called particle-hole excitations.

In Step 2, in order to conclude from (13) that $\Lambda_N^{(n)}(\boldsymbol{\lambda})$ is an eigenvalue of $V_N^{(n)}(a, b, c)$, it must be shown that the Bethe roots' coordinates are pairwise distinct and that $\Psi_N^{(n)}(\boldsymbol{\lambda})$ is non-zero. This was shown to hold for the solution $\boldsymbol{\lambda}$ associated with $n_i = I_i$ by Yang and Yang [43]. For solutions having pairwise distinct coordinates that are associated with other choices of (half-)integers $\boldsymbol{n}$ and which satisfy some form of condensation, *c.f.* later on, the non-vanishing of $\Psi_N^{(n)}(\boldsymbol{\lambda})$ for N large enough may be proved using the determinant representation for the norm of $\Psi_N^{(n)}(\boldsymbol{\lambda})$, which was conjectured in [21, 29] and rigorously proven in [28, 39].

In Step 3, one should argue that the vector $\Psi_N^{(n)}(\boldsymbol{\lambda})$ obtained using the specific choice of (half-)integers I_i is indeed the Perron-Frobenius eigenvector of $V_N^{(n)}(a, b, c)$. This was first conjectured by Húlten [27] and was established by Yang and Yang [43]. Checking that $\Psi_N^{(n)}(\boldsymbol{\lambda})$ is the Perron-Frobenius eigenvector is reasonably simple for Δ equal to 0 or $-\infty$ (for $\Delta = -\infty$ and general n this actually does require some effort). In order to extend the result to an interval of values of Δ, one may prove the continuity or analyticity of $\Psi_N^{(n)}(\boldsymbol{\lambda})$ as a function of Δ. If continuity is used, then one should additionally prove that $\Psi_N^{(n)}(\boldsymbol{\lambda})$ does not vanish outside of a discrete set of values of Δ.

In Step 4, in order to perform the asymptotic expansion, one needs to prove some form of *condensation* of the Bethe roots $\boldsymbol{\lambda}$, *i.e.* the convergence of the point measure $\mathtt{L}_N^{(\boldsymbol{\lambda})} = \frac{1}{N}\sum_{a=1}^{n} \delta_{\lambda_a}$ towards a given measure in the large N limit. To be more precise, we should first introduce the *continuum Bethe equation* whose solution allows one to characterise the limiting measure. For $q \in [0, \infty]$ (when $|\Delta| \leq 1$) and $q \in [0, \pi/2]$ (when $\Delta < -1$), define $\rho(\cdot|q)$ as the solution (the unique solvability was thoroughly discussed, by different methods, in [12, 30, 44]) to the linear integral equation

$$\rho(\lambda|q) + \int_{-q}^{q} K(\lambda-\mu)\rho(\mu|q)d\mu = \xi(\lambda), \qquad \forall \lambda \in \mathbb{R}, \tag{cont.BE}$$

with $K := \frac{1}{2\pi}\vartheta'$ and $\xi := \frac{1}{2\pi}\mathfrak{p}'$.

When $n/N \to m \in [0, 1/2]$ as $n, N \to +\infty$, the point measure $\mathtt{L}_N^{(\boldsymbol{\lambda})}$ associated with the solution $\boldsymbol{\lambda}$ to (10) corresponding to the choice of (half-)integers $n_i = I_i$ converges weakly towards $\rho\big(\lambda|Q(m)\big)\mathbb{1}_{[-Q(m);Q(m)]}(\lambda)d\lambda$, in which $Q(m)$ is the unique solution to

$$\int\limits_{-Q(m)}^{Q(m)} \rho(\lambda|Q(m))d\lambda = m. \tag{15}$$

The existence and uniqueness of $Q(m)$ has been first proven in [12]. We also refer to Appendixes B and D for a proof of $Q(m)$'s existence. The uniqueness of $Q(m)$ may be obtained as a consequence of Theorem 3 below, and will be discussed thereafter. For future reference, it may be useful to note that Q is increasing and $Q(1/2) = \pi/2$ when $\Delta < -1$ and $Q(1/2) = \infty$ when $|\Delta| \leq 1$.

Condensation of Bethe roots was first proven when $0 < \Delta < 1$ by Gusev [26] for any m using convex analysis tools. Much later, Dorlas and Samsonov [10] used different convex analysis techniques to prove the same result and were also able to prove condensation for any m and $\Delta < -\Delta_0$ with Δ_0 large enough, *viz.* perturbatively around $\Delta = -\infty$. More recently, Kozlowski [30] proved condensation for any value of $m \in [0, 1/2]$ and $\Delta \in (-\infty, 1)$, in particular away from the region where convexity or perturbative arguments are applicable. That proof relied on developing a rigorous approach to dealing with the non-linear integral equations governing the so-called counting function of the Bethe roots that were introduced and handled, on a loose level of rigour in [5, 8, 42]. The non-linear integral equation method allowed to rigorously establish the condensation of Bethe roots associated with a large class of (half-)integers in (10), not only $n_i = I_i$, as well as to go beyond the limiting value, and to compute an all order asymptotic expansion in N for $\int f(\mu)d\mathtt{L}_N^{(\boldsymbol{\lambda})}(\mu)$ for any $\Delta < -1$ and $m \in [0, 1/2]$, as well as for any $-1 \leq \Delta < 1$ and $m \in [0, 1/2)$. However, owing to the lack of certain compactness properties, the non-linear integral equation method does not allow one to reach rigorously[2] an estimate beyond $\mathrm{o}(1)$ for

$$\int f(\mu)d\mathtt{L}_N^{(\boldsymbol{\lambda})}(\mu) \;-\; \int\limits_{-Q(\frac{n}{N})}^{Q(\frac{n}{N})} f(\mu)\,\rho(\mu|Q(\tfrac{n}{N}))d\mu \tag{16}$$

when $-1 \leq \Delta < 1$ and $m = 1/2$.

In this work, we develop a method which allows one to estimate (16) up to a $O(1/N)$ for $-1 < \Delta < 1$ and $m = 1/2$ and up to a $O(\ln N/N)$ for $\Delta = -1$ and $m = 1/2$.wit

[2]This was, however, achieved on a formal level in [9].

Reaching these values of the parameters in the model plays a very important role for the results obtained in [16, 17, 18] and this stresses the significance of our result.

1.4 Results for Bethe's equations

For $n \le N/2$, we will henceforth always consider the sequence of (half)-integers

$$n_i \equiv I_i := i - \tfrac{n+1}{2} \qquad 1 \le i \le n, \tag{17}$$

appearing in (10).

For $\Delta < 1$, recall that we are interested in the solutions $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ to

$$\frac{N}{2\pi}\mathfrak{p}(\lambda_i) - \frac{1}{2\pi}\sum_{j=1}^{n} \vartheta(\lambda_i - \lambda_j) = I_i, \qquad \forall 1 \le i \le n, \tag{disc.BE}$$

where $\mathfrak{p}$ and ϑ are defined in Appendix A. We will also ask that solutions are

- *symmetric*, meaning that $\lambda_{n+1-i} = -\lambda_i$ for every $1 \le i \le n$,
- *strictly ordered*, meaning $\lambda_i < \lambda_{i+1}$ for every $1 \le i < n$.

The first main result of this section is the existence of solutions to (disc.BE) without any assumption on $n \le N/2$ or $\Delta \neq -1$, with a quantitative control on how condensed these solutions are.

Theorem 3 (Existence of condensed solutions to discrete Bethe equations when $\Delta \neq -1$)**.** *There exists a constant $C > 0$ such that for every $n \le N/2$ and every $\Delta \in (-\infty, -1) \cup (-1, 1)$, there exists a symmetric strictly ordered solution $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ to* (disc.BE). *Moreover, for every $f : \mathbb{R} \to \mathbb{R}$ with integrable derivative, this solution $\boldsymbol{\lambda}$ satisfies*

$$\Big|\frac{1}{N}\sum_{j=1}^{n} f(\lambda_j) - \int_{-\mathfrak{q}}^{\mathfrak{q}} f(\lambda)\rho(\lambda|\mathfrak{q})\, d\lambda\Big| \le \frac{C}{\zeta N}\, \|f'\|_{L^1(\mathbb{R})}. \tag{Cond}$$

Above, ζ is related to Δ as in (2)–(4)*, while we introduced the shorthand notation*

$$\mathfrak{q} := Q\big(\tfrac{n}{N}\big)\ . \tag{18}$$

A solution satisfying (Cond) will be referred to as *condensed*. Note that the condensation is fairly quantitative but that the control degenerates when Δ is approaching -1. We refer to Theorem 6 for the treatment of the case $\Delta = -1$.

The second theorem will be devoted to the existence of an analytic family of such solutions. The existence of a continuous family of solutions has been previously proven in [43]. Yet, we could not identify any use of the continuity property which warrants mentioning this stronger statement. On the contrary, a property that seems crucial for applications to Bethe's Ansatz is the property of analyticity in Δ of the Bethe roots. Analyticity may also be directly inferred from the results of [30] for $\Delta < -1$ and all m as well as for $-1 \le \Delta < 1$ and $m \in [0, 1/2)$. In this paper, we extend these analyticity results (by another range of arguments) up to $m = 1/2$ in the sense described by Theorem 4 below.

Theorem 4 (Analytic family of solutions to discrete Bethe equations)**.** *For every Δ_0, there exist $N_0(\Delta_0) < \infty$ and $C_0(\Delta_0) < \infty$ such that there exists a* unique *family of condensed symmetric strictly ordered solutions $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ to* (disc.BE)*, which is analytic as a function of Δ on the following intervals:*

- *If $\Delta_0 > -1$, on $(\Delta_0, 1)$ as soon as $N \geq N_0(\Delta_0)$ and $n \leq N/2 - C_0(\Delta_0)$.*
- *If $\Delta_0 < -1$, on $(-\infty, \Delta_0)$ as soon as $N \geq N_0(\Delta_0)$ and $n \leq N/2$.*

Moreover, there exists $\Delta_0 \in (-1, 0)$ such that $N_0(\Delta_0)$ and $C_0(\Delta_0)$ can be taken to be 0.

We are currently unable to prove, with our method, the existence of an analytic solution for arbitrary $n \leq N/2$ over the whole intervals $(-\infty, -1)$ and $(-1, 1)$. We refer to the remarks in Section 3.2 for more details. However, this fact appears to be closely related to the expected property that the model undergoes a phase transition of infinite order at $\Delta = -1$.

Our next result states that the eigenvalue (14) obtained from the Bethe roots provided by Theorem 4 is indeed the Perron-Frobenius eigenvalue of $V_N^{(n)}(a, b, c)$.

Theorem 5 (The Bethe Ansatz gives the Perron-Frobenius eigenvalue)**.** *Fix $n \leq N/2$. For the analytic family of solutions $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ on (u, v) to* (disc.BE) *given by Theorem 4 (with $(u, v) = (-\infty, \Delta_0)$ or $(\Delta_0, 1)$), the quantity $\Lambda_N^{(n)}\big(\boldsymbol{\lambda}(\Delta)\big)$ constructed by* (14) *from $\boldsymbol{\lambda}(\Delta)$ is the Perron-Frobenius eigenvalue of $V_N^{(n)}(a, b, c)$ for every a, b, c such that $\Delta \in (u, v)$.*

The last two theorems have the following direct consequence. For $\Delta \neq -1$ and $n \leq N/2$, consider the solution $\boldsymbol{\lambda}(\Delta)$ provided by Theorem 4. Since the functions $\log|L(x)|$ and $\log|M(x)|$ are differentiable, the condensation and symmetry imply that

$$\tfrac{1}{N}\sum_{j=1}^{n} \log|L(\lambda_j)| = \int_{-\mathfrak{q}}^{\mathfrak{q}} \log|L(\lambda)|\,\rho(\lambda|\mathfrak{q})d\lambda + O(\tfrac{1}{N}),$$

and a similar expression for M. When $a > b$, one may check that the contribution to $\Lambda_N^{(n)}\big(\boldsymbol{\lambda}(\Delta)\big)$ issuing from the L term is larger than the one issuing from the M term. This allows one to deduce from the transfer matrix formalism and Theorem 5 that[3]

$$\begin{aligned} f_N^{(n)}(a, b, c) &= \lim_{M\to\infty} \tfrac{1}{NM} \log \operatorname{trace}[V_N^{(n)}(a, b, c)^M] \\ &= \tfrac{1}{N} \log\big[\Lambda_N^{(n)}\big(\boldsymbol{\lambda}(\Delta)\big)\big] \\ &= \log a + \int_{-\mathfrak{q}}^{\mathfrak{q}} \log|L(\lambda)|\,\rho(\lambda|\mathfrak{q})d\lambda + O(\tfrac{1}{N}) \end{aligned} \tag{19}$$

[3]The absolute value in $\log|L(\lambda)|$ is harmless, as the symmetry of $\boldsymbol{\lambda}$ implies that $\prod_{j=1}^{n} L(\lambda_j)$ is real and positive.

as long as n, N, Δ are in one the cases where Theorem 4 holds and $a \geq b$.

Theorems 1 and 2 follow from (19) once one can estimate efficiently the right-hand side. At the core of this estimate is the following observation going back to [44]. Let $\mathcal{K}$ be the operator acting on $L^2(I)$, where $I = \mathbb{R}$ for $|\Delta| \leq 1$ and $I = [-\frac{\pi}{2}, \frac{\pi}{2}]$ for $\Delta < -1$, constructed from the integral kernel $K(\lambda - \mu)$; let $\mathcal{R}$ be defined by $(\mathrm{id} - \mathcal{R}) = (\mathrm{id} + \mathcal{K})^{-1}$. We refer to $\mathcal{R}$ as the resolvent, and to its integral kernel R as the *resolvent kernel*. Then, (cont.BE) is equivalent to the linear integral equation

$$\rho(\lambda|q) - \int_{I\setminus[-q,q]} R(\lambda - \mu)\rho(\mu|q)d\mu = \rho(\lambda) \qquad \text{for all } \lambda \in \mathbb{R}, \tag{20}$$

where $\rho = (\mathrm{id} - \mathcal{R})\xi$. The resolvent kernel R and ρ are best expressed through their Fourier transforms/coefficients[4]

$$\widehat{R} := \frac{\widehat{K}}{1 + \widehat{K}} \quad \text{and} \quad \widehat{\rho} := \frac{\widehat{\xi}}{1 + \widehat{K}} .$$

We refer to Appendix A for the explicit formulae.

Due to the definition of Q, we have that $I = [-Q(1/2), Q(1/2)]$, and thus $\rho(\lambda) = \rho\big(\lambda|Q(\frac{1}{2})\big)$. The rewriting of (cont.BE) as (20) has the advantage of putting emphasis on the perturbative structure of the equation for q located in the vicinity of $Q(\frac{1}{2})$.

Up to now, our results were always stated for Δ belonging to strict subintervals of $(-\infty, -1)$ or $(-1, 1)$. We conclude this section with a result dealing with the case $\Delta = -1$.

Theorem 6. *There exist $N_0, C_0, C_1 > 0$ such that for every $N \geq N_0$ and*

$$n \leq \frac{N}{2} - C_1(\log N)^2,$$

there exists $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ on $(-1, 1)$ such that

- *for every $\Delta \in (-1, 1)$, $\boldsymbol{\lambda}(\Delta)$ is a solution to* (disc.BE)
- *$\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ is analytic on $(-1, 1)$;*
- *for every $\Delta \in (-1, 1)$ and $f \in L^1(\mathbb{R})$,*

$$\Big|\frac{1}{N}\sum_{j=1}^{n} f\big(\lambda_j(\Delta)\big) - \int_{-\mathfrak{q}}^{\mathfrak{q}} f(\lambda)\, \rho\big(\lambda|\mathfrak{q}\big) d\lambda\Big| \leq C_0 \frac{\log N}{N} \, \|f'\|_{L^1(\mathbb{R})}. \tag{21}$$

[4]When $|\Delta| \leq 1$, we consider square-integrable functions on $\mathbb{R}$. For F in $L^2(\mathbb{R})$, the Fourier transform of F is given by

$$\widehat{F}(t) := \int_{\mathbb{R}} e^{-itx} F(x) dx.$$

When $\Delta < -1$, we consider π-periodic functions that are square integrable on $[-\pi/2, \pi/2]$ (call $L^2_\pi(\mathbb{R})$ the set of π-periodic functions f with $\int_0^\pi |f(t)|^2 dt < \infty$). Then, for $f \in L^2_\pi(\mathbb{R})$, define the Fourier coefficients $\widehat{f} : \mathbb{Z} \to \mathbb{C}$ by

$$f(t) = \frac{1}{\pi} \sum_{n \in \mathbb{Z}} \widehat{f}(n) e^{2int}.$$

In Remark 19, we will also see from the proof that one can obtain a solution $\boldsymbol{\lambda}(-1)$ of (disc.BE) with $\Delta = -1$ by taking the limit of $\frac{1}{\zeta}\boldsymbol{\lambda}(\Delta)$ when ζ tends to 0. This solution also satisfies (21).

Organization The paper is split into seven further sections and several appendixes. In Sections 2 and 3, we present the proofs of Theorems 3 and 4, respectively. The sections themselves start with general considerations and are then divided into the different cases $0 \leq \Delta < 1$, $-1 < \Delta \leq 0$ and $\Delta < -1$, as these exhibit different features. Sections 4 and 5 contain the proofs of Theorems 5 and 6, respectively.

Building upon these results, Theorems 1 and 2 are proved in Sections 6, and 7, respectively. These sections are divided between the cases $|\Delta| < 1$ and $\Delta < -1$ as these correspond to different behaviours.

Finally, Section 8 presents a refined version of Theorem 2. While being interesting in its own right, this result is mostly useful in our subsequent paper [17].

The first Appendix lists the different definitions of functions in order to have a place conveniently gathering all the formulae. The two other Appendixes gather properties of $\rho(\cdot|q)$ and (cont.BE) so as to not overburden the rest of the text.

Acknowledgments The first author was funded by the ERC CriBLaM. The third author was funded by the Swiss FNS. The first, third, fourth and fifth authors were partially funded by the NCCR SwissMap and the Swiss FNS. The first and fourth authors thank Matan Harel for inspiring discussions at the beginning of the project.

2 Proof of Theorem 3

In this whole section, fix $\Delta \in (-\infty, -1) \cup (-1, 1)$ and $n \leq N/2$. Recall that N is even and that $I_i = i - \frac{n+1}{2}$ for $1 \leq i \leq n$.

Below, we introduce the notion of *interlaced solution* which will be useful in the proof. For $q > 0$ (with $q \leq \pi/2$ when $\Delta < -1$) and $x \in (-x_0, x_0)$, where $x_0 = x_0(q) \in \mathbb{R}_+ \cup \{+\infty\}$ is defined by

$$\int_0^\infty \rho(\lambda|q)d\lambda = \tfrac{1}{N}(x_0 - \tfrac{n+1}{2}),$$

introduce the quantile $\Lambda(x|q)$ given by the formula

$$\int_0^{\Lambda(x|q)} \rho(\lambda|q)d\lambda = \tfrac{1}{N}(x - \tfrac{n+1}{2}). \tag{22}$$

Note that $\Lambda(x|q)$ is unambiguously defined since $\rho(\lambda|q) > 0$.

Due to definition of $\rho(\cdot|q)$ (see Appendix A and (20)), x_0 is equal to infinity for $\Delta < -1$ and is finite, but larger than or equal to $\pi/2$ for $\Delta \geq -1$. In the latter case, in order to avoid unnecessarily heavy notation, we set $\Lambda(x|q) = +\infty$ for $x \geq x_0$ and $-\infty$ for $x \leq -x_0$. Note also that by definition of $\mathfrak{q}$, *c.f.* (15) and (18), we have that $\mathfrak{q} = \Lambda(n+\frac{1}{2}|\mathfrak{q}) = -\Lambda(\frac{1}{2}|\mathfrak{q})$.

Definition 7. *For $n \leq \frac{N}{2}$, $k \geq 1$ and $q \in \mathbb{R}_+$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ is (k,q)-*interlaced *if for every $1 \leq i \leq n$,*

$$\Lambda(i - \tfrac{k}{2}|q) \leq \lambda_i \leq \Lambda(i + \tfrac{k}{2}|q). \tag{23}$$

*We say that $\boldsymbol{\lambda}$ is (k,q)-*strictly interlaced *if the strict inequalities hold.*

Remark 8. *When $k = 1$, this corresponds to a perfect interlacement between the λ_i and the quantiles of the measure $\rho(\lambda|q)d\lambda$.*

This notion of interlacement is useful since (k,q)-interlaced solutions obviously satisfy (Cond), as is stated in the next lemma.

Lemma 9 (From interlacement to quantitative condensation)**.** *Fix $k \geq 1$ and $q \in \mathbb{R}_+$. For every (k,q)-interlaced $\boldsymbol{\lambda}$ and every $f : \mathbb{R} \to \mathbb{R}$ with integrable derivative,*

$$\Big|\frac{1}{N}\sum_{j=1}^{N} f(\lambda_j) - \int_{\Lambda(\frac{1}{2}|q)}^{\Lambda(n+\frac{1}{2}|q)} f(\lambda)\rho(\lambda|q)\, d\lambda\Big| \leq \frac{k}{N}\|f'\|_{L^1(\mathcal{I}_k)}. \tag{24}$$

with $\mathcal{I}_k := (\Lambda(1 - \frac{k}{2}|q), \Lambda(n + \frac{k}{2}|q))$. Furthermore, if f is monotonic, the constant $k\|f'\|_{L^1(\mathcal{I}_k)}$ can be replaced by

$$\tfrac{k+1}{2} \max\Big\{\big|f(\lambda_1) - f\big(\Lambda(n + \tfrac{1}{2}|q)\big)\big|, \big|f(\lambda_n) - f\big(\Lambda(\tfrac{1}{2}|q)\big)\big|\Big\}.$$

Proof. By (22), the integral of $\rho(\cdot|q)$ between $\Lambda(j - \frac{1}{2}|q)$ and $\Lambda(j + \frac{1}{2}|q)$ is $\frac{1}{N}$, as long as both arguments are in between $-x_0$ and x_0. Thus, we find

$$\begin{aligned}\Big|\frac{1}{N}\sum_{j=1}^{n} f(\lambda_j) - \int_{\Lambda(\frac{1}{2}|q)}^{\Lambda(n+\frac{1}{2}|q)} f(\lambda)\rho(\lambda|q)d\lambda\Big| &\leq \sum_{j=1}^{n} \int_{\Lambda(j-\frac{1}{2}|q)}^{\Lambda(j+\frac{1}{2}|q)} |f(\lambda_j) - f(\lambda)|\rho(\lambda|q)d\lambda \\ &\leq \frac{1}{N}\sum_{j=1}^{n} \int_{\Lambda(j-\frac{k}{2}|q)}^{\Lambda(j+\frac{k}{2}|q)} |f'(\mu)|d\mu \leq \frac{k}{N}\,\|f'\|_{L^1(\mathcal{I}_k)},\end{aligned}$$

where we invoked the trivial inequality, for $\Lambda(j - \frac{1}{2}|q) \leq \lambda \leq \Lambda(j + \frac{1}{2}|q)$,

$$|f(\lambda_j) - f(\lambda)| = \Big|\int_{\lambda}^{\lambda_j} f'(\mu)d\mu\Big| \leq \int_{\Lambda(j-\frac{k}{2}|q)}^{\Lambda(j+\frac{k}{2}|q)} |f'(\mu)|d\mu.$$

In the case when f is non-decreasing (non-increasing works in the same way), (k,q)-interlacement gives

$$N\int_{\Lambda(j-\frac{k}{2}-1|q)}^{\Lambda(j-\frac{k}{2}|q)} f(\lambda)\rho(\lambda|q)d\lambda \le f(\lambda_j) \le N\int_{\Lambda(j+\frac{k}{2}|q)}^{\Lambda(j+\frac{k}{2}+1|q)} f(\lambda)\rho(\lambda|q)d\lambda. \tag{25}$$

The lower-bound holds for $n \ge j \ge (k+3)/2$ while the upper one for $1 \le j \le n - \frac{k+1}{2}$. Summing the left-hand side over $j \ge (k+3)/2$, bounding from below the remaining sums of $f(\lambda_j)$ by $\frac{k+1}{2} f(\lambda_1)$, and the missing piece of integral by $-\frac{k+1}{2} f\big(\Lambda(n+\frac{1}{2}|q)\big)$ gives the lower bound on the difference. The upper bound follows from analogous considerations. □

The core of the proof of Theorem 3 will be to construct solutions of (disc.BE) that lie in the subset of $\mathbb{R}^n$ given by

$$\Omega_{k,R} := \{\text{strictly } (k,\mathfrak{q})\text{-interlaced, symmetric, strictly ordered } \boldsymbol{\lambda} \text{ with } |\lambda_i| < R, \forall i\}, \tag{26}$$

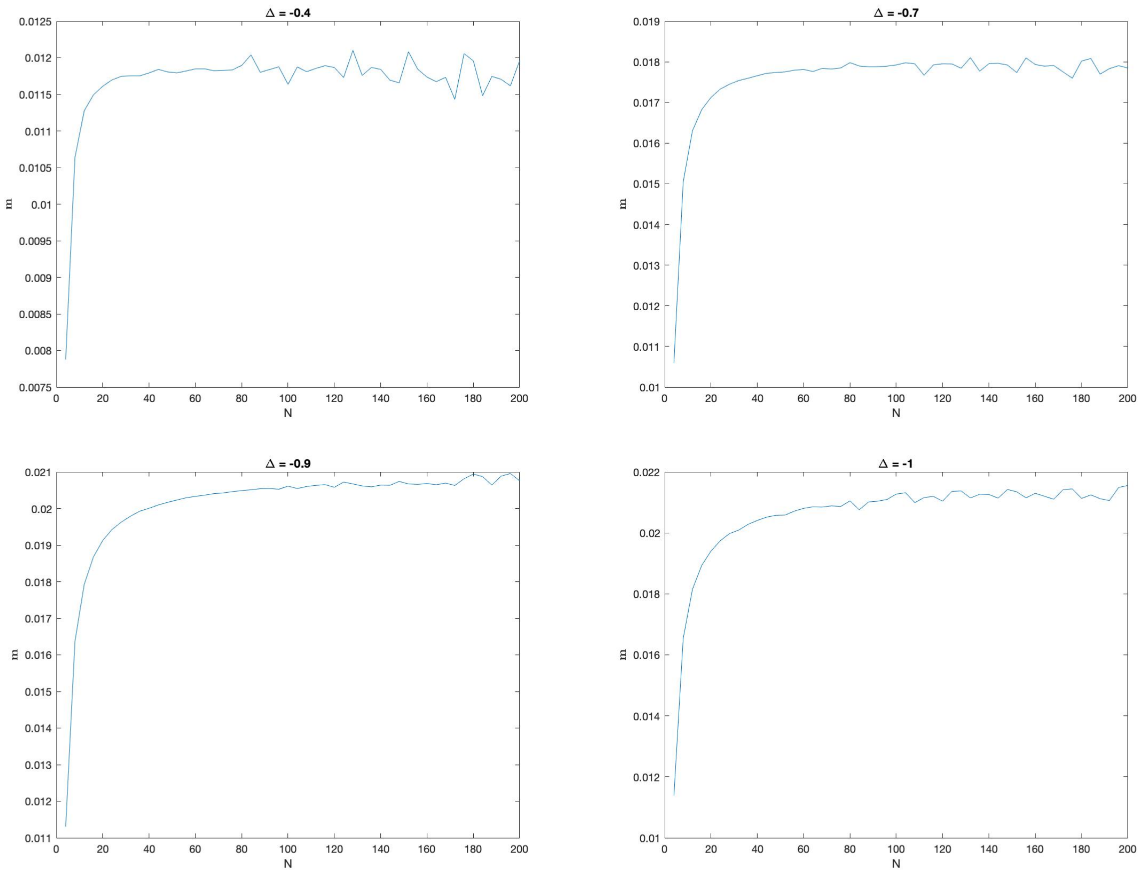
with $\mathfrak{q}$ given by (18), as fixed points of a well-chosen function. This will be done by picking $k = k(\Delta)$ and $R = R(\Delta, N)$ carefully, and then proving that the closure $\overline{\Omega}_{k,R}$ of $\Omega_{k,R}$ is mapped to $\Omega_{k,R}$ by this function. Then, the Brouwer fixed point theorem implies the existence of a fixed point for this function, which is a solution to (disc.BE) due to the choice of the function. While the proof is fairly similar in the different regimes, some tiny differences still exist and we therefore divide it between the cases $\Delta \in [0,1)$, $\Delta \in (-1,0]$, and $\Delta < -1$. Let us mention that a fixed point method was already used, although in a slightly different manner, for proving existence of Bethe roots in [24].

Remark 10. *The reason for distinguishing between $\Delta \ge 0$ and $\Delta < 0$ issues from the fact that $\lambda \mapsto \vartheta(\lambda)$ given in Appendix A is respectively decreasing and increasing. Furthermore, when $\Delta < 0$, dividing between $\Delta < -1$ and $\Delta > -1$ comes from the small caveat that the image of ϑ is an interval of length strictly smaller than 2π when $\Delta > -1$ and equal to 2π when $\Delta < -1$.*

Remark 11. *We expect the existence of $(1,\mathfrak{q})$-interlaced solutions to* (disc.BE) *for every $\Delta < 1$, even though we are currently unable to prove this fact for n close to $N/2$ (when $n/N \le 1/2 - \epsilon$, this follows readily from the results established in [30]). Below are plots of*

$$\mathbf{m} := \max\Big\{N\int_{\Lambda(i)}^{\lambda_i} \rho(\lambda|\infty)d\lambda \ :\ 1 \le i \le N/2\Big\}$$

as a function of the system size N, for $n = N/2$ and different $\Delta \ge -1$. One sees that the quantity remains bounded by $1/2$, meaning that the solution is $(1,+\infty)$-interlaced.



2.1 Case $0 \leq \Delta < 1$

Introduce the smallest *odd* integer $k = k(\Delta) \in \mathbb{Z}_+$ such that

$$2\pi \frac{k}{k+1} > |\vartheta(+\infty) - \vartheta(-\infty)| = |2\pi - 4\zeta|. \tag{27}$$

Consider the map $\Phi : \mathbb{R}^n \to \mathbb{R}^n, \boldsymbol{\lambda} \mapsto \boldsymbol{\mu}$ for which μ_i is defined for every $1 \leq i \leq n$ by

$$\mathfrak{p}(\mu_i) - \tfrac{1}{N} \sum_{j=1}^{n} \vartheta(\mu_i - \lambda_j) = \tfrac{2\pi}{N} I_i, \tag{28}$$

where $\mathfrak{p}$ and ϑ are given in Appendix A.

This function is well-defined as the map $x \mapsto \mathfrak{p}(x) - \frac{1}{N}\sum_{j=1}^{n} \vartheta(x - \lambda_j)$ is continuous strictly increasing (here the fact that $\Delta \geq 0$ ensures that ϑ is decreasing) and tends to $\pm\Upsilon$ with $\Upsilon := (1-\frac{n}{N})\pi - \zeta(1-2\frac{n}{N})$ at $\pm\infty$. Also, since $|\frac{2\pi}{N} I_i| \leq \frac{\pi}{2} - \frac{\pi}{N}$ and $\Upsilon \geq \frac{\pi}{2}$ (recall that $n \leq N/2$ and that $\zeta \leq \pi/2$ in this case), there exists a constant $R = R(\Delta, N)$ such that $|\mu_i| < R$ for every i. From now on, we fix this constant and show that Φ maps $\overline{\Omega}_{k,R}$ onto $\Omega_{k,R}$. The Brouwer fixed point theorem then implies that Φ has a fixed point (since $\overline{\Omega}_{k,R}$

is a compact convex set), which is a strictly $(k, \mathfrak{q})$-interlaced symmetric strictly ordered solution of (disc.BE).

Let $\boldsymbol{\mu} = \Phi(\boldsymbol{\lambda})$ for some $\boldsymbol{\lambda} \in \overline{\Omega}_{k,R}$. That $\boldsymbol{\mu}$ is strictly ordered and symmetric is obvious from (28), whose left-hand side is strictly increasing in μ_i. We therefore only need to check the strict $(k, \mathfrak{q})$-interlacement, which is a direct consequence of the following sequence of inequalities:

$$2\pi\Big|N\int_0^{\mu_i} \rho(\lambda|\mathfrak{q})d\lambda - I_i\Big| = \Big|N\mathfrak{p}(\mu_i) - N\int_{-\mathfrak{q}}^{\mathfrak{q}} \vartheta(\mu_i-\lambda)\rho(\lambda|\mathfrak{q})d\lambda - 2\pi I_i\Big|$$
$$\leq \tfrac{k+1}{2}|\vartheta(+\infty) - \vartheta(-\infty)| \overset{(27)}{<} \pi k. \tag{29}$$

The first equality in (29) is due to the identity, valid for every x and q,

$$2\pi\int_0^{x} \rho(t|q)dt = \mathfrak{p}(x) - \int_{-q}^{q} \vartheta(x-\lambda)\rho(\lambda|q)d\lambda, \tag{30}$$

which is the integrated version of (cont.BE) (recall that ϑ and $\mathfrak{p}$ are odd). The first inequality in (29) is an application of the definition of μ_i together with Lemma 9 applied to the monotone function $\vartheta(\mu_i - \cdot)$; it is also useful to mention that $\mathfrak{q} = \Lambda(n + \frac{1}{2}|\mathfrak{q}) = -\Lambda(1 - \frac{1}{2}|\mathfrak{q})$, due to the definitions of $\mathfrak{q}$ and $\Lambda(\cdot|q)$.

2.2 Case $-1 < \Delta < 0$

As before, introduce the smallest *odd* integer $k = k(\Delta) \in \mathbb{Z}_+$ such that (27) holds.

We are unable to use the map Φ from the previous subsection as ϑ is now increasing. We therefore change the map slightly and consider the map $\Psi : \mathbb{R}^n \to \mathbb{R}^n, \boldsymbol{\lambda} \mapsto \boldsymbol{\mu}$ for which μ_i is defined for every $1 \leq i \leq n$ by

$$\mathfrak{p}(\mu_i) = \tfrac{1}{N}\sum_{j=1}^{n} \vartheta(\lambda_i - \lambda_j) + \tfrac{2\pi}{N} I_i. \tag{31}$$

The map is again well-defined as $\mathfrak{p}$ is continuous, strictly increasing, and $\mathfrak{p}(\mathbb{R})$ is equal to $[\zeta - \pi, \pi - \zeta]$, while for any $\boldsymbol{\lambda} \in \mathbb{R}^n$,

$$-\tfrac{2n}{N}(\pi - \zeta) + \tfrac{\pi}{N} \;<\; \tfrac{1}{N}\sum_{j=1}^{n} \vartheta(\lambda_i - \lambda_j) + \tfrac{2\pi}{N} I_i \;<\; \tfrac{2n}{N}(\pi - \zeta) - \tfrac{\pi}{N}\;,$$

(we use that $|\vartheta| \leq \pi - 2\zeta$, $|I_i| \leq (n-1)/2$) which ensures that the left-hand side of (31) lies in the range of $\mathfrak{p}$ since $n \leq N/2$. Thus, as before, we may find $R = R(\Delta, N)$ large enough such that $|\mu_i| < R$ for every i.

Again, we wish to prove that Ψ is mapping $\overline{\Omega}_{k,R}$ to $\Omega_{k,R}$, which will imply the existence of a fixed point, and therefore a strictly $(k, \mathfrak{q})$-interlaced, symmetric strictly, ordered solution to (disc.BE).

Fix $\boldsymbol{\lambda} \in \overline{\Omega}_{k,R}$ and set $\boldsymbol{\mu} = \Psi(\boldsymbol{\lambda})$. The strict monotonicity and the fact that $\mu_i \in (-R, R)$ are immediate consequences of the definition of Ψ and the choice of R, and we do not give further details. Lemma 9 applied to the decreasing function $\vartheta(\lambda_i - \cdot)$ implies that

$$N\mathfrak{p}(\mu_i) \leq N \int_{-\mathfrak{q}}^{\mathfrak{q}} \vartheta(\lambda_i - \lambda)\rho(\lambda|\mathfrak{q})d\lambda + \tfrac{k+1}{2} \max\big\{|\vartheta(\lambda_i - \lambda_1) - \vartheta(\lambda_i - \mathfrak{q})|, |\vartheta(\lambda_i - \lambda_n) - \vartheta(\lambda_i + \mathfrak{q})|\big\} + 2\pi I_i.$$

Observe now that, due to (27), the maximum above is smaller than $|2\pi - 4\zeta| < 2\pi\frac{k}{k+1}$. Since in addition λ_i was assumed smaller than $\Lambda(i + \frac{k}{2}|\mathfrak{q})$, and since ϑ is increasing, we conclude that

$$N\mathfrak{p}(\mu_i) < N \int_{-\mathfrak{q}}^{\mathfrak{q}} \vartheta(\Lambda(i + \tfrac{k}{2}|\mathfrak{q}) - \lambda)\rho(\lambda|\mathfrak{q})d\lambda + 2\pi I_i + 2\pi\tfrac{k}{2} = N\mathfrak{p}\big(\Lambda(i + \tfrac{k}{2}|\mathfrak{q})\big),$$

where the last equality follows from (30) and the definition of $\Lambda(i + \frac{k}{2}|\mathfrak{q})$. Since $\mathfrak{p}$ is increasing, we get that $\mu_i < \Lambda(i + \frac{k}{2}|\mathfrak{q})$. Similarly, one proves that $\mu_i > \Lambda(i - \frac{k}{2}|\mathfrak{q})$.

2.3 Case $\Delta < -1$

For $\Delta < -1$ and $q > 0$, first observe that the function $\varphi(\lambda) := 2\vartheta(\lambda) - \vartheta(\lambda + \frac{\pi}{2}) - \vartheta(\lambda - \frac{\pi}{2})$ is increasing on $[0, \pi/4]$ and decreasing on $[\pi/4, \pi/2]$ with $\varphi(0) = \varphi(\pi/2) = 0$ and $\varphi(\pi/4) \in (0, 2\pi)$. Moreover, φ is π-periodic and even (due to the corresponding properties for ϑ), and therefore $\varphi(\pi/4)$ is its maximum over all $\mathbb{R}$. Then, introduce the smallest *odd* integer $k \in \mathbb{Z}_+$ such that

$$2\pi\frac{k}{k+1} > \sup\big\{|\varphi(\lambda)| \,:\, \lambda \in \mathbb{R}\big\} = \varphi(\tfrac{\pi}{4})\,. \tag{32}$$

In the present case, we reuse the map Ψ defined in Section 2.2. This map is well defined since, in this regime of Δ, $\mathfrak{p}$ is strictly increasing and $\mathfrak{p}(\mathbb{R}) = \mathbb{R}$.

For small values of N, the existence of a fixed point of Ψ (or equivalently of a solution to (disc.BE)) that is not necessarily (k, R)-interlaced is easily obtained. Its condensation may be derived by adjusting the constant C in (Cond). Henceforth we focus on values of N above a threshold independent of Δ chosen below.

Let $\boldsymbol{\mu} = \Psi(\boldsymbol{\lambda})$ for some $\boldsymbol{\lambda} \in \overline{\Omega}_{k,R}$. As in the previous part, it is immediate that $\boldsymbol{\mu}$ is symmetric and strictly ordered. One should still establish the boundedness and the strict $(k, \mathfrak{q})$-interlacement of $\boldsymbol{\mu}$. We will argue that the former is a direct consequence of the later. We thus first establish interlacement.

To do so, one should start by establishing a generalisation of Lemma 9 to the case of a function $g : [0, +\infty) \to \mathbb{R}$ which is monotonous on $[0, \pi/2]$. Here, we only treat the case of n even and leave the details of n odd to the reader, since it only leads to minor modifications. We claim that, for any such function g,

$$\Big|\sum_{i=1+\frac{n}{2}}^{n} g(\lambda_i) \;-\; N\int_0^{\mathfrak{q}} g(\mu)\rho(\mu|\mathfrak{q})d\mu\Big| \;\leq\; \frac{k+1}{2}\max\big\{\mathfrak{m}^+[g],\mathfrak{m}^-[g]\big\}, \tag{33}$$

where

$$\mathfrak{m}^+[g] := \max\big\{|g(\lambda_j)-g(0)| \;:\; j=n-\tfrac{k-1}{2},\ldots,n\big\} \quad \text{and} \quad \mathfrak{m}^-[g] := |g(\mathfrak{q})-g(\lambda_{\frac{n}{2}+1})|\;.$$

The inequality above is obtained in the same way as Lemma 9, and we give no further details.

Define $\vartheta^{\mathrm{sym}}(\lambda,\mu) := \vartheta(\lambda-\mu)+\vartheta(\lambda+\mu)$. Then a direct computation shows that the functions $\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)$ for $i=1,\ldots,n$ are monotonous on $[0,\pi]$. Applying (33) to $\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)$ we find

$$\begin{aligned}\mathfrak{p}(\mu_i) &= \tfrac{1}{N}\sum_{j=1+\frac{n}{2}}^{n}\vartheta^{\mathrm{sym}}(\lambda_i,\lambda_j)+\tfrac{2\pi}{N}I_i\\ &\geq \int_0^{\mathfrak{q}}\vartheta^{\mathrm{sym}}(\lambda_i,\mu)\rho(\mu|\mathfrak{q})d\mu \;-\; \tfrac{k+1}{2N}\max\big\{\mathfrak{m}^+[\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)],\mathfrak{m}^-[\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)]\big\}+\tfrac{2\pi}{N}I_i\;.\end{aligned} \tag{34}$$

It follows from the lower bound $\rho(x|q)\geq\rho(x)\geq\frac{1}{2\zeta}$ established in Lemma 29 of Appendix D, that for each j

$$\Lambda(n+\tfrac{1}{2}|\mathfrak{q}) \;-\; \Lambda(n-j+\tfrac{1}{2}|\mathfrak{q}) \leq\; 2\zeta\int_{\Lambda(n-j+\frac{1}{2}|\mathfrak{q})}^{\Lambda(n+\frac{1}{2}|\mathfrak{q})}\rho(\lambda|\mathfrak{q})d\lambda = 2j\tfrac{\zeta}{N}. \tag{35}$$

Therefore, the $(k,\mathfrak{q})$-interlacement of $\boldsymbol{\lambda}$ allows one to infer that $\lambda_j=\mathfrak{q}+O(\frac{k}{N})$, with the $O(.)$ here and below being uniform in $j=n-\frac{k-1}{2},\ldots,n$ and $\Delta<-1$. Hence, since $\mathfrak{q}\leq\pi/2$, any λ_j appearing in the definition of $\mathfrak{m}^+[g]$ exceeds $\pi/2$ by at most $O(\frac{k}{N})$.

Direct computation show that $\pi/2$ is a local extremum of $\mu\mapsto\vartheta^{\mathrm{sym}}(\lambda,\mu)$ on $[\pi/2-\eta,\pi/2+\eta]$ for some $\eta>0$ independent of Δ and λ. Thus, we conclude that for all N large enough (which we will assume henceforth for reasons described at the start of the proof), every $2n\leq N$ and $i\in\{1+\frac{n}{2},n\}$,

$$\max\big\{\mathfrak{m}^+[\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)],\mathfrak{m}^-[\vartheta^{\mathrm{sym}}(\lambda_i,\cdot)]\big\}\leq|\vartheta^{\mathrm{sym}}(\lambda_i,\tfrac{\pi}{2})-\vartheta^{\mathrm{sym}}(\lambda_i,0)|=|\varphi(\lambda_i)|.$$

Plugging the above into (34), we find

$$\mathfrak{p}(\mu_i)\geq\int_0^{\mathfrak{q}}\vartheta^{\mathrm{sym}}(\lambda_i,\mu)\rho(\mu|\mathfrak{q})d\mu \;-\; \tfrac{k+1}{2N}\varphi(\tfrac{\pi}{4})+\tfrac{2\pi}{N}I_i\;.$$

Invoking the choice of k and the fact that $\lambda \mapsto \vartheta(\lambda)$ is increasing on $\mathbb{R}$ gives that

$$\mathfrak{p}(\mu_i) > \int_{-\mathfrak{q}}^{\mathfrak{q}} \vartheta\big(\Lambda(n-\tfrac{k}{2}|\mathfrak{q}) - \mu\big)\rho(\mu|\mathfrak{q})d\mu \, + \, \tfrac{2\pi}{N} I_{i-\frac{k}{2}} \, = \, \mathfrak{p}(\Lambda(n-\tfrac{k}{2}|\mathfrak{q})) \, .$$

This yields the lower bound for the $(k,\mathfrak{q})$-interlacement of $\boldsymbol{\mu}$. The upper bound is obtained in an analogous way.

Finally, the $(k,\mathfrak{q})$-interlacement of $\boldsymbol{\mu}$ and the upper bound $\Lambda(n+\frac{k}{2}|\mathfrak{q}) - \Lambda(n+\frac{1}{2}|\mathfrak{q}) \leq \frac{\zeta(k-1)}{N}$ ensure that $\mu_n \leq \mathfrak{q} + \frac{\zeta(k-1)}{N}$ and thus, by symmetry, that $\mu_i \in (-R,R)$ with $R := \frac{\pi}{2} + \frac{\zeta(k-1)}{N}$. The latter establishes that $\Psi(\overline{\Omega}_{k,R}) \subset \Omega_{k,R}$.

3 Proof of Theorem 4

Fix $n \leq N/2$. Since the dependence in Δ plays a role in this argument, we recall it in the subscript of the map $\mathsf{T} : (\Delta, \boldsymbol{\lambda}) \mapsto \mathsf{T}_\Delta(\boldsymbol{\lambda})$ from $[-\infty, 1) \times \mathbb{R}^n$ to $\mathbb{R}^n$ defined by the formula, for every $1 \leq i \leq n$,

$$\big[\mathsf{T}_\Delta(\boldsymbol{\lambda})\big]_i = \tfrac{1}{2\pi}\mathfrak{p}(\lambda_i) - \tfrac{1}{2\pi N}\sum_{j=1}^{n} \vartheta(\lambda_i - \lambda_j) - \tfrac{1}{N} I_i \, , \quad 1 \leq i \leq n \, . \tag{36}$$

We remind that $\mathfrak{p}$ and ϑ appearing above do depend on Δ, *c.f.* Appendix A.

The zeroes of T_Δ correspond to the solutions to (disc.BE) for Δ. The following proposition will play a key role in the proof of Theorem 4.

Proposition 12. *Let k be as defined by* (27) *for $-1 < \Delta < 1$ and* (32) *for $\Delta < -1$. Then, there exists some universal constant C such that, for every $\Delta \in (-\infty, 1) \setminus \{-1\}$, every N large enough, and every*

$$n \leq N/2 - Ck^2 \mathbb{1}_{(-1,0)}(\Delta) \, ,$$

we have that $d\mathsf{T}_\Delta$ is invertible at $\boldsymbol{\lambda}$ for any $(k,\mathfrak{q})$-interlaced, ordered, symmetric $\boldsymbol{\lambda}$.

With this proposition at hand, we are in position to prove the theorem.

Proof of Theorem 4. Taking into account the definition of T_Δ and introducing $\Omega(\Delta) := \Omega_{k(\Delta),R(\Delta,N)}$, with $\Omega_{k,R}$ given by (26) and $R(\Delta,N)$ as constructed in Subsections 2.1, 2.2 and 2.3, depending on the value of Δ, we can restate the theorem as the existence of an analytic family $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ such that $\boldsymbol{\lambda}(\Delta) \in \Omega(\Delta)$ satisfies $\mathsf{T}_\Delta(\boldsymbol{\lambda}(\Delta)) = 0$ for every Δ.

Consider some Δ_0 for which we are in the possession of $\boldsymbol{\lambda}(\Delta_0) \in \Omega(\Delta_0)$ satisfying $\mathsf{T}_{\Delta_0}(\boldsymbol{\lambda}(\Delta_0)) = 0$. Using Proposition 12, the implicit function theorem for analytic functions gives the existence of an analytic family $\Delta \mapsto \boldsymbol{\lambda}(\Delta) \in \mathbb{R}^n$ such that $\mathsf{T}_\Delta(\boldsymbol{\lambda}(\Delta)) = 0$ in a small neighbourhood of Δ_0. Continuity implies that by reducing the neighbourhood if need be, we can further assume that $\boldsymbol{\lambda}(\Delta) \in \Omega(\Delta)$.

Also note that a continuous limit as $\Delta \to \Delta_1$ of $\boldsymbol{\lambda}(\Delta) \in \Omega(\Delta)$ with $\mathsf{T}_\Delta(\boldsymbol{\lambda}(\Delta)) = 0$ converges to $\boldsymbol{\lambda}(\Delta_1) \in \overline{\Omega}(\Delta_1)$ with $\mathsf{T}_{\Delta_1}(\boldsymbol{\lambda}(\Delta_1)) = 0$. Yet, we saw in the previous section that solutions to (disc.BE) in $\overline{\Omega}(\Delta_1)$ are necessarily in $\Omega(\Delta_1)$. Together with the previous paragraph, this implies the existence of an analytic family of solutions on any open interval on which the conditions of Proposition 12 hold, and which contains at least one value Δ for which there exists a solution $\boldsymbol{\lambda} \in \Omega(\Delta)$ to (disc.BE). The intervals of Theorem 4 are indeed such that the conditions of Proposition 12 hold; the existence of solutions for some Δ in these intervals is ensured by Theorem 3 (or alternatively by Lemmata 16 and 17, see below).

To prove the uniqueness of the solutions for all Δ, it suffices to prove it for a single value Δ_1 in each of the two intervals of Theorem 4. Indeed, assuming the existence of multiple solutions at some value of Δ, the argument above implies the existence of multiple analytic families of solutions in the whole interval. These families may no cross inside the interval, due to the implicit function theorem, and would therefore contradict the uniqueness at Δ_1. We choose to check the uniqueness of solutions for $\Delta_1 = 0$ and Δ_1 a very large negative number. This is done by solving (disc.BE) explicitly for $\Delta = 0$ and $\Delta = -\infty$, then extending the property to large negative numbers by continuity; see Lemmata 16 and 17 below for more details. □

We now focus on the proof of Proposition 12, and divide it into three subsections depending on the range of Δ as before. Note that since $K = \frac{1}{2\pi}\vartheta'$ and $\xi = \frac{1}{2\pi}\mathfrak{p}'$, $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ can be rewritten as

$$\Big[d\mathsf{T}_\Delta(\boldsymbol{\lambda})\Big]_{ij} = \begin{cases} \xi(\lambda_i) - \dfrac{1}{N}\displaystyle\sum_{\ell \neq i} K(\lambda_i - \lambda_\ell) & i = j, \\ \dfrac{K(\lambda_i - \lambda_j)}{N} & i \neq j. \end{cases}$$

3.1 Proof of Proposition 12 when $0 \leq \Delta < 1$

For any $\boldsymbol{\lambda} \in \mathbb{R}^n$, the matrix $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ is symmetric (since K is even) and positive definite (since $K \geq 0$ and $\xi > 0$ – these are derivatives of increasing and strictly increasing functions, respectively). As a consequence, it is invertible.

Remark 13. *Alternatively, in this regime, one can obtain existence and uniqueness of the solution to the Bethe equation* (disc.BE) *as follows. Since $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ is a positive definite matrix, the function*

$$S(\boldsymbol{\lambda}) := \sum_{i=1}^{n} \Big(\int_0^{\lambda_i} \mathfrak{p}(\mu) d\mu - \frac{1}{2N} \sum_{j=1}^{n} \int_0^{\lambda_i - \lambda_j} \vartheta(\mu) d\mu - \frac{2\pi I_i}{N}\lambda_i \Big)$$

is strictly convex, and has therefore at most one extremum which, if exists, is its minimum. The existence thereof can be obtained in at least three ways. Either one uses the fixed point theorem in the previous section, or one checks that S tends to infinity as soon as one of *the λ_i tends to infinity (this is slightly technical), or finally one observes that at $\Delta = 0$ there is an explicit solution and that the implicit function theorem guarantees that this solution extends into an analytic function on $0 \le \Delta < 1$.*

3.2 Proof of Proposition 12 when $-1 < \Delta < 0$

Recall that in this regime we restrict our attention to n satisfying

$$n \le N/2 - Ck^2, \tag{37}$$

where k is given by (27) while C is yet to be determined. Also, we remind that $\mathfrak{q}$ is given by (18).

Fix $\boldsymbol{\lambda}$ as in the proposition. The matrix $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ is no longer obviously positive definite and therefore not obviously invertible. In order to prove invertibility, we rather show that the matrix A defined by

$$A_{ij} := \frac{d\mathsf{T}_\Delta(\boldsymbol{\lambda})_{ij}}{\rho(\lambda_j|\mathfrak{q})} = \begin{cases} \dfrac{1}{\rho(\lambda_i|\mathfrak{q})}\big[\xi(\lambda_i) - \dfrac{1}{N}\displaystyle\sum_{\substack{\ell=1\\ \neq i}}^{n} K(\lambda_i - \lambda_\ell)\big] & i = j, \\ \dfrac{K(\lambda_i - \lambda_j)}{N\rho(\lambda_j|\mathfrak{q})} & i \neq j, \end{cases}$$

is diagonal dominant[5] hence invertible (note that Proposition 24(i) of Appendix B gives that $\rho(\lambda|\mathfrak{q}) > 0$ on $\mathbb{R}$ and therefore the matrix A is well-defined). The invertibility of $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ follows trivially from that of A. Checking diagonal dominance relies on two computations.

On the one hand, Lemma 9 applied to $\lambda \mapsto K(\lambda_j - \lambda)$ (since $\boldsymbol{\lambda}$ is $(k, \mathfrak{q})$-interlaced) together with (cont.BE) gives

$$A_{jj} \ge \frac{1}{\rho(\lambda_j|\mathfrak{q})}\big(\xi(\lambda_j) - \int_{-\mathfrak{q}}^{\mathfrak{q}} K(\lambda_j - \lambda)\rho(\lambda|\mathfrak{q})d\lambda - \frac{k}{N}\|K'\|_{L^1(\mathbb{R})}\big) = 1 - \frac{k\|K'\|_{L^1(\mathbb{R})}}{N\rho(\lambda_j|\mathfrak{q})}. \tag{38}$$

On the other hand, Lemma 9 applied to the function

$$\lambda \mapsto f_j(\lambda) := K(\lambda_j - \lambda)/\rho(\lambda|\mathfrak{q})$$

gives[6]

$$\begin{aligned} \sum_{\ell \neq j} A_{j\ell} &= \frac{1}{N}\sum_{\ell \neq j} \frac{K(\lambda_j - \lambda_\ell)}{\rho(\lambda_\ell|\mathfrak{q})} \\ &\le \int_{-\mathfrak{q}}^{\mathfrak{q}} \frac{K(\lambda_j - \lambda)}{\rho(\lambda|\mathfrak{q})}\rho(\lambda|\mathfrak{q})d\lambda + \frac{k}{N}\|f_j'\|_{L^1(\mathcal{I}_k)} \\ &= \frac{\vartheta(\lambda_j + \mathfrak{q}) - \vartheta(\lambda_j - \mathfrak{q})}{2\pi} + \frac{k}{N}\|f_j'\|_{L^1(\mathcal{I}_k)} \\ &\overset{(27)}{\le} \frac{k}{k+1} + \frac{k}{N}\|f_j'\|_{L^1(\mathcal{I}_k)}, \end{aligned} \tag{39}$$

[5] Meaning that $A_{ii} > \sum_{j \neq i} |A_{ij}|$ for every i.

[6] We also use the trivial facts that $K(0) \ge 0$, $\rho(\lambda|\mathfrak{q}) \ge 0$ and $K = \frac{1}{2\pi}\vartheta'$ in the first inequality.

where $\mathcal{I}_k := [\Lambda(1-\frac{k}{2}|\mathfrak{q}), \Lambda(n+\frac{k}{2}|\mathfrak{q})]$.

Now, we estimate the error terms (meaning the terms with factor $\frac{k}{N}$) in (38) and (39) separately. Below, the constants C_i are independent of everything else. We use analytic properties of the solution to the continuum Bethe Equation (cont.BE) that are proved in Proposition 24 of Appendix B. The assumption (37) plays an essential role in what follows.

We start by estimating (38). Using that $|\rho'(\lambda)| \leq \pi\rho(\lambda)/\zeta$ (see Proposition 24(ii)) and further invoking Proposition 24(i), we find:

$$\rho(-\mathfrak{q}) - \rho(\Lambda(1-\tfrac{k}{2}|\mathfrak{q})) \leq \tfrac{\pi}{\zeta}\int_{\Lambda(1-k/2|\mathfrak{q})}^{\Lambda(1/2|\mathfrak{q})} \rho(\lambda)d\lambda \leq \tfrac{\pi}{\zeta}\int_{\Lambda(1-k/2|\mathfrak{q})}^{\Lambda(1/2|\mathfrak{q})} \rho(\lambda|\mathfrak{q})d\lambda = \tfrac{\pi}{\zeta}\tfrac{k-1}{2N}. \tag{40}$$

Furthermore, Proposition 24(iii) leads to

$$\rho(-\mathfrak{q}) \geq \tfrac{c_1}{\zeta}(\tfrac{1}{2}-\tfrac{n}{N}) \quad \text{with} \quad c_1 > 0. \tag{41}$$

Combining the two last displayed equations we infer a lower bound

$$\rho(\Lambda(1-\tfrac{k}{2}|\mathfrak{q})) \geq \tfrac{c_1}{\zeta}(\tfrac{1}{2}-\tfrac{n+C_1k}{N}) \quad \text{with} \quad C_1 = \frac{\pi}{c_1}.$$

Then Proposition 24(i), the monotonicity of $\rho(\cdot)$ on $(-\infty, 0]$ and the symmetry and $(k,\mathfrak{q})$-interlacement of $\boldsymbol{\lambda}$ give that

$$\rho(\lambda_j|\mathfrak{q}) \geq \rho(\lambda_j) \geq \rho(\lambda_1) \geq \rho(\Lambda(1-\tfrac{k}{2}|\mathfrak{q})) \geq \tfrac{c_1}{\zeta}(\tfrac{1}{2}-\tfrac{n+C_1k}{N}). \tag{42}$$

Now, since K is unimodal, even, and has limits 0 at $\pm\infty$, $\|K'\|_{L^1(\mathbb{R})} = 2K(0)$. Using this and the previous paragraph, we find

$$\frac{k}{N}\frac{\|K'\|_{L^1(\mathbb{R})}}{\rho(\lambda_j|\mathfrak{q})} \leq \frac{2\zeta K(0)k}{c_1(N/2-n-C_1k)} \leq \frac{C_2k}{N/2-n-C_1k}\ , \tag{43}$$

where the last inequality is obtained by observing that $K(0) \leq C_0/\zeta$ for some ζ-independent constant C_0.

We now turn to the error term in (39). Proposition 24(ii) implies that for $1 \leq j \leq n$, $t \in \mathcal{I}_k$, and N large enough,

$$\begin{aligned}|f_j'(t)| &\leq \frac{|K'(\lambda_j-t)|}{\rho(t|\mathfrak{q})} + |K(\lambda_j-t)|\frac{|\rho'(t|\mathfrak{q})|}{\rho(t|\mathfrak{q})^2} \\ &\leq \frac{|K'(\lambda_j-t)| + |K(\lambda_j-t)|\frac{C_3}{\zeta}\exp(\frac{\pi}{\zeta}[\Lambda(\frac{1}{2}|\mathfrak{q})-\Lambda(1-\frac{k}{2}|\mathfrak{q})])}{\inf\{\rho(t|\mathfrak{q}) : t \in \mathcal{I}_k\}}.\end{aligned} \tag{44}$$

Using (40) and (42), direct computation allows one to obtain the upper bound

$$\exp\left(\tfrac{\pi}{\zeta}[\Lambda(\tfrac{1}{2}|\mathfrak{q})-\Lambda(1-\tfrac{k}{2}|\mathfrak{q})]\right) \leq \frac{C_4(k-1)}{N/2-n-C_1k}\ .$$

Then, using that $\Lambda(n + \frac{k}{2}|\mathfrak{q}) = -\Lambda(1 - \frac{k}{2}|\mathfrak{q})$, the parity of ρ, Proposition 24(i) and the previously mentioned bound $K(0) \leq C_0/\zeta$, one deduces from (44) that

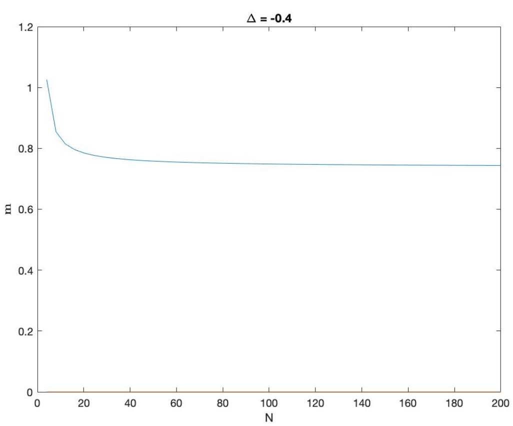
$$\begin{aligned} \frac{k}{N}\|f_j'\|_{L^1(\mathcal{I}_k)} &\leq \frac{k}{\zeta\rho\big(\Lambda(1-\frac{k}{2}|\mathfrak{q})\big)}\Big\{\zeta\|K'\|_{L^1(\mathbb{R})} + \frac{C_4\, k\|K\|_{L^1(\mathbb{R})}}{N/2 - n - C_1 k}\Big\} \\ &\leq C_5\Big(1 + \frac{k}{N/2 - n - C_1 k}\Big)\frac{k}{N/2 - n - C_1 k}. \end{aligned} \tag{45}$$

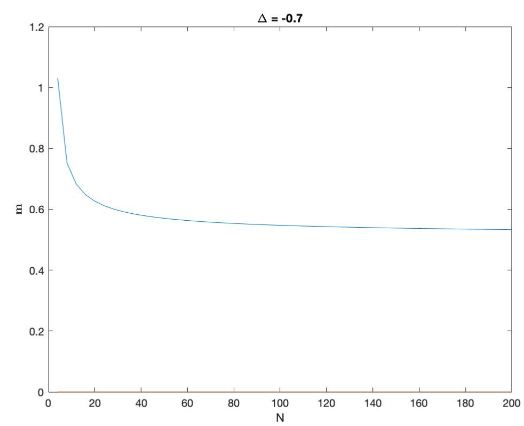
Plug (43) and (45) in (38) and (39), respectively, to find that

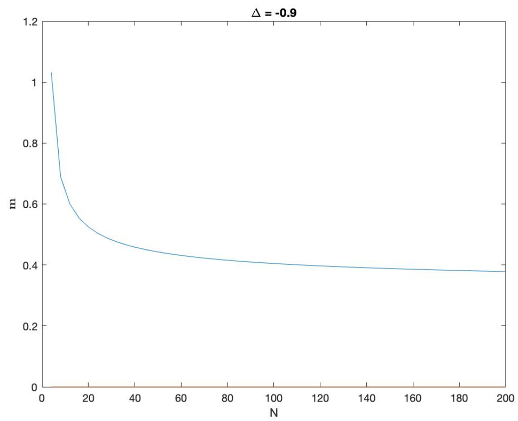
$$A_{jj} - \sum_{\ell \neq j} A_{j\ell} \geq \frac{1}{k+1} - C_6 \frac{k}{N/2 - n - C_1 k}.$$

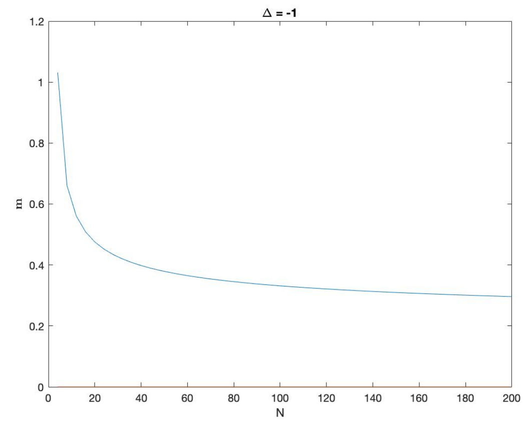
Taking C large enough in the statement of the proposition ensures that A is indeed diagonal dominant.

Remark 14. *The difficulty in proving that A_{ij} is diagonally dominant comes from the estimates involving j close to 1 or n as approximating sums by integrals is not efficient for these values of j. Another way of seeing this is that when j is far from 1 and n then $\rho(\lambda_j|\mathfrak{q})$ is larger and therefore the error term is smaller. Nonetheless, numerics suggest that the matrix is diagonal dominant for every $-1 < \Delta < 1$ and $n \leq N/2$, as shown on the plots of $\mathbf{m} := \min\{A_{ii} - \sum_{j\neq i} |A_{ij}| : 1 \leq i \leq n\}$ as a function of the system-size N, for $n = N/2$ and at four different values of Δ.*









Remark 15. *Since the differential is non-zero at $\Delta = 0$ (it is diagonal since $K \equiv 0$), we obtain in particular the existence of an analytic family of solutions for* every $n \leq N/2$ *for* $|\Delta| \leq \Delta_0$ *with* Δ_0 *small enough.*

3.3 Proof of Proposition 12 when $\Delta < -1$

Fix $\boldsymbol{\lambda}$ as in the proposition. Again, $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ is not obviously positive definite. At this point, we may use a symmetrization trick like in the proof of Theorem 3 for $\Delta < -1$. This argument was presented in [14] and we refer to this paper for a full proof. Here, we present an alternative proof that we find to be of some interest.

We show that $d\mathsf{T}\Delta(\boldsymbol{\lambda})$ is invertible by estimating the large N behaviour of $\det[d\mathsf{T}\Delta(\boldsymbol{\lambda})]$ with the help of Lemma 9. To start with, observe that

$$\widehat{\chi}(\lambda) := \xi(\lambda) - \frac{1}{N}\sum_{j=1}^{n} K(\lambda - \lambda_j) = \xi(\lambda) - \int_{-\mathfrak{q}}^{\mathfrak{q}} K(\lambda,\mu)\rho(\mu|\mathfrak{q})d\mu + O\Big(\frac{k}{N}\|K'(\lambda - \cdot)\|_{L^1(\mathcal{I}_k)}\Big)$$

$$= \rho(\lambda|\mathfrak{q}) \, + \, O\Big(\frac{k}{N}\Big)\,, \tag{46}$$

where $\mathcal{I}_k := [\Lambda\big(1-\frac{k}{2}|\mathfrak{q}\big), \Lambda\big(n+\frac{k}{2}|\mathfrak{q}\big)]$ is a subinterval of $[-\pi,\pi]$ uniformly bounded in $n \leq N/2$.

Since $\rho(\lambda|\mathfrak{q}) > 0$ on $\mathbb{R}$, the above ensures that the matrix

$$M_{ij} \;=\; \frac{K(\lambda_i - \lambda_j)}{\widehat{\chi}(\lambda_j)}$$

is well-defined for any n, N. Then, introduce an integral operator $\mathcal{M}$ acting on $L^2([-\mathfrak{q},\mathfrak{q}])$ with the integral kernel

$$M(\lambda,\mu) \;=\; \sum_{i,j=1}^{n} \mathbb{1}_{\mathcal{J}_i\times\mathcal{J}_j}(\lambda,\mu)\, M_{ij}\, \rho(\mu|\mathfrak{q}) \tag{47}$$

where $\mathscr{J}_i := \big[\Lambda\big(i-\frac{1}{2}|\mathfrak{q}\big), \Lambda\big(i+\frac{1}{2}|\mathfrak{q}\big)\big]$ is a partition of the integration domain $\cup_{i=1}^{n}\mathscr{J}_i = (-\mathfrak{q},\mathfrak{q}]$ as can be inferred from the identities

$$\mathfrak{q} = \Lambda\big(n+\tfrac{1}{2}|\mathfrak{q}\big) = -\Lambda\big(1-\tfrac{1}{2}|\mathfrak{q}\big) \ .$$

The matter is that the Fredholm determinant of $\mathrm{Id}+\mathcal{M}$ is equal to $\det_N[\mathrm{I}_N+M/N]$, where I_N is the identity matrix. Indeed, by writing the Fredholm expansion for the determinant, one has that

$$\begin{aligned}\operatorname*{det}_{L^2((-\mathfrak{q},\mathfrak{q}])}\big[\mathrm{Id}+\mathcal{M}\big] &= \sum_{\ell=1}^{+\infty}\frac{1}{\ell!}\int\limits_{-\mathfrak{q}}^{\mathfrak{q}} d^{\ell}\lambda \operatorname*{det}_{\ell}\big[M(\lambda_i,\lambda_j)\big] \\ &= \sum_{\ell=1}^{+\infty}\frac{1}{\ell!}\int\limits_{-\mathfrak{q}}^{\mathfrak{q}} d^{\ell}\lambda \sum_{\substack{i_1,\dots,i_\ell=1\\ j_1,\dots,j_\ell=1}}^{n}\prod_{s=1}^{n}\Big\{\mathbb{1}_{\mathscr{J}_{i_s}}(\lambda_s)\cdot\mathbb{1}_{\mathscr{J}_{j_s}}(\lambda_s)\rho(\lambda_s|\mathfrak{q})\Big\}\operatorname*{det}_{\ell}\big[M_{i_k j_u}\big].\end{aligned}$$

Since $\mathscr{J}_k\cap\mathscr{J}_\ell=\emptyset$ if $k\neq\ell$, by using that $\int\limits_{-\mathfrak{q}}^{\mathfrak{q}} d\mu\,\mathbb{1}_{\mathscr{J}_i}(\mu)\rho(\mu|\mathfrak{q}) = \frac{1}{N}$, one obtains

$$\operatorname*{det}_{L^2((-\mathfrak{q},\mathfrak{q}])}\big[\mathrm{Id}+\mathcal{M}\big] = \sum_{\ell=1}^{+\infty}\frac{1}{\ell!}\sum_{i_1,\dots,i_\ell=1}^{n}\operatorname*{det}_{\ell}\big[\tfrac{1}{N}M_{i_k i_u}\big] = \operatorname*{det}_{N}\big[\mathrm{I}_N+\tfrac{1}{N}M\big].$$

Introduce the integral operator $\mathcal{K}$ on $L^2((-\mathfrak{q},\mathfrak{q}])$ characterised by the integral kernel $K(\lambda-\mu)$. Both $\mathcal{M}$ and $\mathcal{K}$ are trace class. Indeed, $\mathcal{M}$ is of finite rank while $\mathcal{K}$ has smooth kernel and acting on functions supported on a compact interval [11]. Moreover, it holds

$$\begin{aligned}\mathrm{tr}\big[\mathcal{K}-\mathcal{M}\big] &= 2\mathfrak{q}K(0) - K(0)\frac{1}{N}\sum_{i=1}^{n}\frac{1}{\widehat{\chi}(\lambda_i)} \\ &= 2\mathfrak{q}K(0) - K(0)\Big\{\int\limits_{-\mathfrak{q}}^{\mathfrak{q}} d\mu\frac{\rho(\mu|\mathfrak{q})}{\widehat{\chi}(\mu)} + O\Big(\frac{k}{N}\Big\|\frac{\widehat{\chi}'}{\widehat{\chi}^2}\Big\|_{L^1(\mathcal{I}_k)}\Big)\Big\} = O\Big(\frac{k}{N}\Big) \ .\end{aligned}$$

There we used the estimate (46). We now estimate the Hilbert–Schmidt norm of $\mathcal{K}-\mathcal{M}$. One starts with the representation for the kernel

$$K(\lambda-\mu)-M(\lambda,\mu) = \sum_{i,j=1}^{n}\mathbb{1}_{\mathscr{J}_i\times\mathscr{J}_j}(\lambda,\mu)\Big\{K(\lambda-\mu)-K(\lambda_i-\lambda_j)\frac{\rho(\mu|\mathfrak{q})}{\widehat{\chi}(\lambda_j)}\Big\} \ .$$

By using that $\Lambda\big(x+k|\mathfrak{q}\big)-\Lambda\big(x|\mathfrak{q}\big) = O\big(k/N\big)$ uniformly in x, the mean-value theorem and $(k,\mathfrak{q})$-interlacement of $\boldsymbol{\lambda}$, one gets that

$$K(\lambda-\mu)-K(\lambda_i-\lambda_j) = O\Big(\frac{k+1}{N}\Big)$$

on $\mathcal{J}_i \times \mathcal{J}_j$. Then, the estimate (46) allows one to conclude that

$$\Big| K(\lambda-\mu) - M(\lambda,\mu) \Big| \,\leq\, C\frac{k+1}{N}\sum_{i,j=1}^{n} \mathbb{1}_{\mathcal{J}_i\times\mathcal{J}_j}(\lambda,\mu) = C\frac{k+1}{N}\ .$$

This yields an estimate on the Hilbert-Schmidt norm $\|\mathcal{K}-\mathcal{M}\|_{HS} = O\big(\frac{k+1}{N}\big)$ since $\mathfrak{q}$ is uniformly bounded in $n \leq N/2$. Thus, since both $\mathcal{K}, \mathcal{M}$ have finite Hilbert–Schmidt norms, it holds [25] that their 2-determinants satisfy

$$\det{}_2\big[\,\mathrm{Id}+\mathcal{M}\big] - \det{}_2\big[\,\mathrm{Id}+\mathcal{K}\big] \;=\; O\Big(\frac{k+1}{N}\Big)\ .$$

Since for a trace class operator $\mathcal{O}$ one has $\det_2\big[\,\mathrm{Id}+\mathcal{O}\big] \,=\, \det\big[\,\mathrm{Id}+\mathcal{O}\big]\mathrm{e}^{-\mathrm{tr}[O]}$, where the determinant appearing on the right-hand side is the usual Fredholm determinant, one infers that $\det\big[\,\mathrm{Id}+\mathcal{M}\big]-\det\big[\,\mathrm{Id}+\mathcal{K}\big] \;=\; O\big(\frac{k+1}{N}\big)$. Since $\mathrm{Id}+\mathcal{K}$ is invertible on $L^2([-q,q])$ for any $q\in[0,\pi/2]$ (this is for instance a consequence of the proofs of Propositions 23, 26, and 28), this entails that $\det\big[\,\mathrm{Id}+\mathcal{M}\big] \neq 0$ for N large enough.

4 Proof of Theorem 5

Let us start by stating two lemmata.

Lemma 16. *Let $\boldsymbol{\lambda}(0)$ be the unique solution to* (disc.BE) *when $\Delta = 0$. Then, one has $\Psi_N^{(n)}(\boldsymbol{\lambda}(0))_{|\Delta=0} \neq 0$ and $\Lambda_N^{(n)}(\boldsymbol{\lambda}(0))_{|\Delta=0}$ is the Perron-Frobenius eigenvalue of*

$$V_N^{(n)}\big(\sqrt{2}r\sin[\tfrac{\pi-\theta}{2}], \sqrt{2}r\sin[\tfrac{\theta}{2}], r\sqrt{2}\big)$$

for any r and θ.

Proof. For $\Delta = 0$, the unique solution to (disc.BE) is given by

$$\lambda_i(0) = \mathfrak{p}^{-1}_{|\Delta=0}\Big(2\pi\frac{i-(n+1)/2}{N}\Big).$$

It is then a matter of elementary computation to show that the entries of $\Psi_N^{(n)}$ are strictly positive, which concludes the proof of the lemma. □

Lemma 17. *Let $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ be an analytic solution of* (disc.BE) *defined on $(-\infty, v)$. Then, for Δ sufficiently negative, $\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is non-zero and $\Lambda_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is the Perron-Frobenius eigenvalue of $V_N^{(n)}(a,b,c)$.*

Proof of Lemma 17. Exceptionally, consider (disc.BE) with $\Delta = -\infty$ (see Appendix A). First, note that

$$\lambda_i(-\infty) := \pi\frac{i-(n+1)/2}{N-n}$$

for $1 \le i \le n$ is the k-interlaced, symmetric, strictly ordered solution of (disc.BE) with $\Delta = -\infty$ (simply note that $\theta_{-\infty}(\lambda) \equiv 2\lambda$ in this case). We deduce that if

$$\begin{aligned}
\psi^{(\infty)}(\vec{x}|\boldsymbol{\lambda}) &:= \lim_{\Delta\to-\infty} \frac{\psi(\vec{x}|\boldsymbol{\lambda})}{(-i\Delta)^{\frac{n(n-1)}{2}}} = \prod_{k=1}^{n} \mathrm{e}^{i(n+1)\lambda_k} \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} \mathrm{e}^{2i\lambda_{\sigma(k)}(x_k-k)},\\
V_N^{(n;\infty)} &:= \lim_{\Delta\to-\infty} \frac{V_N^{(n)}(a,b,c)}{(-2\Delta)^{\frac{N}{2}-\frac{\theta}{\pi}(N-2n)}},\\
\Lambda_N^{(n;\infty)} &:= r^N,
\end{aligned}$$

then the Bethe Ansatz at $\Delta = -\infty$ (or equivalently its limit as $\Delta \to -\infty$) implies that

$$V_N^{(n;\infty)}\Psi_N^{(n;\infty)} = \Lambda_N^{(n;\infty)}\Psi_N^{(n;\infty)} \qquad \text{with} \qquad \Psi_N^{(n;\infty)} := \sum_{|\vec{x}|=n} \psi^{(\infty)}(\vec{x}|\boldsymbol{\lambda}(-\infty))\, \boldsymbol{\Psi}_{\vec{x}}.$$

One gets the result by proving that $\Psi_N^{(n;\infty)}$ is non-zero and that $\Lambda_N^{(n;\infty)}$ is the largest eigenvalue of $V_N^{(n;\infty)}$, which we next do.

We start by showing the first claim by considering the entry of $\Psi_N^{(n;\infty)}$ for $\vec{x}_{\mathfrak{e}} = (2,4,\ldots,2n)$. Setting $\tau = \mathrm{e}^{2\pi i/(N-n)}$, one deduces from the above that

$$\psi^{(\infty)}(\vec{x}_{\mathfrak{e}}|\boldsymbol{\lambda}(-\infty)) = \tau^{-\frac{n}{4}(n+1)^2} \det\left[\tau^{j\cdot k}\right]_{j,k}.$$

The determinant of the Vandermonde matrix $(\tau^{j\cdot k})_{j,k}$ does not vanish since it corresponds to the values $\tau, \tau^2, \ldots, \tau^n$, which are all distinct owing to $2n \le N$. Hence, $\Psi_N^{(n;\infty)} \neq 0$.

For the second property, we refer to [14, Lemma 3.2] for the full proof. □

Finally, we are ready to prove the theorem.

Proof of Theorem 5. Consider an analytic family of solutions $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ as in the statement of the theorem. Then $\Delta \mapsto \Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is an analytic family of vectors. By Lemma 16 for $(u,v) = (\Delta_0, 1)$ or Lemma 17 for $(u,v) = (-\infty, \Delta_0)$, $\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta_0)) \neq 0$ for some $\Delta_0 \in (u,v)$. The analyticity implies that $\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta)) \neq 0$ for all but a discrete set D of Δ in (u,v).

It follows that $\Lambda_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is an eigenvalue of $V_N^{(n)}(a,b,c)$ for all $\Delta \in (u,v) \setminus D$. By continuity of $(a,b,c) \mapsto V_N^{(n)}(a,b,c)$ and $\Delta \mapsto \Lambda_N^{(n)}(\boldsymbol{\lambda}(\Delta))$, this property extends to all values $\Delta \in (u,v)$.

Now, since $V_N^{(n)}(a,b,c)$ is an irreducible symmetric matrix, its Perron-Frobenius eigenvalue is isolated for all a,b,c with $\Delta \in (u,v)$. Lemmata 16 and 17 proved that $\Lambda_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is the Perron-Frobenius eigenvalue for *some* $\Delta \in (u,v)$; by connexity and continuity of both $(a,b,c) \mapsto V_N^{(n)}(a,b,c)$ and $(a,b,c) \mapsto \Lambda_N^{(n)}$ this property extends to the whole set of parameters a,b,c with $\Delta \in (u,v)$. □

Remark 18. *The analyticity of $\Delta \mapsto \boldsymbol{\lambda}(\Delta)$ was only used once in the proof above, namely to show that the vector $\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ is non-zero for (almost) all Δ.*

It is highly non-trivial that this property hods for all *Δ, N and n. The norm of $\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta))$ has been argued to be given in terms of the determinant of $d\mathsf{T}_\Delta(\boldsymbol{\lambda})$ in [21, 29] and this was proven in [28, 39]. The results reads*

$$\|\Psi_N^{(n)}(\boldsymbol{\lambda}(\Delta))\|^2 = f(\boldsymbol{\lambda}) \det[d\mathsf{T}_\Delta(\boldsymbol{\lambda})] \tag{48}$$

for some explicit non-zero function f. Therefore, proving that the vector is non-zero amounts to proving that the differential of T_Δ is invertible which, as shown above, automatically implies analyticity.

In conclusion, proving analyticity of the solutions and using the strategy above, rather than proving their continuity and separately that the resulting vector is non-zero, bypasses the use of (48) *and contains no additional complications.*

5 Proof of Theorem 6

Let $k = C_0 \log N$ and $n \leq N/2 - C_1 k^2$. The constants C_0 and C_1 will be chosen large enough in the course of the proof.

To start, we follow the argument of Theorem 3 in Section 2.2 to guarantee that for each $-1 < \Delta < 0$, every $(k, \mathfrak{q})$-interlaced solution is strictly interlaced. For the proof to work, we need to check that

$$|\vartheta(\lambda_i - \lambda_1) - \vartheta(\lambda_i - \mathfrak{q})| \leq 2\pi \tfrac{k}{k+1}. \tag{49}$$

In order to do that, remark that extremum calculation and $(k, \mathfrak{q})$-interlacement imply that

$$\begin{aligned}
|\vartheta(\lambda_i - \lambda_1) - \vartheta(\lambda_i - \mathfrak{q})| &\leq 2|\vartheta\Big(\frac{\lambda_1 - \mathfrak{q}}{2}\Big)| \\
&\leq 2|\vartheta(\Lambda\big(1 - \tfrac{k}{2}|\mathfrak{q})\big)| \\
&= 4|\arctan\big[\tanh\big(\Lambda(1 - \tfrac{k}{2}|\mathfrak{q}\big)\cot(\zeta)\big]\big| \\
&\leq 2\pi + 4\arctan\left[\frac{\zeta}{\Lambda\big(1 - \frac{k}{2}|Q(\frac{1}{2} - C_1\frac{k^2}{N})\big)}\right] \\
&\leq 2\pi + \frac{2\zeta}{\Lambda\big(1 - \frac{k}{2}|Q(\frac{1}{2} - C_1\frac{k^2}{N})\big)} \\
&\leq 2\pi + \frac{C}{\log\big[\frac{C'}{N}(C_1k^2 - C''k)\big]}.
\end{aligned}$$

In the third line, we used that $\tanh(y) \leq y$ and $\cot(\zeta) \leq 1/\zeta$, and that $q \mapsto \Lambda(x|q)$ is decreasing for $x < (n+1)/2$ in this regime of Δ. In the last line, we used (42). Overall, we deduce that there exits a constant $c_3 > 0$ independent of everything such that for every $-1 < \Delta < 0$,

$$|\vartheta(\lambda_i - \lambda_1) - \vartheta(\lambda_i - \mathfrak{q})| \leq 2\pi - \frac{c_3}{\log N}.$$

We deduce (49) by fixing C_0 large enough. In particular, we obtain the equivalent of Theorem 3, namely that for every $-1 < \Delta < 0$, there exists a $(k, \mathfrak{q})$-interlaced strictly ordered symmetric solution $\boldsymbol{\lambda}(\Delta)$ to (disc.BE).

We now need to prove that the family $\boldsymbol{\lambda}(\Delta)$ can be assumed to be analytic. We follow the argument of Section 3.2, with minor changes which we describe next. As in Section 3.2 we may use (49) to deduce (38) and (39). It remains to bound the error terms. We start with (39). Since $N/2 - n - k \geq C_1 k^2 - k$, (45) leads to

$$\frac{k}{N}\|f_j'\|_{L^1(\mathcal{I}_k)} \leq \frac{k|2\pi - 4\zeta|}{c_0 c_3 (N/2 - n - Ck)} \leq \frac{C_4}{C_1 k - 1}.$$

This can be made smaller than $1/(4k)$ by choosing C large enough. The same argument applies to the error term in (38) and the proof follows.

Remark 19. *The uniform $(C_0 \log N, \mathfrak{q})$-interlacement shows that the entries of $\frac{1}{\zeta}\boldsymbol{\lambda}(\Delta)$ are bounded uniformly in Δ so that we may extract sub-sequential limits as ζ tends to 0 to approach $(C_0 \log N, \mathfrak{q})$-interlaced strictly ordered symmetric solutions $\boldsymbol{\lambda}(-1)$ to* (disc.BE) *for $\Delta = -1$. Here, one should use the convergence on compact subsets of $\mathbb{R}$, as $\Delta \to -1$, viz. $\zeta \to 0^+$, of $\mathfrak{p}(\cdot/\zeta)$, $\vartheta(\cdot/\zeta)$, $\frac{1}{\zeta}K(\cdot/\zeta)$ and $\frac{1}{\zeta}\xi(\cdot/\zeta)$ to $\mathfrak{p}_{|\Delta=-1}$, $\vartheta_{|\Delta=-1}$, $K_{|\Delta=-1}$ and $\xi_{|\Delta=-1}$ given in Appendix A.2.*

To prove an analogue of Theorem 3 for $\Delta = -1$, one may employ the bounds of Section 2.2 and rescale all variables by $1/\zeta$. This allows one to conclude that Ψ maps $\frac{1}{\zeta}\bar{\Omega}_{k,R}$ onto $\frac{1}{\zeta}\Omega_{k,R}$ and so, by taking the $\zeta \to 0$ limit, yields a Brouwer fixed point for the map $\Psi_{|\Delta=-1}$. The unique fixed point of this map coincides, by construction, with any sub-sequential limit of $\frac{1}{\zeta}\boldsymbol{\lambda}(\Delta)$ as ζ tends to 0. Thus, such sub-sequential limits are unique, which shows that $\frac{1}{\zeta}\boldsymbol{\lambda}(\Delta)$ does converge to a $(C_0 \log N, \mathfrak{q})$-interlaced, strictly ordered, symmetric solution $\boldsymbol{\lambda}(-1)$ to (disc.BE) *for $\Delta = -1$.*

6 Proof of Theorem 1

We prove the statement for $a > b$ and $\Delta \neq -1$. The results extends to all $a \geq b \geq 0$ and $c \geq 0$ with $\Delta < 1$, since the left and right sides of (6) are Lipschitz in each coordinate of (a, b, c). The particular expression for $\Delta = -1$ is obtained by taking the limit either from above or below in (6).

We start with the expression

$$Z(\mathbb{T}_{N,M}, a, b, c) = \sum_{n=0}^{N} Z^{(n)}(\mathbb{T}_{N,M}, a, b, c),$$

which gives

$$\lim_{M\to\infty} \tfrac{1}{NM} \log Z(\mathbb{T}_{N,M}, a, b, c) = \max_{n\le N/2} f_N^{(n)}(a,b,c) = f_N^{(N/2)}(a,b,c), \tag{50}$$

where in the second equality we restrict our attention to $n \le N/2$ thanks to the symmetry $n \longleftrightarrow N-n$ corresponding to the symmetry under the reversal of all arrows. The last equality uses the classical fact that $f_N^{(n)}(a,b,c)$ is maximal for $n = N/2$, see [14, Lemma 3.6] or [33] for a proof.

For $0 \le \Delta < 1$ or $\Delta < -1$, we can apply (19) to $n = N/2$ to get

$$\begin{aligned} f_N^{(N/2)}(a,b,c) = \max\Big\{ \log a + \int_{-Q(1/2)}^{Q(1/2)} \log|L(\lambda)|\rho(\lambda)d\lambda + O(\tfrac{1}{N})\ , \\ \log b + \int_{-Q(1/2)}^{Q(1/2)} \log|M(\lambda)|\rho(\lambda)d\lambda + O(\tfrac{1}{N}) \Big\}. \end{aligned} \tag{51}$$

We claim that the same holds for $-1 < \Delta < 0$. Notice however that we do not have access to $f_N^{(N/2)}(a,b,c)$ in this case. Yet, the inequality[7]

$$Z^{(n)}(\mathbb{T}_{N,M}, a, b, c) \ge \frac{1}{N}\Big(\frac{\min\{a,b,c\}}{4\max\{a,b,c\}}\Big)^{MN/n} Z^{(n+1)}(\mathbb{T}_{N,M}, a, b, c)$$

is sufficient to obtain (51) as we can apply (19) to $n := N/2 - C_0$, and then compare $f_N^{(N/2)}$ to $f_N^{(n)}$ and $\rho(\lambda)$ to $\rho(\lambda|Q(\frac{1}{2} - \frac{C_0}{N}))$.

Overall, we see that the only remaining difficulty is to compute

$$\log a + \int_{-Q(1/2)}^{Q(1/2)} \log|L(\lambda)|\rho(\lambda)d\lambda \quad \text{and} \quad \log b + \int_{-Q(1/2)}^{Q(1/2)} \log|M(\lambda)|\rho(\lambda)d\lambda.$$

We shall only focus on the evaluation of the first term and split the proof in two depending on whether $|\Delta| < 1$ or $\Delta < -1$. We leave to the reader the verification that the first term does indeed dominate the second when $a > b$.

6.1 Case $|\Delta| < 1$

When $|\Delta| < 1$, we use the Fourier transform on $\mathbb{R}$. Define

$$C(x) := \tfrac{1}{2}\log[L(x)L(-x)] = \log|L(x)| \qquad \text{for } x \in \mathbb{R} \tag{52}$$

[7] The displayed inequality can be obtained from the easy observation (already made in a number of papers, see e.g. [18]) that there exists a map from configurations with $n+1$ up arrows per row to configurations with n up arrows per row constructed by choosing a path of oriented edges cycling around the torus in the vertical direction with length smaller than MN/n (such a path always exists and there are at most $N4^{MN/n}$ choices for it) and reversing all arrows on it. This changes the weights, hence the factor $\min\{a,b,c\}/\max\{a,b,c\}$.

and observe the exact expressions given in Appendix A for the different functions and their Fourier transforms. Recalling that ρ is even and $Q(1/2) = +\infty$, we find that

$$f(a,b,c) - \log a = \int_{-\infty}^{+\infty} C(x)\rho(x)dx = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{C}(t)\widehat{\rho}(t)dt$$
$$= \int_{-\infty}^{+\infty} \frac{1}{2\cosh\left(\zeta\frac{t}{2}\right)}\frac{\sinh(\frac{\theta\zeta}{\pi}t)}{t}\frac{\sinh[(\pi-\zeta)\frac{t}{2}]}{\sinh\left[\pi\frac{t}{2}\right]}dt.$$

Using that

$$\log\frac{b}{a} = \int_{-\infty}^{+\infty} \frac{\sinh\left[t\zeta(\frac{\theta}{\pi}-\frac{1}{2})\right]\sinh\left[(\pi-\zeta)\frac{t}{2}\right]}{t\sinh\left[\pi\frac{t}{2}\right]}dt$$

some algebra and the change of variables $t \mapsto 2t$ give the result.

Remark 20. *For the special case $a = b = c = 1$, we may compute the integral directly. After a fairly elementary computation, we recover the classical result of Lieb [33]*

$$f(1,1,1) = \int_{-\infty}^{\infty} \frac{\frac{1}{2}\log[1-\frac{3}{1+2\cosh x}]}{\frac{8\pi}{3}\cosh(3x/4)}dx = \tfrac{3}{2}\log[\tfrac{4}{3}].$$

The particular expression for $a = b = 1$ and $c = 2$ is obtained from (6) *by direct computation.*

6.2 Case $\Delta < -1$

When $\Delta < -1$, we work with π-periodic functions and consider Fourier coefficients. Again, we introduce $C(x) = C(x) := \frac{1}{2}\log[L(x)L(-x)]$ and use the exact expression of the functions and their Fourier coefficients given in Appendix A. We deduce that

$$f(a,b,c) - \log a = \int_{-\pi/2}^{\pi/2} C(x)\rho(x)dx = \tfrac{1}{\pi}\sum_{n=0}^{\infty}\widehat{\rho}(n)\widehat{C}(n)$$
$$= \frac{\theta\zeta}{\pi} + \sum_{n\in\mathbb{Z}\setminus\{0\}} \mathrm{e}^{-|n|\zeta}\frac{\sinh\left[2n\zeta\theta/\pi\right]}{2n\cosh(\zeta n)}. \tag{53}$$

7 Proof of Theorem 2

7.1 Focusing on the asymptotic in the q variable

We claim that it suffices to estimate the asymptotic behaviour of

$$\delta f(q) := \int_{-Q(1/2)}^{Q(1/2)} C(x)\rho(x)dx - \int_{-q}^{q} C(x)\rho(x|q)dx \tag{54}$$

as $q \nearrow Q(1/2)$, with C defined in (52).

Indeed, (19) shows that $f(a,b,c) - f^{(n)}(a,b,c) = \delta f(Q(n/N)) + O(1/N)$. Recall that (19) is obtained from Theorem 4 whenever $\Delta \neq -1$ and (7) holds. For $\Delta = -1$, if (7) is satisfied, (19) may be deduced from Theorem 6.

The asymptotics of $Q(n/N)$ as n/N approaches $1/2$ are given by Propositions 25, 27 and Lemma 30 of Appendixes B, C and D, respectively, and read

$$\lim_{m\to 1/2} (\tfrac{1}{2} - m)\mathrm{e}^{Q(m)\frac{\pi}{\zeta}} =: C_\Delta \text{ for } \Delta \in [-1,1) \quad \text{and} \quad \lim_{m\to 1/2} \frac{1-2m}{\pi - 2Q(m)} = \rho(\tfrac{\pi}{2}) \text{ for } \Delta < -1$$

for some constant $C_\Delta > 0$.

The estimation of (54) is different for $-1 \leq \Delta < 1$ and $\Delta < -1$ and does not refer anymore to the discrete Bethe equation. Deriving Theorem 2 from the asymptotics of (54) obtained below and those for Q mentioned above is a matter of simple algebra, which we do not detail further.

Remark 21. *The constant C_Δ was explicitly computed in [12]. We do not need the precise value here and therefore work with this weaker and simpler result. We provide however in Proposition 25 of Appendix B an integral representation for C_Δ in terms of the solution to a Wiener-Hopf equation on $\mathbb{R}_+$.*

7.2 Case $|\Delta| < 1$

Consider the function G defined for $x \in \mathbb{R}$ by

$$G(x) := C(x) - \int_{\mathbb{R}} R(x-y)C(y)dy.$$

Using (20), then reorganizing the integrals (in particular using that C and R are even), and then passing to Fourier gives that

$$
\begin{aligned}
\delta f(q) &= \int_{\mathbb{R}} C(x)\rho(x|q)\mathbb{1}_{|x|>q}dx - \int_{\mathbb{R}}\int_{\mathbb{R}} C(x)R(x-y)\rho(y|q)\mathbb{1}_{|y|>q}dxdy \\
&= \int_{\mathbb{R}} G(x)\rho(x|q)\mathbb{1}_{|x|>q}dx \\
&= \int_{\mathbb{R}} \widehat{G}(t)\widehat{\rho(\cdot|q)}(t)\frac{dt}{2\pi} - \int_{\mathbb{R}}\int_{\mathbb{R}} \widehat{G}(t)\frac{\mathrm{e}^{iq(t-s)}-\mathrm{e}^{iq(s-t)}}{i(t-s)}\widehat{\rho(\cdot|q)}(s)\frac{ds}{2\pi}\frac{dt}{2\pi} \\
&= \int_{\mathbb{R}} \widehat{G}(t)\widehat{\rho(\cdot|q)}(t)\frac{dt}{2\pi} - \lim_{\delta\searrow 0}\int_{\mathbb{R}+i\delta}\Big(\int_{\mathbb{R}} \widehat{G}(t)\frac{\mathrm{e}^{iq(t-s)}-\mathrm{e}^{iq(s-t)}}{i(t-s)}\frac{dt}{2\pi}\Big)\widehat{\rho(\cdot|q)}(s)\frac{ds}{2\pi}.
\end{aligned}
$$

In the last identity we use that

$$
\widehat{G}(k) = \frac{\widehat{C}(k)}{1+\widehat{K}(k)} = \frac{\pi}{k}\frac{\sinh\left[k\zeta\frac{\theta}{\pi}\right]}{\cosh\left[k\frac{\zeta}{2}\right]}
$$

is integrable on a neighbourhood of $\mathbb{R}$ and has exponential decay. We now perform two elementary residue computations. Fix $\pi/\zeta < \beta < 3\pi/\zeta$ and $s \in i\delta + \mathbb{R}$ for $\delta > 0$ very small. Since $\mathrm{Res}_{t=\pm i\pi/\zeta}[\widehat{G}] = \mp 2i\sin\theta$ and there is no other pole of $\widehat{G}$ in the strip $\{z \in \mathbb{C} : \mathrm{Im}(z) < \beta\}$ (and $\widehat{G}$ tends to 0 at infinity), we get

$$
\begin{aligned}
\int_{\mathbb{R}} \frac{\mathrm{e}^{iq(t-s)}}{t-s}\widehat{G}(t)\frac{dt}{2i\pi} &= \widehat{G}(s) + \frac{\mathrm{e}^{iq(i\pi/\zeta-s)}}{i\pi/\zeta-s}\mathrm{Res}_{t=i\pi/\zeta}[\widehat{G}] + \int_{\mathbb{R}+\beta i}\frac{\mathrm{e}^{iq(t-s)}}{i(t-s)}\widehat{G}(t)\frac{dt}{2\pi} \\
&= \widehat{G}(s) - 2i\sin\theta\,\frac{\mathrm{e}^{-q\pi/\zeta-iqs}}{i\pi/\zeta-s} + \mathrm{e}^{-\beta q}\psi_q^{(+)}(s)
\end{aligned}
$$

where

$$
\psi_q^{(\pm)}(s) = \pm\int_{\mathbb{R}} \frac{\mathrm{e}^{\pm iq(t-s)}\widehat{G}(t\pm i\beta)}{t-s\pm i\beta}\frac{dt}{2i\pi} \tag{55}
$$

and similarly

$$
-\int_{\mathbb{R}} \frac{\mathrm{e}^{iq(s-t)}}{t-s}\widehat{G}(t)\frac{dt}{2\pi} = -2i\sin\theta\,\frac{\mathrm{e}^{-q\pi/\zeta+iqs}}{i\pi/\zeta+s} + \mathrm{e}^{-\beta q}\psi_q^{(-)}(s)
$$

Putting these two displayed equations in the first one gives that

$$
\begin{aligned}
\delta f(q) = 2i\sin\theta\,\mathrm{e}^{-\frac{q\pi}{\zeta}}\int_{\mathbb{R}}\widehat{\rho(\cdot|q)}(s)\left(\frac{\mathrm{e}^{-iqs}}{i\pi/\zeta-s}+\frac{\mathrm{e}^{iqs}}{i\pi/\zeta+s}\right)\frac{ds}{2\pi} \\
- \mathrm{e}^{-\beta q}\int_{\mathbb{R}}\widehat{\rho(\cdot|q)}(s)(\psi_q^{(-)}(s)+\psi_q^{(-)}(s))ds\ .
\end{aligned} \tag{56}
$$

We first justify that the second term is a $O(\mathrm{e}^{-\beta q})$. Clearly, $\psi_q^{(\pm)} \in L^2(\mathbb{R})$ and $\|\psi_q^{(\pm)}\|_{L^2(\mathbb{R})} \leq C$ uniformly in q. Furthermore, it is established in Proposition 25 of Appendix B that given

$$\mathfrak{e}(x) := \frac{1}{\zeta}\mathrm{e}^{-\frac{\pi}{\zeta}x}\mathbb{1}_{\mathbb{R}_+}(x),$$

we find that

$$\rho(x|q) = \rho(x) \;+\; \mathrm{e}^{-\frac{q\pi}{\zeta}}\big[(T-\mathfrak{e})(x-q) + (T-\mathfrak{e})(-q-x) + \delta T(x)\big] \tag{57}$$

where $\|\delta T\|_{L^\infty(\mathbb{R})} + \|\delta T\|_{L^1(\mathbb{R})} = O(e^{-2q})$ and T is the unique solution of the integral equation

$$T(x) - \int_0^\infty R(x-y)T(y)dy = \mathfrak{e}(x). \tag{58}$$

By interpolation theorems for L^p spaces, we get that T and δT belong to $L^2(\mathbb{R})$ with norms uniformly bounded in q. Therefore, $\rho(\cdot|q) \in L^2(\mathbb{R})$ with a norm controlled uniformly in q. All of this ensures that the last term in (56) is indeed $O(\mathrm{e}^{-\beta q})$.

Since $\widehat{\rho(\cdot|q)}$ is even and $\frac{1}{\pi/\zeta+is}$ is the Fourier transform of $\mathfrak{e}$, we find

$$\int_{\mathbb{R}} \widehat{\rho(\cdot|q)}(s)\left(\frac{\mathrm{e}^{-iqs}}{\pi/\zeta+is} + \frac{\mathrm{e}^{iqs}}{\pi/\zeta-is}\right)\frac{ds}{2\pi} = 2\int_0^\infty \rho(x+q|q)\mathrm{e}^{-x\frac{\pi}{\zeta}}dx. \tag{59}$$

Plugging $x+q$ in (57), and performing an asymptotic development of (57) (at first order) enables us to recast (56) as

$$\delta f(q) = 4\sin\theta\,\mathrm{e}^{-q\frac{\pi}{\zeta}}\int_0^\infty \mathrm{e}^{-x\frac{\pi}{\zeta}}\Big\{\rho(x+q) + \mathrm{e}^{-q\frac{\pi}{\zeta}}\big[T(x) - \mathfrak{e}(x) + T(-x-2q) + \delta T(x+q)\big]\Big\}dx$$
$$+\, o\big(\mathrm{e}^{-2q\frac{\pi}{\zeta}}\big),$$

where we used the fact that β can be taken strictly larger than $2\pi/\zeta$. Then, by using

- $\rho(x+q) = \mathfrak{e}(x+q)\big(1+\mathrm{e}^{-\frac{2\pi}{\zeta}q}\delta\mathfrak{e}(x)\big)$ with $\|\delta\mathfrak{e}\|_{L^1(\mathbb{R}_+)}$ bounded uniformly in q,
- $\delta T \in L^\infty(\mathbb{R})$,
- the fact that as λ tends to infinity,

$$T(\lambda) = O\Big(\max\big\{\mathrm{e}^{-\frac{2\pi}{\pi-\zeta}\lambda}, \mathrm{e}^{-\frac{\pi}{\zeta}\lambda}\big\}\Big).$$

  which follows from the integral equation satisfied by T, the fact that $T \in L^1(\mathbb{R})$ and similar estimates for the behaviour of $R(\lambda)$ when $\lambda \to \pm\infty$, one readily infers that

$$\delta f(q) = 4\sin\theta\, \mathrm{e}^{-2q\frac{\pi}{\zeta}} \cdot \int_0^\infty \mathrm{e}^{-x\frac{\pi}{\zeta}} T(x)dx \;+\; o\big(\mathrm{e}^{-2q\frac{\pi}{\zeta}}\big).$$

Thus, all-in-all, we get that with $\mathfrak{q} = Q(n/N)$,

$$\delta f(\mathfrak{q}) = (1+o(1))\big(\tfrac{1}{2} - \tfrac{n}{N}\big)^2 \sin\theta \cdot \tfrac{4}{C_\Delta^2} \int_0^\infty \mathrm{e}^{-x\frac{\pi}{\zeta}} T(x)dx.$$

Note that the constant is strictly positive since $T > 0$ on $\mathbb{R}$.

7.3 Case $\Delta = -1$

We omit the proof as it is the same as in the previous section, using Appendix C instead of Appendix B.

7.4 Case $\Delta < -1$

Following the same reasoning as in the previous section gives

$$\begin{aligned}\delta f(q) :=& \int_{-\pi/2}^{\pi/2} C(x)\rho(x)dx - \int_{-q}^{q} C(x)\rho(x|q)dx \\ =&(\pi - 2q)\rho(\tfrac{\pi}{2})\big[C(\tfrac{\pi}{2}) - \int_{-\pi/2}^{\pi/2} C(x)R(x-\tfrac{\pi}{2})dx + o(1)\big].\end{aligned}$$

It remains to prove that the following quantity is strictly positive:

$$\begin{aligned}C(\tfrac{\pi}{2}) - \int_{-\pi/2}^{\pi/2} C(x)R(x-\tfrac{\pi}{2})dx &= \frac{1}{\pi}\sum_{n\in\mathbb{Z}}(-1)^n\widehat{C}(n) - \frac{1}{\pi}\sum_{n\in\mathbb{Z}}(-1)^n\widehat{C}(n)\widehat{R}(n) \\ &= \frac{1}{\pi}\sum_{n\in\mathbb{Z}}(-1)^n\frac{\widehat{C}(n)}{1+\widehat{K}(n)} \\ &= \sum_{n\in\mathbb{Z}}(-1)^n\frac{\sinh\left[2n\zeta\frac{\theta}{\pi}\right]}{2n\cosh(n\zeta)}.\end{aligned}$$

Positivity then follows from the Parseval identity combined with the fact that the inverse Fourier transforms of both $\tanh[n\zeta]/(2|n|)$ and $\sinh\left[2n\zeta\frac{\theta}{\pi}\right]/\sinh\left[n\zeta\right]$ are positive.

8 A refined version of Theorem 2 (under additional conditions)

In this section, we prove a sharper version of Theorem 2 under mild conditions on a, b, c and n, N.

Theorem 22. *For $N \geq 2$ and $a \neq b$ and $c \geq 0$ leading to $|\Delta| < 1$, there exists a constant $C = C(\theta) < \infty$ such that for every $n \leq \frac{1}{2}N - C/\zeta$,*

$$f_N^{(n)}(a,b,c) = f(a,b,c) - C(\zeta)\sin\theta(1+o(1))(1-\tfrac{2n}{N})^2 + O(\tfrac{1}{\zeta(N-2n)N}), \tag{60}$$

where $o(1)$ is a quantity tending to zero as n/N tends to $1/2$.

The improvement with respect to Theorem 2 is that the $O(1/N)$ is replaced with the more precise $O(\frac{1}{\zeta(N-2n)N})$. This will be particularly useful in [17], where it is used to prove that $N|f_N^{(n)}(a,b,c) - f_N^{(n+1)}(a,b,c)| \to 0$ as $N \to \infty$, with $n \leq \sqrt{N}$. This convergence is expected to hold for all $n = o(N)$, when $\Delta \in [-1, 1)$.

Proof of Theorem 22. Due to the form of (60), it suffices to prove the statement for N large enough. Fix for now n and N as in the theorem; we will see later which bounds are needed on N.

Consider the analytic family $\boldsymbol{\lambda} = (\lambda_i : 1 \leq i \leq n)$ given by Theorem 4 on $(\Delta_0, 1)$. We start by improving on the condensation formula of Theorem 3. Introduce, for a fixed Δ, the quantity $\mathrm{Diff} : [1, \ldots, n] \to \mathbb{R}$ defined by

$$\mathrm{Diff}(i) := N \int_{\Lambda(i|\mathfrak{q})}^{\lambda_i} \rho(\lambda|\mathfrak{q})\, d\lambda,$$
$$\mathrm{Diff} := \max\{|\,\mathrm{Diff}(i)| : 1 \leq i \leq n\}.$$

Claim 1 *There exists $C_0 > 0$ such that for every $f : \mathbb{R} \to \mathbb{R}$ with integrable first and second derivatives, and for every $\boldsymbol{\lambda} = (\lambda_i : 1 \leq i \leq n)$ with $\mathrm{Diff} \leq \frac{1}{2}$,*

$$\Big|\frac{1}{N}\sum_{j=1}^{n} f(\lambda_j) - \int_{-\mathfrak{q}}^{\mathfrak{q}} f(\lambda)\rho(\lambda|\mathfrak{q})\, d\lambda\Big| \leq \frac{\mathrm{Diff}}{N}\|f'\|_{L^1[-\mathfrak{q},\mathfrak{q}]} + \frac{C_0(1+\mathrm{Diff})}{N(N-2n-C_0)}(\zeta\|f''\|_{L^1[-\mathfrak{q},\mathfrak{q}]} + \|f'\|_{L^1[-\mathfrak{q},\mathfrak{q}]}).$$

Proof. Using that the integral of $\rho(\lambda|\mathfrak{q})$ between $\Lambda(j-\frac{1}{2}|\mathfrak{q})$ and $\Lambda(j+\frac{1}{2}|\mathfrak{q})$ is $\frac{1}{N}$ gives

$$\frac{1}{N}\sum_{j=1}^{n} f(\lambda_j) - \int_{-\mathfrak{q}}^{\mathfrak{q}} f(\lambda)\rho(\lambda|\mathfrak{q})d\lambda = \sum_{j=1}^{n}\int_{\Lambda(j-1/2|\mathfrak{q})}^{\Lambda(j+1/2|\mathfrak{q})} (f(\lambda_j) - f(\lambda))\rho(\lambda|\mathfrak{q})d\lambda.$$

Differentiating the definition of $\Lambda(x|\mathfrak{q})$ gives

$$\Lambda'(y|\mathfrak{q}) = \frac{1}{N\rho(\Lambda(y|\mathfrak{q})|\mathfrak{q})}.$$

Therefore, if we set $I_j := [j-\frac{1}{2}, j+\frac{1}{2}]$ and $g(y) := f(\Lambda(y|\mathfrak{q}))$, a change of variables implies, for every j,

$$\begin{aligned}\Big|\int_{\Lambda(j-1/2|\mathfrak{q})}^{\Lambda(j+1/2|\mathfrak{q})} (f(\Lambda(j|\mathfrak{q})) - f(\lambda))\rho(\lambda|\mathfrak{q})d\lambda\Big| &= \frac{1}{N}\Big|\int_{j-1/2}^{j+1/2}(g(j)-g(x))dx\Big| \\ &\leq \frac{1}{4N}\|g''\|_{L^1(I_j)}\end{aligned}$$

and, since $\lambda_j \in [\Lambda(j-\frac{1}{2}|\mathfrak{q}), \Lambda(j+\frac{1}{2}|\mathfrak{q})]$ thanks to $|\operatorname{Diff}(j)| \leq 1/2$,

$$\begin{aligned}\int_{\Lambda(j-1/2|\mathfrak{q})}^{\Lambda(j+1/2|\mathfrak{q})} |f(\lambda_j) - f(\Lambda(j|\mathfrak{q}))|\rho(\lambda|\mathfrak{q})d\lambda &= \frac{1}{N}|f(\lambda_j) - f(\Lambda(j|\mathfrak{q}))| \\ &\leq \frac{|\operatorname{Diff}(j)|}{N}\|g'\|_{L^\infty(I_j)} \\ &\leq \frac{|\operatorname{Diff}(j)|}{N}(\|g'\|_{L^1(I_j)} + \|g''\|_{L^1(I_j)}).\end{aligned}$$

Summing these estimates on j gives

$$\Big|\frac{1}{N}\sum_{j=1}^{n} f(\lambda_j) - \int_{-\mathfrak{q}}^{\mathfrak{q}} f(\lambda)\rho(\lambda|\mathfrak{q})\,d\lambda\Big| \leq \frac{\operatorname{Diff}}{N}\|g'\|_{L^1([1/2,n+1/2])} + \frac{\frac{1}{4}+\operatorname{Diff}}{N}\|g''\|_{L^1([1/2,n+1/2])}.$$

To conclude, observe that

$$\|g'\|_{L^1([1/2,n+1/2])} = \|f'\|_{L^1([-\mathfrak{q},\mathfrak{q}])}$$

and

$$\|g''\|_{L^1([1/2,n+1/2])} \leq \Big\|\frac{1}{N\rho(\cdot|\mathfrak{q})}\Big\|_{L^\infty([-\mathfrak{q},\mathfrak{q}])}\|f''\|_{L^1([-\mathfrak{q},\mathfrak{q}])} + \Big\|\frac{\rho'(\cdot|\mathfrak{q})}{N\rho(\cdot|\mathfrak{q})^2}\Big\|_{L^\infty([-\mathfrak{q},\mathfrak{q}])}\|f'\|_{L^1([-\mathfrak{q},\mathfrak{q}])}$$

which, combined with the bounds

$$\zeta N\rho(x|\mathfrak{q}) \geq c_0(N-2n-C_0)$$

(which is obtained from (42), the monotonicity of ρ on $(-\infty, 0]$ and the assumption $\operatorname{Diff} \leq 1/2$ which implies interlacement of $\boldsymbol{\lambda}$ with $k=1$) and $|\rho'(x|\mathfrak{q})| \leq \frac{C_1}{\zeta}\rho(x|\mathfrak{q})$ (Proposition 24(ii)) on $[-\mathfrak{q},\mathfrak{q}]$, gives the claim. □

Claim 2 *There exists an absolute constant* $C_2 \in (0,\infty)$ *such that for every* $\Delta \in (-1,1)$ *and every* $n \le N/2 - C_2/\zeta$,

$$\mathrm{Diff} \le \frac{C_2}{\zeta(N-2n-C_0)}.$$

Proof. The constant C_2 will be chosen at the end of the proof; it will be apparent that it is independent of n or N. For $\Delta = 0$, the result is obvious as the explicit (and unique) solution of the discrete Bethe Equation satisfies $\mathrm{Diff} = 0$.

Assume that there exists $\Delta \in (-1,1)$ such that $\mathrm{Diff} = \frac{C_2}{\zeta(N-2n)} \le \frac{1}{2}$. Using (30) in the first equality and then (disc.BE) in the second, we find

$$\begin{aligned}\mathrm{Diff}(i) &= \frac{N}{2\pi}\mathfrak{p}(\lambda_i) - \int_{-\mathfrak{q}}^{\mathfrak{q}} \vartheta(\lambda_i-\mu)\rho(\mu|\mathfrak{q})d\mu - I_i\\ &= \frac{1}{2\pi}\sum_{j=1}^{n}\vartheta(\lambda_i-\lambda_j) - \frac{N}{2\pi}\int_{-\mathfrak{q}}^{\mathfrak{q}}\vartheta(\lambda_i-\mu)\rho(\mu|\mathfrak{q})\,d\mu.\end{aligned}$$

Now, we use that for $K = \frac{1}{2\pi}\vartheta'$, $|K'| \le \frac{C}{\zeta}|K|$ and $\|K\|_{L^1[\mathbb{R}]} = 1 - \frac{2\zeta}{\pi}$ (see the Appendix again). Apply Claim 1 to $\frac{1}{2\pi}\vartheta(\lambda_i - x)$ (and bound the L^1 norm on $[-\mathfrak{q},\mathfrak{q}]$ by the L^1 norm on $\mathbb{R}$) to get

$$\begin{aligned}|\,\mathrm{Diff}(i)| &\le \mathrm{Diff}\cdot\|K\|_{L^1(\mathbb{R})} + \frac{C_0(1+\mathrm{Diff})}{N-2n-C_0}(\zeta\|K'\|_{L^1(\mathbb{R})} + \|K\|_{L^1(\mathbb{R})}),\\ &\le (1-\tfrac{2\zeta}{\pi})\mathrm{Diff} + \frac{C_0'}{N-2n-C_0},\end{aligned}$$

where C_0 is the constant given by Claim 1, and C_0' depends on C_0, but not on C_2. Since this applies to all i, we conclude that

$$\mathrm{Diff} \le \frac{\pi C_0'}{2\zeta(N-2n-C_0)}. \tag{61}$$

Choose now C_2 so that $C_2 > \frac{\pi}{2}C_0'$. Then (61) contradicts our assumption on Diff, and we conclude that there exists no $\Delta \in (-1,1)$ with $\mathrm{Diff} = \frac{C_2}{\zeta(N-2n-C_0)}$. By the continuity of Diff as a function of Δ and considering the fact that $\mathrm{Diff} = 0$ for $\Delta = 0$, we conclude that $\mathrm{Diff} < \frac{C_2}{\zeta(N-2n-C_0)}$ for all $\Delta \in (-1,1)$. □

We are now in a position to conclude the proof of Theorem 22. Let $C = C_2$ be given by Claim 2 and fix a,b,c as in the theorem. By taking N large enough, we may assume that the value Δ corresponding to (a,b,c) is contained in the domain in which $\boldsymbol{\lambda}$ is defined for any $n \le N/2 - C/\zeta$ (see Theorem 4). Then (19) states that

$$\tfrac{1}{N}\log\Lambda_N^{(n)}(\theta) = \tfrac{1}{N}\sum_{j=1}^{n} C(\lambda_j) + O(e^{-cN}),$$

where $C(.)$ is the function defined in (52). Claims 1 and 2 give

$$\begin{aligned}\Big|\tfrac{1}{N}\log\Lambda_N^{(n)}-&\int_{-\mathfrak{q}(\frac{n}{N})}^{\mathfrak{q}(\frac{n}{N})}C(\lambda)\rho(\lambda|\mathfrak{q}(\tfrac{n}{N}))d\lambda\Big|\\ &\le\frac{C_3}{\zeta N(N-2n-C_0)}\|C'\|_{L^1(\mathbb{R})}+\frac{C_3}{\zeta N(N-2n-C_0)}(\zeta\|C''\|_{L^1(\mathbb{R})}+\|C'\|_{L^1(\mathbb{R})})\\ &\le\frac{C_0}{\zeta N(N-2n-C_0)}.\end{aligned}$$

Furthermore, Sections 7.1 and 7.2 give that

$$\int_{-\mathfrak{q}(\frac{n}{N})}^{\mathfrak{q}(\frac{n}{N})}C(\lambda)\rho(\lambda)d\lambda=f(a,b,c)-C(\Delta)(1+o(1))\sin\theta(1-\tfrac{2n}{N})^2.$$

The above implies (60) by choosing C large enough. □

A formulae for the different functions and their Fourier transforms

Recall the parameterisations (2), (3) and (4) of the weights a, b, c. We remind that we always assume that $a \ge b > 0$ which corresponds to $\theta \in (0, \pi/2]$.

A.1 Case $|\Delta| < 1$

If $\zeta := \arccos(-\Delta) \in (0, \pi)$, we have for $x \in \mathbb{R}$, using the principal branch of the logarithm,

$$\begin{aligned}\mathfrak{p}(x)&:=i\log\frac{\sinh(i\zeta/2+x)}{\sinh(i\zeta/2-x)},\\ \vartheta(x)&:=i\log\frac{\sinh(i\zeta+x)}{\sinh(i\zeta-x)},\\ K(x)&:=\tfrac{1}{2\pi}\vartheta'(x)=\tfrac{1}{2\pi}\frac{\sin(2\zeta)}{\sinh(x+i\zeta)\sinh(x-i\zeta)},\\ \xi(x)&:=\tfrac{1}{2\pi}\mathfrak{p}'(x)=\tfrac{1}{2\pi}\frac{\sin(\zeta)}{\sinh(x+i\zeta/2)\sinh(x-i\zeta/2)},\\ \rho(x)&:=\frac{1}{2\zeta\cosh(\pi x/\zeta)},\\ C(x)&:=\tfrac{1}{2}\log[L(x)L(-x)].\end{aligned}$$

Moreover, the following direct consequences of the formulas above are used in the text: $\mathfrak{p}$ is strictly increasing, odd and $\mathfrak{p}(\mathbb{R}) = (-\pi + \zeta/2, \pi - \zeta/2)$; ϑ is decreasing for $\Delta > 0$, increasing for $\Delta < 0$ and constant for $\Delta = 0$; and finally $\vartheta(\mathbb{R}) = (-|\pi - \zeta|, |\pi - \zeta|)$.

We will also use the following Fourier transforms (with the relevant continuous extension at $t = 0$ when needed):

$$\widehat{K}(t) = \frac{\sinh[(\pi - 2\zeta)t/2]}{\sinh[\pi t/2]},$$
$$\widehat{\xi}(t) = \frac{\sinh[(\pi - \zeta)t/2]}{\sinh[\pi t/2]},$$
$$\widehat{\rho}(t) = \frac{1}{2\cosh(\zeta t/2)},$$
$$\widehat{C}(t) = \frac{2\pi \sinh(\frac{\theta}{\pi}\zeta t)}{t}\widehat{\xi}(t).$$

A.2 Case $\Delta = -1$

We have

$$\mathfrak{p}(x) := i \log\left(\frac{i/2 + x}{i/2 - x}\right),$$
$$\vartheta(x) := i \log\left(\frac{i + x}{i - x}\right),$$
$$K(x) := \tfrac{1}{2\pi}\vartheta'(x) = \frac{1}{\pi(1 + x^2)},$$
$$\xi(x) := \tfrac{1}{2\pi}\mathfrak{p}'(x) = \frac{2}{\pi(1 + 4x^2)},$$
$$\rho(x) := \frac{1}{2\cosh[\pi x]},$$
$$C(x) := \tfrac{1}{2}\log[L(x)L(-x)] = \tfrac{1}{2}\log\left[\frac{x^2 + \left(\frac{1}{2} + \frac{\theta}{\pi}\right)^2}{x^2 + \left(\frac{1}{2} - \frac{\theta}{\pi}\right)^2}\right].$$

We will also be interested in the following Fourier coefficients (with relevant extensions at $t = 0$)

$$\widehat{K}(t) = e^{-|t|},$$
$$\widehat{\xi}(t) = e^{-|t|/2},$$
$$\widehat{\rho}(t) = \frac{\widehat{\xi}(t)}{1 + \widehat{K}(t)} = \frac{1}{2\cosh[t/2]},$$
$$\widehat{C}(t) = \pi \cdot \frac{e^{-\frac{a-b}{2c}\cdot|t|} \cdot (1 - e^{-|t|})}{|t|}.$$

A.3 Case $\Delta < -1$

We have

$$\begin{aligned}
K(x) &:= \tfrac{1}{2\pi}\vartheta'(x) = \tfrac{1}{2\pi}\frac{\sinh(2\zeta)}{\sin(x+i\zeta)\sin(x-i\zeta)},\\
\xi(x) &:= \tfrac{1}{2\pi}\mathfrak{p}'(x) = \tfrac{1}{2\pi}\frac{\sinh\zeta}{\sin(x+i\zeta/2)\sin(x-i\zeta/2)},\\
\rho(x) &:= \tfrac{1}{2\pi}\sum_{n\in\mathbb{Z}}\frac{e^{2inx}}{\cosh[n\zeta]} = \tfrac{1}{2\zeta}\sum_{n\in\mathbb{Z}}\frac{1}{\cosh[\pi(\pi n-x)/\zeta]},\\
C(x) &:= \tfrac{1}{2}\log[L(x)L(-x)].
\end{aligned}$$

The functions $\mathfrak{p}$ and ϑ are then defined as the odd smooth functions on $\mathbb{R}$ that have $\frac{1}{2\pi}\xi$ and $\frac{1}{2\pi}K$ as derivatives. In particular, on $(-\frac{\pi}{2},\frac{\pi}{2})$, they are equal to

$$\begin{aligned}
\mathfrak{p}(x) &:= i\ln\frac{\sin(i\zeta/2+x)}{\sin(i\zeta/2-x)},\\
\vartheta(x) &:= i\ln\frac{\sin(i\zeta+x)}{\sin(i\zeta-x)}.
\end{aligned}$$

Moreover, the following direct consequences of the formulas above are used in the text: $\mathfrak{p}$ is increasing and maps $\mathbb{R}$ to $\mathbb{R}$; ϑ is increasing; $\vartheta([-\pi/2,\pi/2]) = [\pi,\pi]$ and ϑ extends to $\mathbb{R}$ as a quasi-periodic continuous function. The function K is even, unimodal, and has zero limits at $\pm\infty$.

We stress that these formulae do not extend, *per se*, beyond $(-\frac{\pi}{2},\frac{\pi}{2})$. We will also be interested in the following Fourier coefficients, when $2\theta/\pi < 1$,

$$\begin{aligned}
\widehat{K}(n) &= e^{-2|n|\zeta},\\
\widehat{\xi}(n) &= e^{-|n|\zeta},\\
\widehat{\rho}(n) &= \frac{\widehat{\xi}(n)}{1+\widehat{K}(n)} = \frac{1}{2\cosh[n\zeta]},\\
\widehat{C}(n) &= \frac{\pi}{2|n|}(e^{-|n|\cdot\left|1-\frac{2\theta}{\pi}\right|\zeta} - e^{-|n|(1+\frac{2\theta}{\pi})\zeta}) = \frac{\pi}{n}\mathrm{e}^{-|n|\zeta}\sinh\left[2\zeta n\tfrac{\theta}{\pi}\right]
\end{aligned}$$

if $n \neq 0$, and $\widehat{C}(0) = \pi\zeta$.

A.4 Case $\Delta = -\infty$

We have

$$\begin{aligned}
\mathfrak{p}(x) &:= 2x,\\
\vartheta(x) &:= 2x,\\
K(x) &:= \tfrac{1}{\pi},\\
\xi(x) &:= \tfrac{1}{\pi},\\
\rho(x) &:= \tfrac{1}{2\pi}.
\end{aligned}$$

B Analysis of continuum Bethe equation for $|\Delta| < 1$

In this appendix we gather some information on $\rho(x|q)$ when $|\Delta| < 1$. The first proposition justifies the existence of this quantity.

Proposition 23 (Existence of solutions to (cont.BE))**.** *For every $|\Delta| < 1$ and $q \geq 0$, there exists a unique solution $x \mapsto \rho(x|q)$ to* (cont.BE)*. Furthermore, for every $m \in [0, 1/2]$, there exists $Q(m)$ satisfying* (15)*.*

Proof. The operator $\mathcal{K}$ on $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ defined by

$$\mathcal{K}[f](x) := \int_{-q}^{q} K(x-y)f(y)dy$$

satisfies that

$$\|\mathcal{K}\|_{L^\infty \to L^\infty} = \|\mathcal{K}\|_{L^1 \to L^1} \leq \|K\|_{L^1} = |\vartheta(+\infty) - \vartheta(-\infty)| = \frac{|4\zeta - 2\pi|}{2\pi} < 1. \tag{62}$$

Hence, $\mathrm{Id} + \mathcal{K}$ is invertible and the solution $\rho(\lambda|q)$ is unique and lies in $L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ with a uniform bound on the norm. Since $K(x)$ is smooth in x and q, the Fredholm series representation for the resolvent of $\mathrm{Id} + \mathcal{K}$ [25] allows one to infer that $(x, q, \Delta) \mapsto \rho(x|q)$ is smooth.

The existence of $Q(m)$ follows readily from the continuity of the map $(x, q, \Delta) \mapsto \rho(x|q)$, the mean-value theorem, and the fact that $\rho(x|0)$ integrates to 0 while $\rho(\cdot) = \rho(\cdot|Q(\frac{1}{2}))$ to $\frac{1}{2}$. □

The following proposition gives the necessary properties for the proof of Theorem 4 when $-1 < \Delta < 0$ (see Section 3.2).

Proposition 24 (Properties necessary for Theorem 4)**.** *Fix $\varepsilon > 0$. Then, there exist $c, C > 0$ such that for every $-1 < \Delta \leq 0$,*

- (i) *For $q \in \mathbb{R}$ and $x \in \mathbb{R}$, $0 < \rho(x) \leq \rho(x|q) \leq \rho(x) + \rho(q)$.*
- (ii) *For every $x \in [-q - \varepsilon, q + \varepsilon]$, $\rho'(x|q) \leq \frac{C}{\zeta} \mathrm{e}^{\frac{\pi}{\zeta}\epsilon} \rho(x|q)$.*
- (iii) *For every m, $\frac{c}{\zeta}(\frac{1}{2} - m) \leq \rho(Q(m)) \leq \frac{C}{\zeta}(\frac{1}{2} - m)$.*

The lower bound of (i) was first established in [12].

Proof. Recall that $\widehat{R} = \widehat{K}/(1 + \widehat{K})$. Following [44], one gets that $R \geq 0$ since $\widehat{R} = \widehat{\rho}\widehat{K}/\widehat{\xi}$ and therefore

$$R(\lambda) = \int_{\mathbb{R}} \rho(\lambda - y)F(y)dy \tag{63}$$

in which ρ is obviously positive while

$$F(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\widehat{K}(\omega)}{\widehat{\xi}(\omega)} \mathrm{e}^{ix\omega} d\omega = \frac{1}{2(\pi - \zeta)} \frac{\sin\left(\frac{\pi - 2\zeta}{\pi - \zeta}\pi\right)}{\cos\left(\frac{\pi - 2\zeta}{\pi - \zeta}\pi\right) + \cosh\left(\frac{\pi x}{\pi - \zeta}\right)} > 0\,,$$

where the second equality follows from a straightforward residue computation.

The operator $\mathcal{U}$ on $L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ defined by

$$\mathcal{U}[f](x) := \int_{\mathbb{R}\setminus[-q,q]} R(x-y) f(y) dy$$

satisfies

$$\|\mathcal{U}\|_{L^\infty(\mathbb{R})\to L^\infty(\mathbb{R})} = \|\mathcal{U}\|_{L^1(\mathbb{R})\to L^1(\mathbb{R})} \leq \|R\|_{L^1(\mathbb{R})} = \widehat{R}(0) = \frac{\widehat{K}(0)}{1+\widehat{K}(0)} = \frac{\pi-2\zeta}{2\pi-2\zeta} < \frac{1}{2} \quad (64)$$

so that the version (20) of (cont.BE) immediately gives that

$$\rho(x|q) = \sum_{k=0}^{\infty} \mathcal{U}^k[\rho](x). \quad (65)$$

This expression and the fact that $R \geq 0$ gives the lower bound of (i). For the upper bound, we isolate the first term in the sum and then use operator bounds to get that

$$\rho(x|q) \leq \rho(x) + \sum_{k=1}^{\infty} \|\mathcal{U}^k\|_{L^\infty(\mathbb{R})\to L^\infty(\mathbb{R})} \cdot \|\rho \mathbb{1}_{|x|>q}\|_\infty \leq \rho(x) + \rho(q).$$

To get (ii), differentiate (20) with respect to x and then integrate by part to obtain that $\rho'(\cdot|q)$ satisfies the functional equation:

$$\rho'(x|q) - \int_{\mathbb{R}\setminus[-q,q]} R(x-y)\rho'(y|q)\, dy = \rho'(x) + [R(x-q) - R(x+q)]\rho(q|q).$$

The claim follows by using the same expression as (65), as well as the bounds

$$|\rho'(x)| \leq \tfrac{\pi}{\zeta}\rho(x) \qquad \text{and} \qquad \rho(q) \leq C' \mathrm{e}^{\frac{\pi}{\zeta}\epsilon} \rho(x)$$

for $|x| \leq q + \varepsilon$, and $\|R\|_\infty < \widetilde{C}/\zeta$.

We now focus on (iii). The definition of m, (20), and

$$\widehat{R}(0) = \frac{\widehat{K}(0)}{1+\widehat{K}(0)} = \frac{\pi-2\zeta}{2\pi-2\zeta} < \frac{1}{2}$$

give that

$$\begin{aligned}
\tfrac{1}{2} - m &= \int_{\mathbb{R}} \rho(x)dx - \int_{-Q(m)}^{Q(m)} \rho(x|Q(m))dx \\
&= \int_{[-Q(m),Q(m)]^{\mathrm{c}}} \rho(x|Q(m))dx - \int_{\mathbb{R}\times[-Q(m),Q(m)]^{\mathrm{c}}} R(x-y)\rho(y|Q(m))dydx \\
&= \tfrac{\pi}{2\pi-2\zeta} \int_{[-Q(m),Q(m)]^{\mathrm{c}}} \rho(x|Q(m))dx > \tfrac{\pi}{2\pi-2\zeta} \int_{[-Q(m),Q(m)]^{\mathrm{c}}} \rho(x)dx \sim \tfrac{2\zeta}{2\pi-2\zeta}\rho(Q(m)).
\end{aligned} \quad (66)$$

as $m \to 1/2$. Above, we set $[-Q(m), Q(m)]^{\mathrm{c}} := \mathbb{R} \setminus [-Q(m), Q(m)]$. Plugging again the expression (65) in this estimate to replace $\rho(\cdot|Q(m))$ by ρ in the integral, and then using the explicit formula for ρ gives the result easily. □

We finish with the properties necessary to obtain Theorem 2 for $|\Delta| < 1$.

Proposition 25 (Properties necessary for Theorem 2)**.** *There exists $C > 0$ such that for every $|\Delta| < 1$:*

(i) *There exists a unique solution $T \in \big(L^\infty \cap L^1\big)(\mathbb{R})$ of the functional equation*

$$T(x) - \int_0^\infty R(x-y)T(y)dy = \mathfrak{e}(x) \quad \textit{with} \quad \mathfrak{e}(x) \;=\; \tfrac{1}{\zeta}\mathrm{e}^{-x\frac{\pi}{\zeta}}\mathbb{1}_{\mathbb{R}_+}(x). \tag{67}$$

(ii) *For every $q \geq 0$ and $x \in \mathbb{R}$,*

$$\rho(x|q) = \rho(x) + \mathrm{e}^{-q\frac{\pi}{\zeta}}\big[(T-\mathfrak{e})(q-x) + (T-\mathfrak{e})(-q-x) + \delta T(x)\big]\,, \tag{68}$$

where $\|\delta T\|_\infty + \|\delta T\|_1 \leq Ce^{-2q}$.

(iii) *It holds*

$$\lim_{m\to 1/2}(\tfrac{1}{2} - m)\mathrm{e}^{Q(m)\frac{\pi}{\zeta}} \;=\; \frac{\pi}{\pi - \zeta}\int_0^{+\infty} T(\lambda)d\lambda.$$

Note that, in fact, one may solve (67) in terms of a scalar Riemann–Hilbert problem by implementing the Wiener-Hopf method. However, we will not need such a precise information on T and will thus establish Proposition 25 by more elementary means.

Proof. We stress that, below, all domination relations $O(f)$ will be uniform in Δ, *viz.* bounded by Cf with C being Δ independent. We start with Item (i). Introduce the operator $\mathcal{V}$ on $L^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ defined by

$$\mathcal{V}[f](x) := \int_{\mathbb{R}_+} R(x-y)f(y)dy.$$

which has $\|\cdot\|_{L^1(\mathbb{R})\to L^1(\mathbb{R})}$-operator norm smaller than $1/2$, owing to the chain of bounds $\|\mathcal{V}[f]\|_1 \;\leq\; \|R\|_\infty \|f\|_1 \;\leq\; \frac{1}{2}\|f\|_1$. This justifies the existence and uniqueness of $T \in L^1(\mathbb{R})$ and gives the formula

$$T = \sum_{k=0}^{\infty} \mathcal{V}^k[\mathfrak{e}]. \tag{69}$$

For item (ii), introduce $\mathcal{U}_\pm[f](x) := \mathcal{V}[f(\pm q \pm \cdot)](-q \pm x)$ and observe that $\mathcal{U}[f] = \mathcal{U}_+[f] + \mathcal{U}_-[f]$. Further, by introducing the operators $\tau_\pm, \check{\tau}_\pm$ such that $\tau_\pm[f](x) = f(\pm q \pm x)$ and $\check{\tau}_\pm[f](x) = f(-q \pm x)$ one finds that $\mathcal{U}_\pm = \check{\tau}_\pm \mathcal{V} \tau_\pm$. Therefore, one finds that

$$\rho(\lambda|q) = \rho(\lambda) \;+\; \sum_{k\geq 1}(\mathcal{U}_+ + \mathcal{U}_-)^k[\rho](\lambda) \;=\; \sum_{k\geq 1}(\mathcal{U}_\pm^k + \mathcal{U}_\pm^k)[\rho] \;+\; \delta T_{\mathrm{pert}}(\lambda)\,,$$

in which

$$\delta T_{\mathrm{pert}}(\lambda) \;=\; \sum_{k\geq 1}\sum_{p=1}^{k-1}\sum_{\substack{\varepsilon_i\in\{\pm\}\\ \#\{i:\varepsilon_i=+\}=p}} \big(\mathcal{U}_{\varepsilon_1}\cdots\mathcal{U}_{\varepsilon_k}\big)[\rho](\lambda). \tag{70}$$

A direct calculation yields

$$\mathcal{U}_{+}\mathcal{U}_{-}[f](\lambda) = \int_{0}^{+\infty} R(-q+\lambda-\mu)d\mu \int_{-\infty}^{-q} R(q+\mu-\nu)f(\nu)d\nu. \tag{71}$$

In order to estimate the norm of $\mathcal{U}^{+}\mathcal{U}^{-}$ one recalls the convolution representation for R (63) which shows that R is monotonously decreasing on $\mathbb{R}_+$ and enjoys the bound

$$R(\lambda) \;=\; O(\mathfrak{b}(\lambda)) \quad \text{as} \quad \lambda\to+\infty \quad \text{and where} \quad \mathfrak{b}(\lambda):=\max\Big\{\mathrm{e}^{-\frac{2\pi}{\pi-\zeta}|\lambda|},\mathrm{e}^{-\frac{\pi}{\zeta}|\lambda|}\Big\}\;. \tag{72}$$

This immediately yields

$$\|\mathcal{U}_{+}\mathcal{U}_{-}[f]\|_{L^1(\mathbb{R})} \;\leq\; \mathfrak{b}(2q)\|R\|_{L^1(\mathbb{R})}\|f\|_{L^1(\mathbb{R})}\;.$$

Clearly, similar bounds do hold for $\|\mathcal{U}_{-}\mathcal{U}_{+}[f]\|_{L^1(\mathbb{R})}$.

Now, observe that $\check{\tau}_{\pm}\tau_{\pm}=\mathrm{Id}$, so that $\big(\mathcal{U}_{\pm}\big)^k[\rho]=\check{\tau}_{\pm}\mathcal{V}^k[\rho(\pm q\pm\cdot)]$. The uniform in $x\geq 0$ expansion

$$\rho(\pm q\pm x) \;=\; \mathrm{e}^{-\frac{\pi}{\zeta}q}\mathfrak{e}(x)\Big(1+\delta\mathfrak{e}(x)\Big) \qquad \text{with} \qquad \delta\mathfrak{e}(x):=\frac{-\mathrm{e}^{-2\frac{\pi}{\zeta}(q+x)}}{1+\mathrm{e}^{-2\frac{\pi}{\zeta}(q+x)}}\;, \tag{73}$$

yields

$$\sum_{k\geq 1}\big(\mathcal{U}_{\pm}\big)^k[\rho] \;=\; \mathrm{e}^{-\frac{\pi}{\zeta}q}\check{\tau}_{\pm}\big[T-\mathfrak{e}\big] \;+\; \mathrm{e}^{-\frac{\pi}{\zeta}q}\delta T_{\pm} \quad \text{with} \quad \delta T_{\pm} \;=\; \sum_{k\geq 1}\check{\tau}_{\pm}\cdot\mathcal{V}^k[\mathfrak{e}\delta\mathfrak{e}] \tag{74}$$

By using that

$$\|\mathcal{U}_{\pm}\|_{L^{1/\infty}(\mathbb{R})\to L^{1/\infty}(\mathbb{R})} \;=\; \|\mathcal{V}\|_{L^{1/\infty}(\mathbb{R})} < \varkappa_{\zeta} = \frac{\pi-2\zeta}{2(\pi-\zeta)} < \tfrac{1}{2}, \tag{75}$$

one infers that

$$\|\delta T_{\pm}\|_{L^{1/\infty}(\mathbb{R})} \;\leq\; \sum_{k\geq 1}\frac{1}{2^k}\|\mathfrak{e}\delta\mathfrak{e}\|_{L^{1/\infty}(\mathbb{R}_+)} \;\leq\; C\mathrm{e}^{-\frac{2\pi}{\zeta}q}\;.$$

Further, given $\varepsilon_i\in\{\pm\}$ such that $\#\{i:\varepsilon_i=+\}\in\{1,\dots,k-1\}$, there is necessarily at least one change of sign in the string $\varepsilon_1,\dots,\varepsilon_k$ so that one gets

$$\|\mathcal{U}_{\varepsilon_1}\cdots\mathcal{U}_{\varepsilon_k}[\rho]\|_{L^{1/\infty}(\mathbb{R})} \;=\; \varkappa_{\zeta}^{k-2}\cdot\max\Big\{\|\mathcal{U}_{+}\mathcal{U}_{-}\|_{L^{1/\infty}(\mathbb{R})},\|\mathcal{U}_{-}\mathcal{U}_{+}\|_{L^{1/\infty}(\mathbb{R})}\Big\} \;\leq\; C\cdot\varkappa_{\zeta}^{k-2}\cdot\mathfrak{b}(2q).$$

This leads to the estimate on δT_{pert} introduced in (70):

$$\|\delta T_{\mathrm{pert}}\|_{L^{1/\infty}(\mathbb{R})} \;\leq\; C\cdot\sum_{k\geq 1}\sum_{p=1}^{k-1} C_k^p \varkappa_\zeta^{k-2}\mathfrak{b}(2q)\cdot \mathrm{e}^{-\frac{\pi}{\zeta}q} \;\leq\; C\mathrm{e}^{-\frac{\pi}{\zeta}q}\mathfrak{b}(2q)\cdot\sum_{k\geq 1}(2\varkappa_\zeta)^k = C'(\zeta)\mathrm{e}^{-\frac{\pi}{\zeta}q}\mathfrak{b}(2q)\;.$$

There, $\mathrm{e}^{-\frac{\pi}{\zeta}q}$ issues from the estimates of the action of $\mathcal{U}^{\pm}$ on ρ. Thus, one obtains the representation (68) with δT given by

$$\delta T \;=\; \delta T_{+} \,+\, \delta T_{-} \,+\, \delta T_{\mathrm{pert}} \quad \text{and} \quad \|\delta T\|_{L^{1/\infty}(\mathbb{R})} \;=\; O\big(\mathfrak{b}(2q) \,+\, \mathrm{e}^{-2\frac{\pi}{\zeta}q}\big). \tag{76}$$

For Item (iii), recall the exact representation for $\frac{1}{2}-m$ given in (66). Then, by virtue of (68) one gets

$$\int\limits_{[-q,q]^{\mathrm{c}}} \rho(\lambda|q)d\lambda \;=\; 2\int\limits_{q}^{\infty}\rho(\lambda)d\lambda + 2\mathrm{e}^{-\frac{\pi}{\zeta}q}\int\limits_{0}^{\infty}\big(T-\mathfrak{e}\big)(\lambda)d\lambda + 2\int\limits_{q}^{\infty}\big(T(-q-\lambda)+\delta T(\lambda)\big)d\lambda\;. \tag{77}$$

There, we used that $\mathfrak{e}$ has support on $\mathbb{R}_{+}$. The part involving δT can be directly estimated when $q\to+\infty$ to give $O(\mathrm{e}^{-2q})$. By repeating the previous reasoning, it is easy to see that

$$\int\limits_{q}^{\infty}\rho(\lambda)d\lambda \sim \mathrm{e}^{-\frac{\pi}{\zeta}q}\int\limits_{0}^{\infty}\mathfrak{e}(\lambda)d\lambda$$

Further, by using the integral equation satisfied by T, one gets that

$$\int\limits_{q}^{\infty} T(-q-\lambda)d\lambda \;=\; \int\limits_{-\infty}^{-2q} T(\lambda)d\lambda \;=\; \int\limits_{-\infty}^{0}\int\limits_{0}^{\infty} R(\lambda-2q-\mu)T(\mu)d\mu d\lambda$$

$$\leq\; \|T\|_{L^1(\mathbb{R})}\int\limits_{0}^{\infty} R(2q+\lambda)d\lambda \;\leq\; C\mathfrak{b}(2q)\;.$$

Thus, by substituing $q\hookrightarrow Q(m)$ in the above estimates and using that $Q(m)\to+\infty$ as $m\to\frac{1}{2}$ by virtue of Item (iii) of Proposition 24, one obtains that

$$\tfrac{1}{2}-m \;=\; \mathrm{e}^{-\frac{\pi}{\zeta}Q(m)}\frac{\pi}{\pi-\zeta}\Bigg\{\int\limits_{0}^{\infty} T(\lambda)d\lambda \,+\, O\Big(\mathfrak{b}(2Q(m)) \,+\, \mathrm{e}^{-2\frac{\pi}{\zeta}Q(m)}\Big)\Bigg\}\;.$$

Therefore

$$\lim_{m\to\frac{1}{2}}\big\{\tfrac{1}{2}-m\big\}\mathrm{e}^{\frac{\pi}{\zeta}Q(m)} \;=\; \frac{\pi}{\pi-\zeta}\int\limits_{0}^{\infty}T(\lambda)d\lambda\;.$$

The constant on the *rhs* is strictly positive since T is given by a sum of strictly positive terms.

□

C Analysis of continuum Bethe equation when $\Delta = -1$

Proposition 26 (Existence of solutions to (cont.BE))**.** *Fix $\Delta = -1$ and $q \geq 0$, there exists a unique solution $x \mapsto \rho(x|q)$ to* (cont.BE)*. Furthermore, for every $m \in [0, 1/2]$, there exists $Q(m)$ satisfying* (15)*.*

Proof. The proof valid for $|\Delta| < 1$ does not generalise directly since estimating that the operator $\mathcal{K}$ has $\|\cdot\|_{L^1 \to L^1}$-norm strictly less than 1 demands more effort, see [30]. However, by working with (20) instead of (cont.BE), one readily checks that the fact that the operator $\mathcal{U}$ has norm smaller than $\frac{1}{2}$, which gives the existence of solutions. The rest of the proof is the same. □

The following proposition gives the necessary properties for the proof of Theorem 2 when $\Delta = -1$. The proof is the same as for $|\Delta| < 1$.

Proposition 27. *Fix $\Delta = -1$. There exists $C > 0$ such that*

(i) *There exists a unique solution $T \in L^1(\mathbb{R})$ of the functional equation*

$$T(x) - \int_0^\infty R_{-1}(x-y)T(y)dy = \mathfrak{e}(x) \equiv \mathrm{e}^{-x\pi}\mathbb{1}_{\mathbb{R}_+}(x). \qquad (78)$$

(ii) *We have that*

$$\rho(x|q) \ = \ \rho(x) + \mathrm{e}^{-q\pi}\left[(T-\mathfrak{e})(q-x) + (T-\mathfrak{e})(-q-x) + \delta T(x)\right], \qquad (79)$$

where $\|\delta T\|_{L^\infty(\mathbb{R})} + \|\delta T\|_{L^1(\mathbb{R})} \leq Ce^{-2q}$.

(iii) $\lim_{m \to 1/2} (\frac{1}{2} - m)\mathrm{e}^{Q(m)\pi}$ *exists and belongs to $(0, \infty)$.*

Note that, above, R_{-1} stands for the resolvent kernel to the operator $\mathrm{Id} + \mathcal{K}$ at $\Delta = -1$ and acting on $L^2(\mathbb{R})$.

D Computations for $\Delta < -1$

Proposition 28. *For every $\Delta < -1$ and $q \in [0, \pi/2]$, there exists a unique solution $x \mapsto \rho(x|q)$ to* (cont.BE)*. Furthermore, there exists $Q(m)$ satisfying* (15)*.*

Proof. In order to circumvent the problems with estimating $\|\mathcal{K}\|_{L^1([-q,q]) \to L^1([-q,q])}$, we again work with the alternative representation of the integral equation (20) in which

$$R(\lambda) \ = \ \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \frac{\mathrm{e}^{2i\lambda - |n|\zeta}}{\cosh(n\zeta)} \ .$$

Then, it remains to use that $\|R\|_{L^1([-\pi/2, \pi/2]} = 1/2$ so as to conclude as in the $\Delta = -1$. Finally, the existence of $Q(m)$ follows from the same argument as in the $|\Delta| < 1$ case. □

Lemma 29. *It holds* $\rho(x|q) \geq \rho(x) \geq \frac{1}{2\zeta}$

Proof. The explicit expression for ρ implies that $\rho(x) \geq \frac{1}{2\zeta}$. It is thus enough to establish the upper bound. This follows from an expression similar to (65) for $\Delta < -1$ and the fact that $R \geq 0$. To check the latter, note that

$$\widehat{R}(n) = \widehat{\rho}(n)\frac{\widehat{K}(n)}{\widehat{\xi}(n)} = \widehat{\rho}(n)\widehat{\xi}(n) = \frac{e^{-2|n|\zeta}}{1+e^{-2|n|\zeta}}$$

so that R is the convolution of ρ and ξ which are both positive. □

Lemma 30. *For* $\Delta < -1$*, we get*

$$\lim_{m\to 1/2} \frac{1-2m}{\pi - 2Q(m)} = \rho(\tfrac{\pi}{2}).$$

Proof. Equation (20) and the observation that $\|R\|_{L^1([-\frac{\pi}{2},\frac{\pi}{2}])} = \frac{1}{2}$ along with the continuity of ρ give

$$\begin{aligned}
\tfrac{1}{2} - m &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \rho(x)dx - \int_{-Q(m)}^{Q(m)} \rho(x|Q(m))dx \\
&= \big(\pi - 2Q(m)\big)\rho(\tfrac{\pi}{2})[1+o(1)] - \int_{-Q(m)}^{Q(m)} \int_{[-Q(m),Q(m)]^{\mathrm{c}}} \rho(y|Q(m))R(x-y)\,dy\,dx. \\
&= \big(\tfrac{\pi}{2} - Q(m)\big)\rho(\tfrac{\pi}{2})[1+o(1)].
\end{aligned}$$

where we set $[-Q(m),Q(m)]^{\mathrm{c}} := [-\pi/2,\pi/2] \setminus [-Q(m),Q(m)]$.

□

References

[1] M. Aizenman and H. Duminil-Copin. Dimerization and Néel order in different quantum spin chains through a shared loop representation. arXiv:2002.02543.

[2] R.J. Baxter. Generalized Ferroelectric Model on a Square Lattice. *S. Appli. Math. L* (1):51–69, 1971.

[3] R.J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1989. Reprint of the 1982 original.

[4] R. J. Baxter, S. B. Kelland, and F. Y. Wu. Equivalence of the Potts model or Whitney polynomial with an ice-type model. *Journal of Physics A: Mathematical and General*, 9(3) :397, 1976.

[5] M.T. Batchelor and A. Klümper, An analytic treatment of finite-size corrections in the spin-1 antiferromagnetic XXZ chain, *Journal of Physics A: Mathematical and General* 23:L189–195, 1990.

[6] H. Bethe, On the theory of metals. 1. eigenvalues and eigenfunctions for the linear atomic chain, *Zeischrift für Physik* 71:205–226, 1931.

[7] T. Deguchi, Introduction to solvable lattice models in statistical and mathematical physics. *Classical and Quantum Integrable Systems: Theory and Applications*, edited by A. Kundu, IOP Publishing, Chapter 5,113–151, 2003.

[8] C. Destri and H.J. de Vega, New thermodynamic Bethe Ansatz equations without strings *Physical Review Letters*, 69:2313–2317, 1992.

[9] C. Destri and H.J. de Vega, Unified approach to Thermodynamic Bethe Ansatz and finite size corrections for lattice models and field theories *Nuclear Physics B*, 438:413–454, 1995.

[10] T.C. Dorlas and M. Samsonov, On the thermodynamic limit of the 6-vertex model, ArXiv:0903.2657, 2009.

[11] R.-M. Dudley and J.M Gonzalez-Barrios, Metric entropy conditions for an operator to be of trace class *Proceedings of the American Mathematical Society*, 118:175–180, 1993.

[12] M. Dugave, F. Gohmann, and K.K. Kozlowski, Functions Characterizing the Ground State of the XXZ Spin-1/2 Chain in the Thermodynamic Limit. *Symmetry, Integrability and Geometry: Methods and Applications, SIGMA* 10: 043, 18 pages, 2014.

[13] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the phase transition for planar random-cluster and Potts models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, 2017.

[14] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *arXiv:1611.09877*, 2016.

[15] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu and V. Tassion. The Bethe Ansatz for the six-vertex and XXZ models: an exposition *Prob. Surveys*, 15:102–130, 2018.

[16] H. Duminil-Copin, M. Harel, B. Laslier, A. Raoufi, and G. Ray. Logarithmic fluctuations of the height function in square-ice. arXiv:1911.00092, 2019.

[17] H. Duminil-Copin, D. Krachun, K.K. Kozlowski, I. Manolescu and M. Oulamara. Rotational invariance in critical planar lattice models. *preprint*, 2020.

[18] H. Duminil-Copin, A. Karrila, I. Manolescu and M. Oulamara. Delocalization of the height function of the six-vertex model. *preprint*, 2020.

[19] L.D. Faddeev and E.K. Sklyanin and L.A. Takhtadzhan. Quantum inverse problem method I. *Theoretical and Mathematical Physics*, **40**:2, 688-706, 1979.

[20] P.L. Ferrari and H. Spohn. Domino tilings and the six-vertex model at its free-fermion point. *J. Phys. A* 39(33):10297–10306, 2006.

[21] M. Gaudin, B. McCoy, T. Wu. Normalization sum for the Bethe's hypothesis wave functions of the Heisenberg-Ising chain. *Phys. Rev. D*, 23(2), 1981.

[22] A. Glazman and R. Peled. On the transition between the disordered and antiferroelectric phases of the 6-vertex model. *arXiv:1909.03436v1*, 2019.

[23] P.S. Goldbaum. Existence of solutions to the Bethe Ansatz equations for the 1d Hubbard model: finite lattice and thermodynamic limit. *Communications in mathematical physics*, 258(2):317–337, 2005.

[24] R.B. Griffiths. Magnetization curve at zero temperature for the antiferromagnetic Heisenberg linear chain *Physics Reveiw*, 133:A768–A775, 1964.

[25] I. Gohberg and S. Goldberg and N. Krupnik, Traces and determinants of linear operators *Operator theory advances and applications*, 116, Birkhäuser, 258pp, 2000.

[26] E. Gusev, Weak convergence of wave values in the Heisenberg model *Reports on Mathematical Physics*, 18(3):399–410, 1980.

[27] R.B. Griffiths, Über das Austauschproblem eines Kristalles *Arkiv för matematik, astronomi och fysik*, 26A:1–106, 1938.

[28] N. Kitanine, J.-M. Maillet and V. Terras, Form factors of the XXZ Heisenberg spin-1/2 finite chain *Nuclear Physics B* , 554: 647–678, 1999.

[29] V.E. Korepin, Calculation of norms of Bethe wave-functions *Communications in Mathematical Physics*, 86(3): 391–418, 1982.

[30] K.K. Kozlowski, On condensation properties of Bethe roots associated with the XXZ chain *Communications in Mathematical Physics*, 357(3): 1009–1069, 2018.

[31] E.H. Lieb. Exact Solution of the problem of the entropy of the two-dimensional ice. *Physical Review*, 18: 692–694, 1967.

[32] E.H. Lieb. Exact Solution of the Two-Dimensional Slater KDP Model of a Ferroelectric. *Physical Review*, 19(3),108–110, 1967.

[33] E.H. Lieb. Residual entropy of square ice. *Physical Review*, 162(1):162, 1967.

[34] E.H. Lieb. Exact solution of the F model of an antiferroelectric. *Condensed Matter Physics and Exactly Soluble Models*, 453-455, 1967.

[35] E.H. Lieb and F.Y. Wu. Two-dimension ferroelectric models in "*Phase transitions and critical phenomena I*, C. Domb and M.S. Green , Academic Press, pp 331-490, 1972.

[36] B.M. McCoy and T.T. Wu Hydrogen-Bonded Crystals and the Anisotropic Heisenberg Chain *Il Nuovo Cimento*, 56: 311-315, 1968.

[37] I.M. Nolden, The asymmetric six-vertex model. *Journal of Statistical Physics* , 67(1–2):144–201, 1992.

[38] R. Orbach, Linear antiferromagnetic chain with anisotropic coupling. *Physics Review Letter*, 112:309–316, 1958.

[39] N.A. Slavnov, On an identity for dual fields *Zapiski Nauchnykh Seminarov POMI*, 245:270–281, 1997.

[40] B. Sutherland. Exact solution of a two-dimensional model for hydrogen-bonded crystals. *Physical Review Letters*, 19(3): 103, 1967.

[41] N. Reshetikhin. Lectures on the integrability of the 6-vertex model. *arXiv1010.5031*, October 2010.

[42] H.J. de Vega and F. Woynarowich, Method for calculating finite size corrections in Bethe Ansatz systems- Heisenberg chains and 6-vertex model, *Nuclear Physics B*, 251: 439–456, 1985.

[43] C. N. Yang and C. P. Yang, One-dimensional chain of anisotropic spin-spin interactions. I. Proof of Bethe's hypothesis for ground state in a finite system, *Physical Review*, 150: 321–327, 1966.

[44] C. N. Yang and C. P. Yang, One dimensional chain of anisotropic spin-spin interactions: II properties of the ground state energy per lattice site for an infinite system, *Physical Review*, 150: 327–339, 1966.