Anchoring and Binning the Coordinate Bethe Ansatz

Jarah Evslin\textsuperscript{1,2} *

1) Institute of Modern Physics, NanChangLu 509, Lanzhou 730000, China
2) University of the Chinese Academy of Sciences, YuQuanLu 19A, Beijing 100049, China

Abstract

The Coordinate Bethe Ansatz (CBA) expresses, as a sum over permutations, the matrix element of an XXX Heisenberg spin chain Hamiltonian eigenstate with a state with fixed spins. These matrix elements comprise the wave functions of the Hamiltonian eigenstates. However, as the complexity of the sum grows rapidly with the length \(N\) of the spin chain, the exact wave function in the continuum limit is too cumbersome to be exploited. In this note we provide an approximation to the CBA whose complexity does not directly depend upon \(N\). This consists of two steps. First, we add an anchor to the argument of the exponential in the CBA. The anchor is a permutation-dependent integral multiple of \(2\pi i\). Once anchored, the distribution of these arguments simplifies, becoming approximately Gaussian. The wave function is given by the Fourier transform of this distribution and so the calculation of the wave function reduces to the calculation of the moments of the distribution. Second, we parametrize the permutation group as a map between integers and we bin these maps. The calculation of the moments then reduces to a combinatorial exercise on the partitioning into bins.

1 Introduction

1.1 Motivation

Man has always sought to understand the origin of the Yang-Mills mass gap. In the instantaneous frame, it is a consequence of the vacuum state. This vacuum state may be realized, in the Schrodinger picture, as a wave functional which satisfies the Schrodinger equation. Despite decades of efforts, no such solution appears to be forthcoming.

On the other hand, Yang-Mills theory in 3+1 dimensions is quite similar to the \(\mathbb{C}P^1\) nonlinear sigma model in 1+1 dimensions. Here also fractional instantons are somehow...

\*jarah@impcas.ac.cn
responsible for the generation of a mass gap \([1]\). Knowledge of the ground state and first excited state wave functionals of this model would unlock exciting doors, allowing a concrete understating of how the instantons generate the mass gap in the Minkowski theory, presumably as a kind of infinite-dimensional generalization of the familiar story in quantum mechanics with a double well potential.

The barrier between the wells in the case of quantum mechanics is just the potential barrier to a closed string, corresponding to the image of the sigma model field with periodic boundary conditions on the target \(\mathbb{CP}^1\), growing from a point to wrap around the equator of the target and then to shrink on the other side. In the case of Yang-Mills, it is similarly the jump from one integral Chern-Simons number to the next. The instanton effects result from the fact that the wave functional does not vanish inside of this barrier, and so enforces communication between the wave functionals on both sides, albeit suppressed by the exponentiated instanton action. In quantum mechanics, the mass gap may be seen as a consequence of a discrete choice in how the wave functions are connected across the barriers. Is there a similar story in quantum field theory? Does this cross-barrier bridge also render a monopole-operator tachyonic in Yang-Mills? To answer these questions, we need at least to understand the basic features of the ground state and first excited wave functionals. For example, does the first excited state wave functional have a node at the maximum of this potential, corresponding to a sigma model field encircling the equator of \(\mathbb{CP}^1\) or a half-integral Chern-Simons number in Yang-Mills?

This sigma model is not only solvable but has already been solved \([2]\). So, what are the wave functionals? A map between the \(\mathbb{CP}^1\) sigma model and the XXX Heisenberg spin chain was shown in \([3,4]\) at the level of low energy fluctuations and in \([5]\) in the full quantum theory. The former applies to a spin chain of any spin \(s\), with strong coupling at small spin while the latter strictly speaking yields an equivalence only at infinite \(s\), although finite \(s\) can be used as a definition for an \(\mathbb{CP}^1\) sigma model whose target is a quantum deformed \(\mathbb{CP}^1\). The spectra of these spin chains are also well-known. There are many formalisms for writing the wave functions corresponding to these states and so these states are also known. With the XXX states known, and the map to the sigma model known, also the sigma model wave functionals are by definition known.

So what are the sigma model wave functionals? To actually take these spin chain solutions and map them to something intelligible on the sigma model side was Faddeev’s challenge to his students in \([6]\). The map is known in the coordinate basis of spins in the spin chain, and so to meet the challenge one needs the matrix elements of the spin chain Hamiltonian eigenvectors with the coordinate states, which have definite spins at each lattice site. The challenge is indeed a challenge because, while many forms are by now known for the spin
chain Hamiltonian eigenstates in the coordinate basis $[7, 8, 9, 10, 11, 12]$, each grows in complexity either with the length $N$ of the chain or else with the distance of a coordinate state from a preferred spin state, such as the classical vacuum.

Our goal is to present a method of approximating the matrix elements which depends on the complexity of the state, but not directly on $N$. The individual lattice sites are replaced by bins. The intuition is that the states which survive to the continuum limit are those which are essentially homogeneous inside of each bin. Homogeneous means that using the mapping to the CP$^1$ model, each pair of adjacent lattice sites in the same bin corresponds to the same point in CP$^1$. Therefore the points in the sigma model correspond not to the original lattice sites, but rather to the bins. To describe the sigma model vacuum, one then needs to calculate the spin chain matrix element for each such configuration of bins. This calculation is very different from the Coordinate Bethe Ansatz (CBA) because the complicated symmetric sum has been smoothed away. The goal of the present note is to present a formalism which allows these bin states to be derived from the CBA.

1.2 Outline

We begin in Sec. 3 with the first key ingredient in our construction, the anchor. The CBA is a sum of phases $e^{i\alpha}$, one for each element of the permutation group $S_n$, where $n = N/2$ for the antiferromagnetic ground state, in which we will be interested from now on. At large $N$, for a given spin basis element, and a given state, in the continuum limit the phases have a density $\rho(\alpha)$ in $S_n$ and so the matrix element is

$$\int_{-\infty}^{\infty} e^{i\alpha} \rho(\alpha) d\alpha. \quad (1.1)$$

In other words it is given by the Fourier transform of the density function $\rho(\alpha)$.

The anchor is a permutation-dependent integral multiple of $2\pi$ which we will subtract from the arguments $\alpha$ of the phases. Clearly subtracting the anchor does not affect the wave function. However, in the CBA the arguments are written as real numbers and their scatter is far larger than $2\pi$, despite the fact that the wave function depends on the argument modulo $2\pi$. In fact, the standard deviation of the arguments

$$\sqrt{\int_{S_n} \alpha^2 \rho(\alpha) - \left(\int_{S_n} \alpha\rho(\alpha)\right)^2} \quad (1.2)$$

is of order $N^{3/2}$ and the distribution is rich in fine structure which dominates the Fourier transform but is very difficult to extract from the moments of $\rho(\alpha)$. If one tried a Gaussian
approximation in Eq. (1.1) anyway, the integral would yield matrix elements of order $e^{-N^3}$ which is clearly inconsistent, as the $2^N$ matrix elements must sum in quadrature to the identity.

On the other hand, subtracting the anchor from $\alpha$ the fine structure disappears and to a good approximation the anchored $\rho(\alpha)$ is indeed a Gaussian. The anchor changes $\rho(\alpha)$ so that the standard deviation (1.2) is of order $O(N^{1/2})$ or $O(N)$. In the first case, the Gaussian approximation in (1.1) yields the correct $e^{-N}$ scaling for the states. The absence of substructure implies that once anchored, the calculation of the matrix elements reduces to the calculation of the moments of $\rho(\alpha)$.

The other key ingredient is introduced in Sec. 4. To calculate these moments, the elements of the group $S_n$ are realized as one to one maps from the integers $[1,n]$ to themselves. This interval is divided into $q$ bins. For each permutation $g$ one can determine how many elements of the $i$th bin map to the $j$th bin. We will call this number $f_{ij}(g)$. We rewrite the CBA in terms of the quantities $f_{ij}(g)$. The moments of $\rho(\alpha)$ are then determined from the correlation functions of $f_{ij}(g)$, which in turn are determined from the combinatorics of the number of $g \in S_n$ with each value of $f_{ij}$. We will show how these correlation functions are computed.

Finally in Sec. 5 we will, in the case of the antiferromagnetic ground state, use these correlation functions to calculate the $O(N^2)$ contribution to the second moment of $\rho(\alpha)$. We will see explicitly that the anchor makes its coefficient quite small, but it may not vanish.

2 The Antiferromagnetic XXX Heisenberg Spin Chain

The $\mathbb{C}P^1$ sigma model is the continuum limit of a spin chain with an infinite spin at each lattice site. Classically, the spin squared corresponds to the inverse coupling and so low spin corresponds to a high coupling. In particular, at low spin one describes the sigma model at strong coupling and one does not expect a sensible description of individual instantons. Therefore, it will be essential for us to eventually extend our analysis to higher spin. However, in the present note we will restrict our attention to spin $s = 1/2$.

2.1 Finite Chain

The spin $1/2$ Heisenberg spin chain consists of $N$ lattice sites. At each lattice site lies a Hilbert space $\mathbb{C}^2$ with basis $\{\uparrow, \downarrow\}$. The total Hilbert space is the $N$-fold tensor product $\mathbb{C}^{2^N}$.

\footnote{As was described in Ref. [13], when $N = \infty$ this space decomposes into superselection selectors. We will be interested in finite $N$ in the present note, however it is tempting to conjecture that the superselection}
of these \( C^2 \). Acting on the \( l \)th copy of \( C^2 \) is the \( \mathfrak{su}(2) \) Lie algebra with generators \( \sigma_i \) satisfying
\[
[\sigma_i, \sigma_j] = 2i \delta_{lm} \epsilon^{ijk} \sigma_k. \tag{2.1}
\]

The XXX spin chain corresponds to the Hamiltonian
\[
H = J \sum_{i=1}^{3} \sum_{l=1}^{N} (\sigma_i^l \sigma_{i+1}^l - 1) \tag{2.2}
\]
where 1 is the identity. We will let the constant \( J \) be positive, corresponding to the antiferromagnetic spin chain. Although the eigenvalues of \( H \) depend on \( J \), in this note we will only be interested in the eigenvectors, which are independent of \( J \). In particular, we will restrict our attention to the antiferromagnetic ground state \(|\Omega\rangle\), which is the eigenstate of \( H \) with minimal eigenvalue. This state is the same for any positive value of \( J \).

Any state can be decomposed into the basis consisting of the tensor product of the \(|\uparrow\rangle_l, |\downarrow\rangle_l\) bases at each lattice site. An element of the basis is a string of \( \uparrow \)'s and \( \downarrow \)'s. It is described by the set of positions \( m(i) \) of the \( i \)th \( \downarrow \) for all \( i \). Therefore an arbitrary state \(|\Psi\rangle\) is fully characterized by the matrix elements
\[
a(\{m(i)\}) = \langle \{m(i)\} |\Psi\rangle. \tag{2.3}\]

For all Hamiltonian eigenstates \( \Psi \), the elements \( a(\{m(i)\}) \) are given by the coordinate Bethe Ansatz [7]. For a state with \( n \) spin downs
\[
a(\{m(i)\}) = \sum_{g \in S_n} \exp \left( i \sum_{j=1}^{n} m(j) K(P(j)) + \frac{i}{2} \sum_{j<k} \Phi(P(j), P(k)) \right) \tag{2.4}\]
where \( P(j) : [1, n] \to [1, n] \) is the permutation corresponding to \( g \in S_n \). The information about the state is contained in the functions \( K \in [0, 2\pi] \) and \( \Phi \in [-\pi, \pi] \) which are related by the Bethe equation
\[
NK(i) = 2\pi Q(i) + \sum_{j \neq i}^{n} \Phi(i, j) \tag{2.5}\]
where \( Q(i) \) is an integer. In fact, a state is characterized by just the set of \( \{Q(i)\} \). The ground state for example corresponds to
\[
N = 2n, \quad Q(i) = 2n - 2i + 1. \tag{2.6}\]

The right hand side of Eq. (2.1) contains an \( \hbar \), which we have set to unity. However in the classical limit it is instead set to zero, in which case the lowest energy state of \( H \) becomes a classical vacuum, such as
\[
|0\rangle = |\uparrow\downarrow\uparrow\downarrow \cdots \rangle. \tag{2.7}\]

sector of interest corresponds to the constant bin states that we will introduce in Sec. 4.
which corresponds to
\[ m(i) = 2i. \] (2.8)

Below we will largely consider \( a(m(i) = 2i) \) for the vacuum state
\[ a(m(i) = 2i) = \langle 0 | \Omega \rangle. \] (2.9)

### 2.2 The Thermodynamic Limit

It is conventional to introduce the spectral parameters \( \lambda(i) \), related to \( K(i) \) and \( \Phi(i) \) by
\[
e^{iK(j)} = \left( \frac{\lambda(j) + \frac{i}{2}}{\lambda(j) - \frac{i}{2}} \right), \quad e^{i\Phi(j,k)} = \left( \frac{\lambda(j) - \lambda(k) + i}{\lambda(j) - \lambda(k) - i} \right)\] (2.10)

so that
\[
K(j) = -i \ln \left( \frac{\lambda(j) + \frac{i}{2}}{\lambda(j) - \frac{i}{2}} \right) = 2 \text{ArcCot}(2\lambda(j)) \] (2.11)
and
\[
\Phi(j,k) = -i \ln \left( \frac{\lambda(j) - \lambda(k) + i}{\lambda(j) - \lambda(k) - i} \right) = 2 \text{ArcCot}(\lambda(j) - \lambda(k)). \] (2.12)

We recall that \( K \in [0, 2\pi] \) and \( \Phi \in [-\pi, \pi] \) and so in Eq. (2.11) \( \text{ArcCot} \in [0, \pi] \) but in Eq. (2.12) it is in \([-\pi/2, \pi/2]\). Using (2.10) Bethe’s equation (2.5) can be rewritten as a condition on the spectral parameters
\[
\left( \frac{\lambda(j) + \frac{i}{2}}{\lambda(j) - \frac{i}{2}} \right)^N = \prod_{k \neq j} \left( \frac{\lambda(j) - \lambda(k) + i}{\lambda(j) - \lambda(k) - i} \right). \] (2.13)

It will prove more convenient to rewrite Bethe’s equation using (2.11) and (2.12) to obtain
\[
\frac{\pi}{2} - \text{ArcCot}(2\lambda(j)) = \pi \left( \frac{1}{4} - \frac{Q(j)}{N} \right) + \frac{1}{N} \sum_{k \neq j} \left( \frac{\pi}{2} - \text{ArcCot}(\lambda(j) - \lambda(k)) \right) \] (2.14)

We have kept \( \text{ArcCot}(\lambda(j) - \lambda(k)) \in [-\pi/2, \pi/2] \) and \( \text{ArcCot}(2\lambda(j)) \in [0, \pi] \). We would like to replace the \( \pi/2 - \text{ArcCot} \) above with \( \text{ArcTan} \), where \( \text{ArcTan} \in [-\pi/2, \pi/2] \). However, \( \pi/2 - \text{ArcCot}(\lambda(j) - \lambda(k)) \in [0, \pi] \). Therefore we will need to subtract \( \pi \) whenever \( \text{ArcCot}(\lambda(j) - \lambda(k)) \) is negative, which occurs when \( \lambda(j) < \lambda(k) \). We will choose the \( \lambda(j) \) to be monotonically increasing in \( j \), and so we will need to subtract \( \pi \) for each \( k > j \). In other words, to bring \( \text{ArcTan} \) into the fundamental domain we must subtract \( \pi(n - j) \) from the sum, which must be added to the \( Q \) term. Restricting attention to the ground state \( N = 2n \),
\( Q(j) = 2n - 2j + 1 \), the spectral parameters are the solutions of
\[
\text{ArcTan}(2\lambda(j)) = \pi \left( \frac{1}{4} - \frac{Q(j)}{2n} + \frac{n - j}{2n} \right) + \frac{1}{2n} \sum_{k \neq j}^{n} \text{ArcTan}(\lambda(j) - \lambda(k))
\]
\[
= \pi \left( -\frac{1}{4} + \frac{j}{2n} \right) + \frac{1}{2n} \sum_{k \neq j}^{n} \text{ArcTan}(\lambda(j) - \lambda(k)). \tag{2.15}
\]

To pass to the continuum limit, one replaces the lattice site index \( j \in [1, n] \) with
\[
x(j) = -\frac{1}{4} + \frac{j}{2n} \in \left[ -\frac{1}{4}, \frac{1}{4} \right]. \tag{2.16}
\]
Sometimes it is convenient to replace \( j \) by \( j - 1/2 \) in this expression to make it symmetric in \( x \to -x \), however this will only affect subdominant contributions in \( 1/n \) and will not affect our main results here. Now all functions \( f(j) \) can be replaced by interpolating functions \( \tilde{f}(x) \), by demanding
\[
\tilde{f}(x(j)) = f(j). \tag{2.17}
\]
By abuse of notation, we will drop the tildes and write simply \( f \) for both the original discrete function and its continuous interpolation. The interpolation is not uniquely defined, however if one imposes (2.15) then the choice of interpolation is irrelevant, since the equation only restricts the values at integral points where (2.17) fully determines \( f(x) \).

To fix \( f(x) \) at all \( x \in \left[ -\frac{1}{4}, \frac{1}{4} \right] \), one replaces the sum in Eq. (2.15) with an integral
\[
\frac{1}{2n} \sum_{k \neq j}^{n} \text{ArcTan}(\lambda(j) - \lambda(k)) \to \int_{-1/4}^{1/4} dy \text{ArcTan}(\lambda(x) - \lambda(y)) \tag{2.18}
\]
so that the spectral function \( \lambda(x) \) is determined by
\[
\text{ArcTan}(2\lambda(x)) = \pi x + \int_{-1/4}^{1/4} dy \text{ArcTan}(\lambda(x) - \lambda(y)). \tag{2.19}
\]
The replacement (2.18) is not an equality. It changes the equation. The solutions \( \lambda(x) \) will not be solutions of the original equation, even at the lattice sites \( x(j) \). It is expected that this correction is subdominant in the \( 1/n \) expansion. However, these subleading corrections to the \( \lambda(x) \) may in principle provide leading contributions to the matrix elements.

### 2.3 The Vacuum State

One can now solve (2.19) to find the above functions of \( x \). First, let us define the density
\[
\rho(x) = \frac{1}{\partial \lambda(x)/\partial x} \tag{2.20}
\]
which is unrelated to the density of phases $\rho(\alpha)$ introduced above. The derivative of
Eq. (2.19) with respect to $x$ is
\[
\frac{2}{1 + 4\lambda(x)^2} \rho(x) = \pi + \int_{-1/4}^{1/4} \frac{dy}{1 + (\lambda(x) - \lambda(y))^2} \rho(x).
\] (2.21)
Now multiply through by $\rho(x)$. The function $\lambda : [-1/4, 1/4] \rightarrow [-\infty, \infty]$ is a bijection and
so we can pull back any function $f(x)$ to obtain $f(\lambda)$. Let $\lambda = \lambda(x)$ and $\mu = \lambda(y)$. This
allows us to rewrite the entire equation using functions of $\lambda$ and $\mu$,
\[
\frac{2}{1 + 4\lambda^2} = \pi \rho(\lambda) + \int_{-\infty}^{\infty} \frac{\rho(\mu)d\mu}{1 + (\lambda - \mu)^2}
\] (2.22)
where the integration measure was converted using
\[
dy = \rho(\mu)d\mu.
\] (2.23)

The equation (2.22) is usually solved using Fourier transforms. We will review the argument here, as we need to go a few steps beyond the textbook treatment to obtain all functions of $x$ explicitly. The Fourier transform of the left hand side, omitting the factor of two for now, is
\[
\int_{-\infty}^{\infty} e^{i\lambda x} d\lambda
\] (2.24)
The integrand has simple poles at $\lambda = \pm i/2$. If $\alpha > 0$ ($\alpha < 0$) then the integrand vanishes
exponentially for a large semicircular contour on the upper (lower) half of the complex plane.
The corresponding contour encircles the pole at $+i/2$ ($-i/2$), where the residue is $-ie^{-\alpha/2}/4$ ($ie^{\alpha/2}/4$). The contour is counterclockwise (clockwise) and so the residue theorem yields
\[
\int_{-\infty}^{\infty} e^{i\lambda x} d\lambda = \pi \left( \frac{e^{-|\alpha|/2}}{2} \right).
\] (2.25)
Defining the Fourier transform of the density by
\[
\tilde{\rho}(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda x} \rho(\lambda)d\lambda
\] (2.26)
the Fourier transform allows Eq. (2.22) to be rewritten
\[
\text{LHS} = \frac{2}{1 + 4\lambda^2} = \frac{1}{2} \int_{-\infty}^{\infty} e^{-i\lambda \alpha} e^{-|\alpha|^2/2} d\alpha
\] (2.27)
and the right hand side
\[
= \text{RHS} = \pi \int_{-\infty}^{\infty} d\alpha e^{-i\lambda \alpha} \tilde{\rho}(\alpha) + \int_{-\infty}^{\infty} d\alpha \tilde{\rho}(\alpha) \int_{-\infty}^{\infty} e^{-i\mu \alpha} d\mu \int_{-\infty}^{\infty} \frac{e^{-|\alpha|^2}}{1 + (\lambda - \mu)^2} d\alpha
\] (2.28)
where the integral over $\mu$ was performed as in Eq. (2.25). Taking the Fourier transform of this equation yields

$$\frac{e^{-|\alpha|/2}}{2} = \pi \tilde{\rho}(\alpha) \left[ 1 + e^{-|\alpha|} \right]$$

(2.29)

and so the Fourier transformed density is

$$\tilde{\rho}(\alpha) = \frac{1}{2\pi} \frac{1}{e^{|\alpha|/2} + e^{-|\alpha|/2}}.$$  (2.30)

To obtain the density, one need only Fourier transform Eq. (2.30). First note that on the real line it is equal to the analytic function given by simply removing the absolute values. With a small perturbation which can later be removed, this function shrinks exponentially on either the positive or negative semicircle of the complex plane. Let us choose the positive semicircle. This contour encircles the poles at

$$\alpha = \pi i (2k + 1), \quad k \in \mathbb{Z}$$  (2.31)

where the residues are $-i(-1)^k e^{\pi(2k+1)\lambda}/(2\pi)$. Therefore the density is

$$\rho(\lambda) = \int_{-\infty}^{\infty} e^{-i\alpha\lambda} \tilde{\rho}(\alpha) d\alpha = \sum_{k=0}^{\infty} (-1)^k e^{\pi(2k+1)\lambda} = \frac{e^{\pi\lambda}}{1 + e^{2\pi\lambda}} = \frac{1}{2 \cosh(\pi\lambda)}.$$  (2.32)

Thus

$$\frac{d\lambda}{dx} = \frac{1}{\lambda} = 2 \cosh(\pi \lambda).$$  (2.33)

This equation is the starting point of studies of the thermodynamics of this model.

We will need explicit expressions for the various functions of $x$. To find these, we must solve Eq. (2.33). Multiplying through by $dx/2\cosh$ and integrating one obtains

$$x + \frac{1}{4} = \int_{-1/4}^{x} dx = \int_{-\infty}^{\lambda(x)} \frac{d\lambda}{e^{\pi\lambda} + e^{-\pi\lambda}} = \int_{-\infty}^{\lambda(x)} d\lambda \sum_{k=0}^{\infty} (-1)^k e^{\pi(2k+1)\lambda}$$

$$= \sum_{k=0}^{\infty} (-1)^k \frac{e^{\pi(2k+1)\lambda}}{\pi(2k+1)} \bigg|^{\lambda(x)}_{-\infty} = \sum_{k=0}^{\infty} (-1)^k \frac{e^{\pi(2k+1)\lambda(x)}}{\pi(2k+1)} = \frac{-i}{2\pi} \ln \left( \frac{1 + ie^{\pi\lambda(x)}}{1 - ie^{\pi\lambda(x)}} \right)$$

$$= \frac{1}{\pi} \text{ArcTan} \left( e^{\pi\lambda(x)} \right)$$  (2.34)

which is easily inverted to obtain

$$\lambda(x) = \frac{1}{\pi} \ln \left( \text{Tan} \left[ \pi \left( x + \frac{1}{4} \right) \right] \right).$$  (2.35)
Substituting this into Eqs. (2.11) and (2.12) gives the needed results

\[ K(x) = 2 \arccot \left( \frac{2}{\pi} \ln \left( \tan \left( \frac{\pi}{4} \left(x + \frac{1}{4}\right) \right) \right) \right) \]  
(2.36)

\[ \Phi(x, y) = 2 \arccot \left( \frac{1}{\pi} \ln \left( \frac{\tan \left( \frac{\pi}{4} \left(x + \frac{1}{4}\right) \right)}{\tan \left( \frac{\pi}{4} \left(y + \frac{1}{4}\right) \right)} \right) \right). \]  
(2.37)

Similarly one finds

\[ \rho(x) = \frac{\cos(2\pi x)}{2}. \]  
(2.38)

Note that the function \( \lambda(x) \) given in Eq. (2.35) is an exact solution of the continuum equation (2.19) but not of the exact discrete equation (2.15). At large \( n \) with \( j \) constant, the \( j \)th equation in Eq. (2.15) is violated by \( c_j/n \) where \( c_j \) is independent of \( n \) to leading order. The left hand side is always larger. Numerically we have found \( c_1 = 0.45, c_{10} = 0.18, c_{100} = 0.009, c_{1000} = 0.004 \) and so on. Shifting an individual \( \lambda \) to adjust for this shift yields a change of order \( 1/n \). However it is not obvious that when all \( \lambda \) are consistently adjusted together, the correction will vanish at large \( n \) at fixed \( x \). In fact in Subsec. 5 we will see a numerical discrepancy of order \( 1\% \) which is likely an error on our part, but could conceivably be a sign of such a finite correction. That said, we have used (2.15) to solve for the left hand side, which we then substituted into the right hand side and so on iteratively 200 times with chain lengths of thousands and we found that convergence appears to arrive after of order 100 recursions, with a total change in \( \lambda \) of less than about \( 1\% \) at every site, and much less than \( 1\% \) far from the boundaries.

3 The Anchor

Recall that the coordinate Bethe Ansatz expresses the matrix elements in the form

\[ a = \sum_{g \in S_n} e^{i\alpha(g)} \]  
(3.1)

where the phase

\[ \alpha(g) = \alpha_1(g) + \alpha_2(g), \quad \alpha_1(g) = \sum_{j=1}^{n} m(j) K(P(j)), \quad \alpha_2(g) = \frac{1}{2} \sum_{j<k} \Phi(P(j), P(k)) \]  
(3.2)

depends on the permutation \( g \in S_n \). To perform this sum at large \( n \), where it becomes an integral, one would like the density \( \rho(\alpha) \). Eq. (1.1) states that the matrix elements are given by the Fourier transform of \( \rho \). This function is unfortunately rich in substructure, at just
the scale which contributes to the Fourier transform, and so its evaluation is a difficult task. The role of the anchor is to shift
\[ \alpha(g) \rightarrow \alpha'(g) = \alpha(g) - \alpha_3(g) \] (3.3)
so as to cancel out the substructure. The anchor \( \alpha_3(g) \) will be a multiple of \( 2\pi \) and so the shift will not affect the matrix elements. To see how it works, and to motivate it, we will first consider the substructure created by two actions of the cyclic group \( \mathbb{Z}_n \).

### 3.1 Type I Cyclic Permutations

One can define a free action of the cyclic group \( \mathbb{Z}_n \) on the permutation group \( S_n \) as follows. Let the generator \( 1 \in \mathbb{Z}_n \) act on \( g \in S_n \) by
\[ 1 : S_n \rightarrow S_n : g \mapsto g' : P(j) \mapsto P'(j) = P(j + 1 \mod n). \] (3.4)
Restrict our attention to matrix elements with classical ground state \( |0\rangle \), which corresponds to \( m(j) = 2j \). In this case, and only in this case, we will now show that \( \mathbb{Z}_n \) is an exact symmetry of the phases
\[ e^{i\alpha(g)} = e^{i\alpha(g')} \] (3.5)
Indeed, \( \alpha(g') \) is easily calculated
\[
\begin{align*}
\alpha(g') &= \sum_{j=1}^{n} m(j) K(P(j + 1 \mod n)) + \frac{1}{2} \sum_{j<k}^{n} \Phi(P(j + 1 \mod n), P(k + 1 \mod n)) \\
&= \sum_{j=2}^{n} m(j - 1) K(P(j)) + m(n) K(P(1)) \\
& \quad + \frac{1}{2} \sum_{2<j<k}^{n} \Phi(P(1), P(j), P(k)) + \frac{1}{2} \sum_{j=2}^{n} \Phi(P(j), P(1)).
\end{align*}
\] (3.6)
Now, fixing \( m(j) = 2j \) we find
\[
\alpha(g') - \alpha(g) = -2 \sum_{j=2}^{n} K(P(j)) + 2(n-1) K(P(1)) - \frac{1}{2} \sum_{j=2}^{n} \Phi(P(1), P(j)) + \frac{1}{2} \sum_{j=2}^{n} \Phi(P(j), P(1)).
\] (3.7)
Using the antisymmetry of \( \Phi \) this simplifies to
\[
\alpha(g') - \alpha(g) = -2 \sum_{j=1}^{n} K(P(j)) + 2n K(P(1)) - \sum_{j=2}^{n} \Phi(P(1), P(j)).
\] (3.8)
$K$ is symmetrically distributed about $\pi$ and so the first term on the right hand side is just $-2n\pi$. Bethe’s equation (2.5) on the other hand gives the sum of the second and third terms to be $2\pi Q(P(1))$. Putting this all together we obtain

$$\alpha(g') - \alpha(g) = 2\pi(-n + 2n - 2P(1) + 1) = 2\pi(n - 2P(1) + 1).$$  \hspace{1cm} (3.9)

This is an integer multiple of $2\pi$. Thus we have shown that these cyclic permutations leave each summand in the matrix elements invariant, and yet they affect the arguments $\alpha(g)$ and so complicate the distribution $\rho(\alpha)$. Clearly, to calculate this distribution, it would be desirable to remove these spurious shifts. How can this be done?

Consider a second action of the generator of the cyclic group. Now

$$\alpha(g'') - \alpha(g') = 2\pi(n - 2P'(1) + 1) = 2\pi(n - 2P(2) + 1)$$  \hspace{1cm} (3.10)

and so

$$\alpha(g'') - \alpha(g) = 2\pi(2n - 2P(2) - 2P(1) + 2).$$  \hspace{1cm} (3.11)

In general the element $k$ of the cyclic group shifts the arguments by

$$\alpha(g^{(k)}) - \alpha(g) = 2\pi\sum_{i=1}^{k}(n - 2P(i) + 1).$$  \hspace{1cm} (3.12)

How can we modify $\alpha(g)$ to prevent these spurious shifts? Recall that $P : [1, n] \to [1, n]$ is a bijection and so it is invertible and the inverse transforms under the cyclic action by

$$(P')^{-1}(j) = P^{-1}(j) - 1.$$  \hspace{1cm} (3.13)

Choose any integer $k \in [1, n]$ and define

$$\alpha_3(g) = -2\pi\sum_{i=1}^{P^{-1}(k)}(n - 2P(i) + 1).$$  \hspace{1cm} (3.14)

How does this transform?

$$\alpha_3(g') = -2\pi\sum_{i=1}^{(P')^{-1}(k)}(n - 2P'(i) + 1)$$  \hspace{1cm} (3.15)

$$= -2\pi\sum_{i=1}^{P^{-1}(k)-1}(n - 2P(i + 1) + 1) = -2\pi\sum_{i=2}^{P^{-1}(k)}(n - 2P(i) + 1)$$

and so the difference is

$$\alpha_3(g') - \alpha_3(g) = 2\pi(n - P(1) + 1) = \alpha(g') - \alpha(g).$$  \hspace{1cm} (3.16)
As $\alpha_3$ and $\alpha$ transform identically under the cyclic permutations, their difference $\alpha(g) - \alpha_3(g)$ is invariant. Thus $\alpha_3(g)$ so defined is an anchor which fixes these cyclic permutations. However it is not the only such anchor. One may add to it any other integral multiple of $2\pi$ which is invariant under these cyclic transformations and so obtain another such anchor. Below we will see that there is another cyclic action which is not fixed, and so this choice of $\alpha_3(g)$ is not optimal. In fact we have checked numerically that this $\alpha_3$ does reduce the scatter in the distribution of the phase arguments. At $N \sim 15$ the reduction is about a factor of 2 but by $N \sim 20$ it is negligible as each orbit of the $\mathbb{Z}_n$ action is quite small in $S_n$. It also leaves considerable substructure and so is not sufficient for the calculation of matrix elements.

3.2 Type II Cyclic Permutations

The symmetric group $S_n$ admits another free $\mathbb{Z}_n$ action, whose generator acts by

$$1 : g \rightarrow g' : P'(j) = P(j) + 1 \mod n. \tag{3.17}$$

This action does not leave the phases invariant. But, for $x$ not too close to the boundaries, it leaves the phases $e^{i\alpha(g)}$ reasonably invariant while dramatically shifting the arguments $\alpha(g)$. Repeating the calculation as above, with this action, one obtains

$$\alpha(g') - \alpha(g) = \sum_{j \neq P^{-1}(n)} m(j)(K(P(j) + 1) - K(P(j))) + m(P^{-1}(n))(K(1) - K(n))$$

$$\begin{align*}
+ \frac{1}{2} \sum_{j < l, j \neq P^{-1}(n)} (\Phi(P(j) + 1, P(l) + 1) - \Phi(P(j), P(l))) \\
+ \frac{1}{2} \sum_{j=1}^{P^{-1}(n)-1} (-\Phi(1, P(j) + 1) - \Phi(P(j), n)) \\
+ \frac{1}{2} \sum_{j=P^{-1}(n)+1}^{n} (\Phi(1, P(j) + 1) + \Phi(P(j), n)) 
\end{align*} \tag{3.18}$$

As before, the last term on the first line combines with the last two terms, via Bethe’s equations, to yield

$$\Delta = -2\pi \left( n - 2P^{-1}(n) - 1 \right). \tag{3.19}$$

However the other terms do not quite vanish in this case, although for a single action of the generator 1 at $N \sim 20$ we have found numerically that they are generally less than 0.05.
In every case observed they are much smaller than $\Delta$, except in those rare cases when $\Delta$ vanishes exactly.

Again it is not difficult to construct an anchor which reproduces this transformation law

$$\alpha_3(g) = 2\pi \sum_{i=1}^{P(k)-1} (n - 2P^{-1}(i) + 1) \tag{3.20}$$

where again $k$ is arbitrary. However this anchor does not leave the type I cyclic permutations invariant.

### 3.3 A Universal Anchor

We propose the anchor

$$\alpha_3(g) = 2\pi \sum_{j<k} \theta(P(j) - P(k)) \tag{3.21}$$

where

$$\theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise}. \end{cases} \tag{3.22}$$

The trivial permutation $P(j) = j$ gives $\alpha_3(g) = 0$. More generally, this counts the number of pairs of sites whose order is flipped by $g$.

How does it work? For example, begin with the identity permutation $P(j) = j$. Now consider the type I cyclic permutation, it yields

$$P'(j) = P(j + 1) = j + 1 \mod n. \tag{3.23}$$

In this case $P(n) = 1$ and so a single entry has moved from the right to the left of all other $n - 1$ entries, all of which were smaller. Thus the sum gains contributions from all elements with $j = 1$

$$\alpha_3(g') = \alpha_3(g') - \alpha_3(g) = 2\pi \sum_{j<k} \theta(P(j) - P(k)) = 2\pi \sum_{k=2}^{n} 1 = 2\pi(n-1) = 2\pi(n-2P(1)+1). \tag{3.24}$$

And so we see that $\alpha_3$ transforms just like $\alpha$ under this cyclic permutation of type I. In fact, the transformation $\text{[3.23]}$ is not only the generator of type I cyclic permutations, but also type II cyclic permutations, which happen to coincide in this example because $g$ is just a shift. Now

$$P^{-1}(n) = n, \quad \Delta = 2\pi(n-1) = \alpha_3(g') - \alpha_3(g) \tag{3.25}$$
and so the anchor compensates for the type II permutation as well, as it must since this is also a type I permutation.

What about general elements of $S_n$? Beginning with an arbitrary element $g \in S_n$, a type I permutation yields

$$P'(j) = P(j + 1 \mod n) \quad (3.26)$$

and so our anchor transforms to

$$\alpha_3(g') = 2\pi \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \theta(P'(j) - P'(k))$$

$$= 2\pi \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \theta(P(j + 1) - P(k + 1 \mod n))$$

$$= 2\pi \sum_{j=1}^{n-2} \sum_{k=j+1}^{n-1} \theta(P(j + 1) - P(k + 1)) + 2\pi \sum_{j=1}^{n-1} \theta(P(j + 1) - P(1))$$

$$= 2\pi \sum_{j=2}^{n-1} \sum_{k=j+1}^{n} \theta(P(j) - P(k)) + 2\pi \sum_{j=2}^{n} \theta(P(j) - P(1))$$

$$\quad (3.28)$$

yielding a difference of

$$\alpha_3(g') - \alpha_3(g) = 2\pi \sum_{j=2}^{n} \theta(P(j) - P(1)) - 2\pi \sum_{k=2}^{n} \theta(P(1) - P(k))$$

$$= 2\pi(n - P(1)) - 2\pi(P(1) - 1) = 2\pi(n - 2P(1) + 1)$$

which equals $\alpha(g') - \alpha(g)$ calculated in Eq. (3.9). Therefore $\alpha(g) - \alpha_3(g)$ is invariant under type I permutations.

What about type II permutations? Now

$$P'(j) = P(j) + 1 \mod n \quad (3.30)$$
\[
\alpha_3(g') = 2\pi \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \theta(P'(j) - P'(k))
\] (3.31)

\[
= 2\pi \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \theta((P(j) + 1 \mod n) - (P(k) + 1 \mod n))
\]

\[
= 2\pi \sum_{j=1, j \neq P^{-1}(n)}^{n-1} \sum_{k=j+1, k \neq P^{-1}(n)}^{n} \theta((P(j) - P(k)))
\]

\[
= 2\pi \sum_{j=1, j \neq P^{-1}(n)}^{n-1} \sum_{k=j+1, k \neq P^{-1}(n)}^{n} \theta(P(j) - P(k)) + 2\pi (P^{-1}(n) - 1).
\] (3.32)

The difference is then

\[
\alpha_3(g') - \alpha_3(g) = -2\pi \sum_{k=P^{-1}(n)+1}^{n} \theta(n - P(k)) - 2\pi \sum_{j=1}^{P^{-1}(n)-1} \theta(P(j) - n) + 2\pi (P^{-1}(n) - 1)
\]

\[
= -2\pi (n - P^{-1}(n)) + 2\pi (P^{-1}(n) - 1)
\]

\[
= -2\pi (n - 2P^{-1}(n) + 1) = \Delta
\] (3.33)

which agrees with the approximation to the shift in \(\alpha(g)\) found in Subsec. 3.2. Therefore \(\alpha(g) - \alpha_3(g)\) is approximately invariant under both kinds of cyclic permutations.

We need more. We need \(\alpha(g) - \alpha_3(g)\) to be free of substructure, so that its moments can be used to construct correlation functions. If it is Gaussian distributed, then its standard deviation should be of order \(O(n^{1/2})\), not of order \(O(n^{3/2})\) like \(\alpha(g)\). In Sec. 5 we will investigate this second property. In the case of the vacuum matrix element \(\langle 0 | \Omega | 0 \rangle\) at \(N = 22\), so that \(n = 11\), these properties are demonstrated numerically in Fig. 1. One sees that the full width have maximum of \(\alpha\) is about 70, which is about \(2n^{3/2}\) as expected. On the other hand \(\alpha(g) - \alpha_3(g)\) is much thinner, with a full width half maximum of only about 6, which is of order \(2n^{1/2}\) as desired. Therefore using a Gaussian approximation and a Fourier transform to obtain the matrix elements in Eq. (1.1) would yield very different answers with and without the anchor, although of course the Fourier transforms must agree as \(\alpha_3(g)\) is an integer multiple of \(2\pi\). This disagreement is reasonable as the Gaussian approximation
Figure 1: Histograms of the distributions of $\alpha(g)$ (top-left) and $\alpha(g) - \alpha_3(g)$ (others) computed numerically at $N = 22$ for $\langle 0 | \Omega \rangle$. The matrix element is dominated by structure at scales near $2\pi$. $\alpha(g)$ has rich substructure at this scale, which dominates the matrix element. $\alpha(g) - \alpha_3(g)$ is much thinner, with no evidence of substructure at this scale. In the bottom panel one sees that $\alpha(g) - \alpha_3(g)$ closely fits a Gaussian of deviation $1.08\pi$ (black curve), although there is slight leptokurtosis. The bin width is 0.1 and cyclic permutations of type I fix $P(1) = 1$. 
is a very poor fit to $\alpha(g)$, which has rich substructure at order unity which dominates the Fourier transform. On the other hand, we see in the bottom panel that a Gaussian provides an excellent fit to the anchored $\alpha(g) - \alpha_3(g)$. If this has any substructure, it lies at scales far beneath $2\pi$ where it has little effect on Eq. (1.1) and so the matrix elements.

This is our first main result. With the anchor (3.21), the distribution of phases $\rho(\alpha)$ in the CBA becomes approximately a Gaussian and so the calculation of the matrix elements in Eq. (1.1) requires only that one determine its moments. In the rest of this note, we will describe a method for the calculation of these moments.

4 Binning

Exact calculations of matrix elements have been a major industry for decades. However as we are interested in continuum field theory, our goal is somewhat different. It is more difficult, because we will need a method which calculates matrix elements for states which differ at arbitrarily many lattice sites from any given reference state. This distance is in general infinite, and so if our proposal requires a computation time which is polynomial in this distance then we are lost. That said, we do not need a closed form answer. It is sufficient to present a method for the calculation of any matrix element, so long as the time required for a given precision, as measured in units accessible to the continuum field theory, does not increase with $N$ but only with some suitable measure of the complexity of the state. Our task is also easier because we are not interested in all states. We are only interested in those states which survive the continuum limit. In particular, nearby lattice sites should have similar behaviors, in the sense that they map nearby pairs of lattice sites to the same target space point via the map in Ref. [5].

4.1 The Binning

This motivates the following approach. We will divide the interval $[1, n]$ into $q$ bins

$$S_i = \left[\frac{n}{q} (i - 1) + 1, \frac{n}{q} i\right]. \quad (4.1)$$

Recall that an element $g \in S_n$ is completely characterized by a bijection $P : [1, n] \to [1, n]$. Let

$$f_{ij}(g) = \sum_{k \in S_i} \sum_{l \in S_j} \delta_{P(k), l} = |f(S_i) \cap S_j| \quad (4.2)$$

In other words, $f_{ij}(g)$ is the number of entries of $S_i$ which $g$ maps into $S_j$. Clearly $f_{ij}(g)$ contains only some of the information in $P$, while $P$ is equivalent to $g$. We will rely upon
**The Binning Postulate:** For the calculation of a given quantity $X$ to any precision $\epsilon > 0$, there exists a sufficiently high $q(\epsilon)$ such that, if $X$ is calculated replacing all $g$ with the same $\{ f_{ij} \}$ by the same $g_f$ then the introduced error in $X$ will be bounded by $\epsilon$.

It may be that the binning postulate is false, or that it is true only at some leading orders in $N$. Certainly it is false for a sufficiently pathological $X$, and so it is our hope that it is true for all $X$ accessible in the continuum field theory. In other words, we conjecture that each point in the continuum field theory corresponds to a bin on the spin chain, and so none of the bins’ internal structure survives in the continuum field theory.

With these strong conjectures in hand our strategy is clear. We will recast our problem in terms of $f$, assuming that with a suitable choice of $g_f$ the intrabin contributions to various quantities vanish in the $q \to \infty$ limit. We have checked this in some cases.

Now our binning approximation is

$$
\alpha_1(g) = \sum_j^n m(j) K(P(j)) \sim \alpha_1(f) = \sum_{i,j=1}^q m \left( \frac{n}{q} \right) f_{ij}(g) K \left( \frac{n}{q} \right).
$$

$$
\alpha_2(g) = \frac{1}{2} \sum_{j<k}^n \Phi(P(j), P(k)) \sim \alpha_2(f) = \frac{1}{2} \sum_{i<k}^q \sum_{j,l=1}^q f_{ij}(g) f_{kl}(g) \Phi \left( \frac{n}{q}, \frac{n}{q} \right).
$$

$$
\alpha_3(g) = 2\pi \sum_{j<k} \theta(P(j) - P(k)) \sim \alpha_3(f) = 2\pi \sum_{i<k} \sum_{j,l} f_{ij}(g) f_{kl}(g).
$$

$$
\alpha(f) = \alpha_1(f) + \alpha_2(f) - \alpha_3(f).
$$

Using this decomposition, the moments of the distribution $\rho(\alpha)$ of $\alpha(g)$ can be approximated by those of $\alpha(f)$ which can be calculated from those of $f$.

### 4.2 Simplifications at First Order

This can be somewhat simplified. First note that each of the $n/q$ elements of $S_i$ is mapped to some $S_j$ by $P$. This yields the sum rule

$$
\sum_{j=1}^q f_{ij}(g) = \frac{n}{q}.
$$

Similarly all $n/q$ elements of $S_j$ are in some $P(S_i)$ yielding the second sum rule

$$
\sum_{i=1}^q f_{ij}(g) = \frac{n}{q}.
$$

19
These sum rules hold individually for every $g \in S_n$.

Let us define the expectation value of $f_{ij}$ by

$$\langle f_{ij} \rangle = \frac{1}{n!} \sum_{g \in S_n} f_{ij}(g). \quad (4.6)$$

Higher correlators are defined similarly. It is quite clear that $\langle f_{ij} \rangle$ is independent of $i$ and $j$. Therefore the expectation value of either sum rule yields

$$\langle f_{ij} \rangle = \frac{1}{q} \langle \sum_{i=1}^{q} f_{ij} \rangle = \frac{1}{q} \langle n \rangle = \frac{n}{q^2}. \quad (4.7)$$

This quantity will appear so often that we will name it

$$\beta = \frac{n}{q^2}. \quad (4.8)$$

Many quantities are more simply expressed in terms of the reduced

$$\tilde{f}_{ij}(g) = \frac{f_{ij}(g)}{\beta} - 1. \quad (4.9)$$

From the corresponding properties of $f_{ij}$ one finds

$$\sum_{i=1}^{q} \tilde{f}_{ij}(g) = \sum_{j=1}^{q} \tilde{f}_{ij}(g) = 0, \quad \langle \tilde{f}_{ij} \rangle = 0. \quad (4.10)$$

We can now use (4.3) to express the Bethe phases in terms of $\tilde{f}$. The first is

$$\alpha_1(f) = \beta \sum_{i,j=1}^{q} m \left( \frac{n}{q} ight) (1 + \tilde{f}_{ij}) K \left( \frac{n}{q j} \right). \quad (4.11)$$

Let us fix our reference state to be the classical vacuum $|0\rangle$ and so $m(j) = 2$. Then this becomes

$$\alpha_1(f) = \beta \sum_{i,j=1}^{q} 2i \frac{n}{q} (1 + \tilde{f}_{ij}) K \left( \frac{n}{q j} \right). \quad (4.12)$$

$$= 2\beta \frac{n}{q} \left( \sum_{i=1}^{q} i \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q j} \right) \right) + \frac{2n^2}{q^3} \sum_{i,j=1}^{q} i \tilde{f}_{ij} K \left( \frac{n}{q j} \right)$$

$$= n^2 \pi \left( 1 + \frac{1}{q} \right) + \frac{2n^2}{q^3} \sum_{i,j=1}^{q} i \tilde{f}_{ij} K \left( \frac{n}{q j} \right)$$
where we used the fact that \( K(i) \) is symmetric about \( \pi \). The \( 1/q \) correction to the first term is an artefact of our treatment of interbin effects, and could be changed if we changed our prescription for these by, for example, adding terms \( \alpha_4 \) to consider cases in which \( i = k \) but nonetheless a given element of \( S_i \) is less than one of \( S_k = S_i \) and so should be included in the sum. The binning postulate states that such corrections should not appear in continuum field theory observables.

Next we will treat \( \alpha_2(f) \)

\[
\alpha_2(f) = \frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} (1 + \tilde{f}_{ij})(1 + \tilde{f}_{kl}) \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right). \tag{4.13}
\]

Note that the term with no \( \tilde{f} \) vanishes because

\[
\sum_{j,l=1}^{q} \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right) \sim \frac{q^2}{n^2} \sum_{j,l=1}^{n} \Phi (j,l) = 0. \tag{4.14}
\]

It may appear that the term linear in \( \tilde{f} \) vanishes as a result of the sum rule, but it does not as \( i < k \) and so it is not summed over all bins. However \( j \) and \( l \) are summed over all bins, and so we can apply the binned version of the Bethe equation (2.5), which in the case of the vacuum state \( |\Omega\rangle \) is

\[
2qK(j) = 2\pi \left( 2q - 2j + \frac{q}{n} \right) + \sum_{l \neq j} \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right). \tag{4.15}
\]

Now we are ready to evaluate the terms linear in \( \tilde{f} \). It turns out that they are equal, so we will show the evaluation of the \( \tilde{f}_{ij} \) term

\[
\frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right) = \frac{1}{2} \beta^2 \sum_{i} \left( \sum_{j=i+1}^{q} \tilde{f}_{ij} \left( 2qK(j) + 2\pi \left( -2q + 2j - \frac{q}{n} \right) \right) \right)
\]

\[
= \beta^2 \sum_{i} \left( q - i \right) \sum_{j=1}^{q} \tilde{f}_{ij} \left( qK(j) + 2\pi \left( -q + j - \frac{q}{2n} \right) \right) \tag{4.16}
\]

which can be cleaned using the sum rule

\[
\frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right) = -\frac{n^2}{q^2} \sum_{i,j}^{q} i \tilde{f}_{ij} K(j) - 2\pi \beta^2 \sum_{i,j}^{q} i \tilde{f}_{ij} j. \tag{4.17}
\]

As the \( \tilde{f}_{kl} \) term is equal to the \( \tilde{f}_{ij} \) term, we have found

\[
\alpha_2(f) = -4\pi \beta^2 \sum_{i,j}^{q} i \tilde{f}_{ij} j - 2\frac{n^2}{q^2} \sum_{i,j}^{q} i \tilde{f}_{ij} K(j)
\]

\[
+ \frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \tilde{f}_{kl} \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right). \tag{4.18}
\]
Here we see our first major cancellation. The second term of $\alpha_2(f)$ exactly cancels the second term in $\alpha_1(f)$ as written in Eq. (4.12). Thus the function $K$ disappears from the phase factor $\alpha(f)$, and only a constant remains of $\alpha_1(f)$.

Finally we turn to $\alpha_3(f)$

$$\alpha_3(f) = 2\pi^2 \beta^2 \sum_{i<k} \sum_{j>l} (1 + \tilde{f}_{ij})(1 + \tilde{f}_{kl}).$$  \hspace{1cm} (4.19)

The term with no $\tilde{f}$ is easily evaluated

$$2\pi^2 \beta^2 \sum_{i<k} \sum_{j>l} 1 = 2\pi^2 \beta^2 \left( \frac{q(q-1)}{2} \right)^2 = \frac{\pi}{2} n^2 \left( 1 - \frac{1}{q} \right)^2.$$  \hspace{1cm} (4.20)

This cancels half of the remaining constant term in $\alpha_1(f)$ in Eq. (4.12). These constant terms then yield

$$\langle \alpha_1 + \alpha_2 - \alpha_3 \rangle = \frac{n^2}{2} \pi \left( 1 + \frac{1}{q^2} \right).$$  \hspace{1cm} (4.21)

In fact the expectation values for $\langle \alpha_1 + \alpha_2 \rangle$ and for this full anchored combination are visible in Fig. 4 and one indeed sees that the later is a bit more than half of the former. Why a bit more? Should not $1/q^2$ be an artifact? When $n \to \infty$, $n/q^2$ should either tend to a constant or else go to zero more slowly than $1/q$. And so one expects that a $1/q^2$ correction will be a $1/n$ correction. Such a $1/n$ correction is expected, as we have made a rather arbitrary choice in definition of $\alpha_3(g)$ in Eq. (3.21). We have not included contributions from the terms $j = k$. If we include these contributions, then the anchor is increased by $2\pi n$ and so the expectation value decreases by $2\pi n$. In this case the expectation value of the anchored phase is slightly less than half of the unanchored phase.

The expectation value of $\alpha$ contributes a phase to the matrix elements, and so needless to say we need to be concerned about an $O(n)$ change in its expectation value. The fact that such subleading effects, as subtle as the choice of whether to include in $j = k$ in the anchor, may have such a large effect on our results means that care will be needed, in particular in such zero point effects which can leak into the next order on $n$. After all, we will see in the next section that the leading contributions to the variance of our distribution $\rho(\alpha)$ nearly cancel, and so the subleading terms may dominate the physics.

Now it is clear that the two linear terms are equal, and so to evaluate their sum we will simply multiply the $\tilde{f}_{ij}$ term by 2

$$4\pi^2 \beta^2 \sum_{i<j} \sum_{j>l} \tilde{f}_{ij} = 4\pi^2 \beta^2 \sum_{i,j=1}^q (q - i)j \tilde{f}_{ij} = -4\pi^2 \beta^2 \sum_{i,j=1}^q i \tilde{f}_{ij}.$$.  \hspace{1cm} (4.22)
This is equal to the first term in $\alpha_2(f)$ as written in Eq. (4.18), leading to our second major cancelation. Putting all remaining terms together we have found our master formula for the anchored phase

$$\alpha(f) = \frac{n^2}{2} \pi \left( 1 + \frac{1}{q^2} \right) + \frac{1}{2} \beta^2 \sum_{i<k} \sum_{j<l} \tilde{f}_{ij} \tilde{f}_{kl} \Phi \left( \frac{n}{q} j, \frac{n}{q} l \right) - 2\pi \beta^2 \sum_{i<k} \sum_{j>l} \tilde{f}_{ij} \tilde{f}_{kl}. \quad (4.23)$$

We will soon see that $\langle \tilde{f} \tilde{f} \rangle \sim 1/n$ at leading order in $n$ and so we may already try to estimate the fluctuations of the anchored phase. The first term is a constant and so does not contribute. The second two have $\beta^2 = n^2 / q^4$. The $q^4$ cancels with the sums, up to a factor of order unity. Now the variance depends on the square of this, and so it will be of order $O(n^4)$. On the other hand the four point function of $\tilde{f}$ in the Gaussian approximation would give $O(1/n^2)$, and so we find a variance of $O(n^2)$ and so a standard deviation of $O(n)$. This is smaller than the $O(n^{3/2})$ which is present with no anchor, but still is bigger than the $O(n^{1/2})$ which is necessary to give matrix elements proportional to $e^{-n}$ in the Gaussian approximation. If the binning postulate is true and $\rho(\alpha)$ is Gaussian, the contributions from the last two terms in (4.23) to the variance must cancel at $O(n^2)$, leaving only $O(n)$ contributions. We will see that such $O(n)$ contributions arise from the Poisson behavior of the 4-point function, and so such a cancellation would mean that the Gaussian correlations between pairs of $\tilde{f}$ must cancel in the matrix elements, leaving only the irreducible contribution to the 4-point function of the $\tilde{f}$. This will be calculated in Sec. 5 and reduced to an identity relating $K$ and $\Phi$ which has only been studied numerically.

### 4.3 Bin Statistics from Partitions: One Point

Finally we are ready to calculate correlation functions of $f$. These are averages of products of $f_{ij}(g)$ over the symmetric group $S_n$. To calculate them, one must count how many members $g \in S_n$ give each value for a given polynomial in $f_{ij}(g)$. Let us warm up by considering a single $f_{ij}$. How many elements of $g$ satisfy $f_{ij}(g) = p$?

Let us call this number

$$h_{ij}(p) = \text{Number of } g \in S_n \text{ such that } f_{ij}(g) = p. \quad (4.24)$$

The probability that a given $g$ satisfies $f_{ij}(g) = p$ is then

$$\tilde{h}_{ij}(p) = \frac{h_{ij}(p)}{n!}. \quad (4.25)$$

Recall that $P$ must map each integer in $[1,n]$ to a distinct integer in $[1,n]$. If $p$ elements of $S_i$ are to map to $S_j$, one needs to choose which $p$ elements of $S_j$ are in $P(S_i)$. Recalling that
each bin has \( \frac{n}{q} \) elements, the number of choices is \( \binom{n/q}{p} \). One must also choose the \( n - p \) elements of the complement of \( S_j \) which are in \( P(S_i) \). The corresponding number of choices is \( \binom{n - n/q}{n/q - p} \). Finally, one may permute the elements of \( S_j \) and its complement, yielding factors of \( (n/q)! \) and \( (n - n/q)! \) respectively. The result is

\[
 h_{ij}(p) = \left( \frac{n}{q} \right) \binom{n - n/q}{n/q - p} (n/q)! (n - n/q)!.
\]

These later factors are independent of \( p \) and so will be ignored in future calculations, as they only contribute to the overall normalization which is fixed by the fact that

\[
 \sum_{p=0}^{\infty} h_{ij}(p) = n!.
\]

So let us separate all of the \( p \)-independent terms into a constant \( c \)

\[
 \tilde{h}_{ij}(p) = \frac{c}{p!} \left( \frac{(n/q)!}{(n/q - p)!} \right)^2 \left( \frac{(n - 2n/q)!}{(n - 2n/q + p)!} \right) .
\]

We will be interested in the limit where \( n/q \to \infty \) while \( p \), which is of order \( \beta = n/q^2 \), will be finite or slowly tend to 0. In this limit the ratios of factorials are of the form \( x!/(x + r)! \) with \( r << \sqrt{x} \). Indeed, \( r \) will be finite and \( \sqrt{x} \) infinite. When \( x = n/q \) and \( r = p \sim n/q^2 \) this inequality implies \( n \lesssim q^3 \).

To find a suitable approximation for the ratios of factorials in this limit, we combine the expansion

\[
 \left( 1 + \frac{c}{n} \right)^n = e^c \left( 1 - \frac{c^2}{2n} + \frac{c^3}{3n^2} + \frac{c^4}{2n^2} - \frac{c^4}{4n^3} - \frac{c^5}{6n^3} - \frac{c^6}{48n^3} + O \left( \frac{1}{n^4} \right) \right)
\]

with Stirling’s approximation

\[
 n! = \sqrt{2\pi n} n^e \left( 1 + \frac{1}{12n} + \frac{1}{288n^2} + O \left( \frac{1}{n^3} \right) \right)
\]

and the binomial expansion to obtain our main tool

\[
 \frac{x!}{(x + r)!} \sim x^{-r} \left[ 1 + \frac{-r - r^2}{2x} + \frac{2r + 9r^2 + 10r^3 + 3r^4}{24x^2} + \frac{-6r^2 - 17r^3 - 17r^4 - 7r^5 - r^6}{48x^3} \right].
\]

With this tool in hand, we can approximate \( h \). If we let \( q \sim O(n^{1/2}) \) and expand to order \( O(n^{-1}) \), for example, we find

\[
 \tilde{h}_{ij}(p) = \frac{c}{p!} \beta^p \left[ 1 + \frac{(1 + 2\beta)p - p^2}{n/q} + \left( \frac{12\beta - 3 - \frac{1}{\beta}}{6n} \right) p + \left( \frac{12\beta + 9 + \frac{2}{\beta}}{6n} \right) p^2 + \left( -12 - \frac{8}{\beta} \right) p^3 + \frac{3}{\beta} p^4 \right].
\]
Note that the leading term is a Poisson distribution times $e^\beta c$. Therefore the expectation value of any function of $f$ can be given in terms of Poisson correlators

$$
\langle p \rangle_0 = \beta, \quad \langle p^2 \rangle_0 = \beta^2 + \beta, \quad \langle p^3 \rangle_0 = \beta^3 + 3\beta^2 + \beta, \quad \langle p^4 \rangle_0 = \beta^4 + 6\beta^3 + 7\beta^2 + \beta, \ldots
$$

(4.33)

In particular, by setting the expectation value of 1 to be equal to 1, we can fix $c$ at any desired order. In this case the relation between Eq. (4.32) and the Poisson distribution yields

$$
1 = \langle 1 \rangle = e^\beta c \left[ \langle 1 \rangle_0 + \frac{(1 + 2\beta)\langle p \rangle_0 - \langle p^2 \rangle_0}{n/q} \right. \\
+ \left. \frac{(12\beta - 3 - \frac{1}{\beta})}{6n} \langle p \rangle_0 + \left(12\beta + 9 + \frac{6}{\beta}\right) \langle p^2 \rangle_0 + \left(-12 - \frac{8}{\beta}\right) \langle p^3 \rangle_0 + \frac{3}{\beta} \langle p^4 \rangle_0 \right].
$$

(4.34)

Then inserting the Poisson expectation values from Eq. (4.33) one finds

$$
1 = e^\beta c \left[ 1 + \frac{\beta^2}{n/q} + \frac{3\beta^3 + 7\beta^2 - 3\beta}{6n} \right]
$$

(4.35)

and so obtains $c$ at $O(n^{-1})$

$$
c = \frac{e^{-\beta}}{1 + \frac{\beta^2}{n/q} + \frac{3\beta^3 + 7\beta^2 - 3\beta}{6n}}.
$$

(4.36)

Any other correlator can be found similarly, using (4.32) to relate the desired correlator to a combination of Poisson correlators. For example,

$$
\langle f_{ij} \rangle = \langle p \rangle = e^\beta c \left[ \langle p \rangle_0 + \frac{(1 + 2\beta)\langle p^2 \rangle_0 - \langle p^3 \rangle_0}{n/q} \right. \\
+ \left. \frac{(12\beta - 3 - \frac{1}{\beta})}{6n} \langle p^2 \rangle_0 + \left(12\beta + 9 + \frac{6}{\beta}\right) \langle p^3 \rangle_0 + \left(-12 - \frac{8}{\beta}\right) \langle p^4 \rangle_0 + \frac{3}{\beta} \langle p^5 \rangle_0 \right]
$$

= $\beta$.

(4.37)

This spectacular order by order cancellation is in fact required by the sum rule, as was argued above, and so provides a consistency check of our approximations.

Higher orders in $n$ have useful information for correlators of distinct $f_{ij}$. However, for our purposes in this Subsection, for correlators at a single $\{i,j\}$ it suffices to use the leading term, given by the Poisson distribution. At this order

$$
\langle f_{ij}^n \rangle = \langle p^n \rangle_0.
$$

(4.38)
We can then find arbitrary correlators of $\tilde{f}$ at the same point. For example
\[
\langle \tilde{f}^2 \rangle = \langle \left( \frac{f}{\beta} - 1 \right)^2 \rangle = \frac{\langle f^2 \rangle - 2\langle f \rangle \langle 1 \rangle}{\beta^2} + 1 = \frac{\langle p^2 \rangle_0}{\beta^2} - \frac{2\langle p \rangle_0}{\beta} + 1 = \frac{1}{\beta}.
\] (4.39)

This is reasonable. It means that so long as $\beta >> 1$, $\tilde{f}$ will stay away from its minimal value of $-1$, where $f$ vanishes, and so is reasonably well approximated by a Gaussian. As $\alpha$ is quadratic in $\tilde{f}$, to determine its variance we will need four point functions of $\tilde{f}$. If all $\tilde{f}$ are at the same point, the leading order contribution is
\[
\langle \tilde{f}^4 \rangle = \langle \left( \frac{f}{\beta} - 1 \right)^4 \rangle = \frac{\langle f^4 \rangle - 4\langle f^3 \rangle \langle 1 \rangle}{\beta^4} + \frac{6\langle f^2 \rangle \langle 1 \rangle}{\beta^2} - \frac{4\langle f \rangle \langle 1 \rangle}{\beta} + 1
\]
\[
= \frac{3}{\beta^2} + \frac{1}{\beta^3} = 3\langle \tilde{f} \rangle^2 + \frac{1}{\beta^3}.
\] (4.40)

The first term is usual disconnected contribution to the four point function, in which the $\tilde{f}$s are paired into 3 possible pairs of pairs and their two point correlations are used. These give a result of order $1/n^2$ which, combined with the $n^4$ in $\alpha^2$ in Eq. (4.23) yields $\beta^2$ and so a variance of order $O(n^2)$. Therefore the consistency of the Gaussian approximation of $\rho(\alpha)$ demands that such terms cancel in the total expression for $\langle \alpha^2 \rangle - \langle \alpha \rangle^2$, leaving only those which, like the last term in (4.40), are of order $1/n^3$.

### 4.4 Bin Statistics from Partitions: Multiple Points

In general we will need correlators of $\tilde{f}_{ij}$ with different indices. There are two ways to generalize the above calculation to multiple indices. The first is the use of the sum rule to extrapolate new correlators from old correlators. This is sufficient to derive all of the Gaussian terms, as these simply come from the two point function, and the sum rule together with one two point function yields all two point functions. So the sum rule approach will be sufficient for our application in Sec. 5 which concerns the cancellation of the $O(n^2)$ terms. However, once that order is understood, the reader may wish to get to work calculating the matrix elements that motivate this project. These come from the essentially Poisson terms like the last term in Eq. (4.40) and most, but not all, of these can be derived from the previous case using sum rules.

Let us begin with the sum rule approach. Once we know that
\[
\langle \tilde{f}_{ij} \tilde{f}_{ij} \rangle = \frac{1}{\beta}
\] (4.41)
the sum rule implies that
\[
\langle \tilde{f}_{ij} \tilde{f}_{k\ell} \rangle = \langle \tilde{f}_{ij} \tilde{f}_{k\ell} \rangle = -\frac{1}{\beta^2}
\] (4.42)

26
for all $i \neq k$ and $j \neq l$ and therefore that

$$\langle \tilde{f}_{ij} \tilde{f}_{kl} \rangle = \frac{1}{\beta q^2} = \frac{1}{n} \quad (4.43)$$

for $i \neq k$ and $j \neq l$. We will denote these correlations using the following diagrams

$$\langle \tilde{f}_{ij} \rangle = \begin{pmatrix} \bullet \end{pmatrix}, \quad \langle \tilde{f}^2_{ij} \rangle = \begin{pmatrix} \bullet^2 \end{pmatrix} \quad (4.44)$$

$$\langle \tilde{f}_{ij} \tilde{f}_{il} \rangle = \begin{pmatrix} \bullet \rightarrow \bullet \end{pmatrix}, \quad \langle \tilde{f}_{ij} \tilde{f}_{kl} \rangle = \begin{pmatrix} \bullet \quad \bullet \end{pmatrix}. \quad (4.45)$$

Here the rows are the $\{j,l\}$ indices which are contracted with $\Phi$ in our master formula (4.23), while the columns are the $\{i,k\}$ indices which are ordered. Recall that $g \in S_n$ is represented as a map $P : [1,n] \rightarrow [1,n]$ and so the rows correspond to the bins in the image and the columns to the bins in the domain of the map. Inverting $g$ corresponds to a transpose of the diagram, but this does not affect the statistics as it is an automorphism of $S_n$ and so each diagram will be equal to its transpose. Similarly, rows can be freely interchanged, as can columns, without changing the value. Solid lines connect entries directly related by the sum rule, and so introduce factors of $-1/q$ whereas dashed lines connect entries which are connected by two sum rules. At this leading order in $n$ the dashed lines introduce factors of $1/q^2$.

The $O(n^2)$ approximation to the four point functions then follow from simply summing together the three pairs of products of two point functions. For example, if $i \neq k$ and $j \neq l$ then at leading order

$$\langle \tilde{f}^2_{ij} \tilde{f}^2_{kl} \rangle = \langle \tilde{f}^2_{ij} \rangle \langle \tilde{f}^2_{kl} \rangle = \frac{1}{\beta^2} \quad (4.45)$$

while

$$\langle \tilde{f}^2_{ij} \tilde{f}_{k1} \tilde{f}_{k2} \rangle = \langle \tilde{f}^2_{ij} \rangle \langle \tilde{f}_{k1} \tilde{f}_{k2} \rangle = \frac{1}{\beta} \left( -\frac{1}{\beta q} \right) = -\frac{1}{\beta^2 q} \quad (4.46)$$

corresponding to the diagrams

$$\langle \tilde{f}^2_{ij} \rangle \langle \tilde{f}^2_{kl} \rangle = \begin{pmatrix} \bullet^2 \end{pmatrix}, \quad \langle \tilde{f}_{ij} \tilde{f}_{k1} \tilde{f}_{k2} \rangle = \begin{pmatrix} \bullet \rightarrow \bullet \end{pmatrix}. \quad (4.47)$$

There are contributions from other combinations of pairings of the points, but these are subdominant in $q$. 

27
In general to calculate correlators at distinct points, the sum rules are not sufficient. However the above partition argument can be generalized. For concreteness, let use consider a correlator corresponding to a diagram with 2 rows and 2 columns. This means that we will be interested in two domain bins \( \{ i, k \} \) and two image bins \( \{ j, l \} \). We will need to calculate the joint probability distributions of

\[
  f_{ij}(g) = p_1, \quad f_{il}(g) = p_2, \quad f_{kj}(g) = p_3, \quad f_{kl}(g) = p_4. \tag{4.48}
\]

The joint probability \( \tilde{h} \) is just the number of elements \( g \) satisfying \( 4.48 \) divided by \( n! \).

It can be calculated as in the \( 1 \times 1 \) case treated above. First, one needs to choose \( p_1 \) elements of \( S_j \) to be in \( P(S_i) \). There are \( \binom{n/q}{p_1} \) such choices. Similarly there are \( \binom{n/q}{p_2} \) choices for the intersection of \( P(S_i) \) and \( S_l \). This leaves \( (n/q - p_1 - p_2) \) elements of \( S_i \) which must map into the complement of \( S_j \) and \( S_l \), which has \( (n - 2n/q) \) elements, yielding \( \binom{n-2n/q}{n/q-p_1-p_2} \) possibilities. Now we have counted the possible images of \( S_i \), we must do the same for \( S_l \). Recall that \( p_3 \) elements of \( S_k \) are mapped into \( S_j \). However, \( p_1 \) elements of \( S_j \) are already full, and so only \( n/q - p_1 \) slots are available. Thus the number of possible images of this map is \( \binom{n/q-p_1}{p_3} \). Similarly the choice of images of \( S_k \) in \( S_l \) yields a factor of \( \binom{n/q-p_2}{p_4} \). Now \( n/q - p_3 - p_4 \) elements rest in \( S_k \) which must be mapped into the remaining \( n - 4n/q + p_1 + p_2 \) elements in the complement of \( S_j \) and \( S_l \), yielding a factor of \( \binom{n-4n/q+p_1+p_2}{n/q-p_3-p_4} \). Finally, once one has chosen which slots are occupied, one multiplies by the various permutations of the domains, yielding \( (n/q)!^2(n - 2n/q)! \). As always, this last factor is independent of the \( p_i \) and can be absorbed into a normalization constant to be fixed later. Expanding these 6 choose functions into factorials and absorbing all terms independent of the \( p_i \) in the constant \( c \), one obtains

\[
  \tilde{h}_{ij}(p_i) = \frac{c}{p_1!p_2!p_3!p_4!} \left( \frac{n/q}{(n/q - p_1 - p_2)!} \right) \left( \frac{n/q}{(n/q - p_3 - p_4)!} \right) \left( \frac{n/q - p_1 - p_3}{(n/q - p_1 - p_3)!} \right) \left( \frac{n/q - p_2 - p_4}{(n/q - p_2 - p_4)!} \right) \left( \frac{n - 4n/q + p_1 + p_2 + p_3 + p_4}{(n - 4n/q + p_1 + p_2 + p_3 + p_4)!} \right). \tag{4.49}
\]

One sees that the terms with isolated \( p_i \)'s cancel, only those with entire rows \( \{ p_1 + p_3, p_2 + p_4 \} \) or columns \( \{ p_1 + p_2, p_3 + p_4 \} \) remain. This is true even for larger matrices.

The first four ratios enforce the correlations caused by the sum rules corresponding to each of the two rows and each of the two columns, while the last enforces the sum rule on the entire matrix. This may be expanded using our main tool \( 4.31 \) and any correlation function may then be calculated as a sum of the corresponding Poisson correlation functions as above.

In particular the \( 1/q \) terms in general always yield factors of \(-1/q\) associated with any two elements of the same row and column. However, since each ratio of factorials only appears
once, no diagram may contain two such lines in the same row and column. Triplets instead appear in the $1/q^2$ terms, and quadruplets at $1/q^3$. Similarly the last term in Eq. (4.49) yields dashed diagonal lines with factors of order $1/q^2$, although the coefficient is more complicated than at leading order. While at $O(n^2)$ we have seen that the diagrams reduce to pairs of two point functions, at $O(n)$ it appears that only connected diagrams contribute to the four point function. This may be expected since the $1/n^3$ term in Eq. (4.40), which is at the correct order, only appears when the irreducible correlation of four points is considered.

5 Testing the Anchor

What is the variance of $\alpha$?

5.1 The Variance of $\alpha_1$ at $O(n^3)$

Let us warm up with $\alpha_1$ as given in Eq. (4.12). There are two terms. First, a constant term, which doesn’t contribute. We will drop it. Next is

$$x = \frac{2n^2}{q^3} \sum_{i,j=1}^{q} i \tilde{f}_{ij} K \left( \frac{n}{q} \right). \quad (5.1)$$

The variance is then

$$\langle x^2 \rangle = \frac{4n^4}{q^6} \sum_{i,j=1}^{q} \sum_{k,l=1}^{q} i k \langle \tilde{f}_{ij} \tilde{f}_{kl} \rangle K \left( \frac{n}{q} \right) K \left( \frac{n}{q} \right). \quad (5.2)$$

This is the sum of four terms depending on whether $i = k$ and whether $j = l$, each summand corresponding to a diagram.

When $i \neq k$ and $j \neq l$ one uses

$$\begin{pmatrix} \bullet & \bullet & \bullet \end{pmatrix} = \langle \tilde{f}_{ij} \tilde{f}_{kl} \rangle = \frac{1}{n} \quad (5.3)$$

to obtain the contribution

$$x_1 = \frac{4n^4}{q^6} \left( \sum_{i=1}^{q} i \right) \left( \sum_{k=1}^{q} k \right) \frac{1}{n} \left( \sum_{j=1}^{q} K \left( \frac{n}{q} \right) \right) \left( \sum_{l=1}^{q} K \left( \frac{n}{q} \right) \right) \quad (5.4)$$

$$= \frac{4n^3}{q^6} \left( \frac{q^2}{2} \right)^2 (\pi q)^2 = \pi^2 n^3. \quad (5.5)$$

29
When \( i = k \) and \( j \neq l \), the matrix element

\[
\begin{pmatrix}
\bullet \\
\downarrow \\
\bullet
\end{pmatrix} = \langle \tilde{f}_{ij} \tilde{f}_{il} \rangle = -\frac{q}{n}
\]  

(5.6)

yields

\[
x_2 = \frac{4n^4}{q^6} \left( \sum_{i=1}^{q} i^2 \right) \left( -\frac{q}{n} \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q} j \right) \right) \left( \sum_{l=1}^{q} K \left( \frac{n}{q} l \right) \right)
\]

(5.7)

\[
= -\frac{4n^3}{q^5} \left( \frac{q^3}{3} \right) (\pi q)^2 = -\frac{4\pi^2}{3}n^3.
\]

(5.8)

Next one considers \( j = l \) but \( i \neq k \), with matrix element

\[
\begin{pmatrix}
\bullet \\
\rightarrow \\
\bullet
\end{pmatrix} = \langle \tilde{f}_{ij} \tilde{f}_{kj} \rangle = -\frac{q}{n}
\]

(5.9)

to find

\[
x_3 = \frac{4n^4}{q^6} \left( \sum_{i=1}^{q} i \right) \left( \sum_{k=1}^{q} k \right) \left( -\frac{q}{n} \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q} j \right) \right)^2
\]

(5.10)

\[
= -\frac{4n^3}{q^5} \left( \frac{q^3}{2} \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q} j \right) \right)^2 = -n^3 \langle K^2 \rangle
\]

(5.11)

where we have defined the average

\[
\langle K^2 \rangle = \frac{1}{q} \sum_{j=1}^{q} K \left( \frac{n}{q} j \right)^2.
\]

(5.12)

Finally the case \( i = k, j = l \)

\[
\begin{pmatrix}
\bullet \\
\bullet
\end{pmatrix} = \langle \tilde{f}_{ij}^2 \rangle = \frac{q^2}{n}
\]

(5.13)

provides the last contribution

\[
x_4 = \frac{4n^4}{q^6} \left( \sum_{i=1}^{q} i^2 \right) \left( \frac{q^2}{n} \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q} j \right) \right)^2
\]

(5.14)

\[
= \frac{4n^3}{q^5} \left( \frac{q^3}{3} \right) \left( \sum_{j=1}^{q} K \left( \frac{n}{q} j \right) \right)^2 = \frac{4}{3}n^3 \langle K^2 \rangle.
\]

30
Summing these contributions one finds the variance of $\alpha_1$

$$\langle x^2 \rangle = \sum_{i=1}^{4} x_i = \left( \frac{(K^2) - \pi^2}{3} \right) n^3. \tag{5.15}$$

Recall that the average value of $K$ is $\pi$, and $K$ is not constant, so $\langle K^2 \rangle > \pi^2$ and therefore the $O(n^3)$ contribution does not vanish.

What about the unanchored $\alpha = \alpha_1 + \alpha_2$? Recall that the $x$ term in $\alpha_1$ is canceled by a term in $\alpha_2$, and so could the $O(n^3)$ contribution to the unanchored $\alpha_1 + \alpha_2$ vanish? The $\Phi$ term enters at $O(n^2)$ so it may seem promising. The trouble is the first term in (4.18). It is identical to the $\alpha_1$ term considered here except with $K(nj/q)$ replaced by $-2\pi j/q$. The $j \neq l$ cases then give $-\pi^2$, as $\pi$ is the average value of $2\pi j/q$. The $j = l$ cases give $4\pi^2/3$, as the average of $j^2/q^2$ is $1/3$. These again appear in the numerator and do not cancel. Thus, just the same calculation as above shows that the $O(n^3)$ terms in the variance do not cancel without the anchor.

### 5.2 The Variance of $\alpha$ at $O(n^2)$

Once the anchor is included, one arrives at our master formula for $\alpha$ in Eq. (4.23). Here all terms that could potentially give $n^3$ contributions on dimensional grounds are gone. The constant term does not contribute to the variance and so we will drop it. We will also shift $\alpha_3$ to make it antisymmetric in $j$ and $l$, thus eliminating the zero point which created a nonzero expectation value for $\langle \alpha_3 \rangle$. We do not know if such a shift is necessary for the binning postulate. However it does not affect the previous arguments concerning the role of the anchor. The variance in $\alpha$ will therefore be equal to that of

$$x = \frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \tilde{f}_{kl} \Phi \left( \frac{n}{q} j, \frac{n}{q} l \right) + 2\pi \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \tilde{f}_{kl} \left( \frac{1}{2} - \theta(j - l) \right)$$

$$= \frac{1}{2} \beta^2 \sum_{i<k} \sum_{j,l=1}^{q} \tilde{f}_{ij} \tilde{f}_{kl} \left( \Phi \left( \frac{n}{q} j, \frac{n}{q} l \right) + 2\pi - 4\pi \theta(j - l) \right). \tag{5.16}$$

The variance is just

$$\langle x^2 \rangle = \frac{\beta^4}{4} \sum_{i_1<k_1} \sum_{i_2<k_2} \sum_{j_1,l_1,j_2,l_2=1} \langle \tilde{f}_{i_1j_1} \tilde{f}_{i_2j_2} \tilde{f}_{k_1l_1} \tilde{f}_{k_2l_2} \rangle \tag{5.17}$$

$$\times \left( \Phi \left( \frac{n}{q} j_1, \frac{n}{q} l_1 \right) + 2\pi - 4\pi \theta(j_1 - l_1) \right) \left( \Phi \left( \frac{n}{q} j_2, \frac{n}{q} l_2 \right) + 2\pi - 4\pi \theta(j_2 - l_2) \right).$$
We are interested in the $O(n^2)$ contribution, which arises entirely from the Gaussian correlations, which are disconnected pairs of 2 point functions. When more than one pairing is available, the sum over pairings may increase the diagram by a factor of 2 or 3 however this requires fixing one of the indices, which costs a factor of $q$ and so diagrams with equal contributions from multiple pairings will always be subleading in $1/q$. Thus we need only consider diagrams with only a single choice of dominant pairing. In addition, diagrams with more than three rows or columns will lead to vanishing contributions, as the sums of both indices of $\Phi$ vanish and also the sums of the anchor terms vanish due to the zero point shift corresponding to the $2\pi$ in (5.16). Thus in all we will only need to sum five diagrams.

We begin with easiest, corresponding to

$$\begin{pmatrix}
  \bullet^2 \\
  \bullet^2
\end{pmatrix} = \langle \bar{f}^2_{ij} \rangle \langle \bar{f}^2_{kl} \rangle = \frac{1}{\beta^2}. \tag{5.18}$$

There are only two distinct values of $i_1$, $i_2$, $k_1$ and $k_2$. As $i_1 < k_1$ and $i_2 < k_2$, this implies that $i_1 = i_2 = i$ and $k_1 = k_2 = k$. Since both points are degenerate, this means that also $j_1 = j_2 = j$ and $l_1 = l_2 = l$. Now in this case and in all cases that follow, the matrix element is entirely determined by the diagram and the $\Phi$ factors have no $i$ or $k$ dependence, thus the sums over the $i$ and $k$ can be factored out and evaluated separately. Thus this contribution is

$$x_1 = \frac{\beta^4}{4} \left( \sum_{i<k} \frac{1}{\beta^2} \sum_{j=1}^q \left( \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right) + 2\pi - 4\pi \theta(j-l) \right)^2 \right).$$

$$= \frac{n^2}{4q^4} \frac{q^2}{2} \sum_{j,l=1}^q \left( \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right)^2 - 8\pi \theta(j-l) \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right) + 4\pi^2 - 16\pi \theta(j-l) + 16\pi \theta(j-l) \right)$$

$$= n^2 \left( \frac{\langle \Phi^2 \rangle}{8} - \pi \langle \Phi \rangle + \frac{\pi^2}{2} \right). \tag{5.19}$$

where we have defined

$$\langle \Phi^2 \rangle = \frac{1}{q^2} \sum_{j,l=1}^q \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right)^2, \quad \langle \Phi \rangle = \frac{1}{q^2} \sum_{j>l}^q \Phi \left( \frac{n_j}{q}, \frac{n_l}{q} \right). \tag{5.20}$$
We next consider the diagram

\[
\begin{pmatrix}
  & \bullet \\
  \downarrow & \bullet \\
  \bullet & \uparrow
\end{pmatrix} = \langle \tilde{f}_{ij} \rangle \langle \tilde{f}_{k\ell_1} \tilde{f}_{k\ell_2} \rangle = -\frac{1}{q\beta^2}.
\]

\hspace{1cm} (5.21)

Here again there are only two values of \( i_1, i_2, k_1 \) and \( k_2 \) and so again \( i_1 = i_2 = i \) and \( k_1 = k_2 = k \). One of these is a double point. If it is \( i \) then \( j_1 = j_2 \), but if it is \( k \) then \( l_1 = l_2 \). These two cases give equal contributions, and so we consider the first and multiply by a factor of two. Altogether

\[
x_2 = 2 \frac{\beta^4}{4} \left( \sum_{i<k}^{q} \left( -\frac{1}{q\beta^2} \right) \right) \times \sum_{j,l_1,l_2=1}^{q} \left( \Phi \left( \frac{n}{q}, \frac{n}{q}, j, l_1 \right) + 2\pi - 4\pi\theta(j - l_1) \right) \left( \Phi \left( \frac{n}{q}, \frac{n}{q}, j, l_2 \right) + 2\pi - 4\pi\theta(j - l_2) \right).
\]

The first line gives \(-n^2/(4q^3)\). To simplify the second line, we can use the binned version of the Bethe equation (4.15) to sum over \( l_1 \) and \( l_2 \), leaving

\[
x_2 = -\frac{n^2}{4q^3} \sum_{j=1}^{q} \left( 2qK \left( \frac{n}{q}, j \right) + 4\pi(j - q) + 2\pi q - 4\pi j \right)^2
\]

\[
= n^2 \left( -\langle K^2 \rangle + \pi^2 \right)
\]

\hspace{1cm} (5.23)

where we have again used the fact that the average value of \( K \) is \( \pi \).

The third diagram is

\[
\begin{pmatrix}
  \bullet & \longrightarrow & \bullet \\
  \bullet & \bullet & \bullet
\end{pmatrix} = \langle \tilde{f}_{ij} \rangle \langle \tilde{f}_{k\ell_1} \tilde{f}_{k\ell_2} \rangle = -\frac{1}{q\beta^2}.
\]

\hspace{1cm} (5.24)

Now there are three columns, and so there are inequivalent pairings of \( i \) and \( k \). One may have \( i_1 = i_2 = i, k_1 = k_2 = k, i_1 = k_2 \) or \( i_2 = k_1 \). The first two give equal contributions, as there is a symmetry in which \( i \) and \( k \) are exchanged and the sites are inverted. Similarly the third and fourth are equal. More subtly, the third is equal to minus one half of the first.
This is because in the first case the $i$ and $k$ sum is

\[
\sum_{i=1}^{q} \sum_{k_{1},k_{1}=i}^{q} 1 = \sum_{i=1}^{q} (q-i)^2 = \frac{q^3}{3}.
\]  

(5.25)

While in the second it is

\[
\sum_{i_{1}=1}^{q} \sum_{i_{2}=1}^{i_{1}-1} \sum_{k_{1}=i_{1}+1}^{q} 1 = \sum_{i_{1}=1}^{q} i_{1}(q-i_{1}) = \frac{q^3}{6}.
\]  

(5.26)

This explains the factor of two difference. The signs are different because in the second case one exchanges one pair of $(j,l)$. Both $\Phi$ and also the zeroed form of $\alpha_{3}$ are antisymmetric with respect to this interchange.

Summarizing, we only need to consider the first of the four possibilities, and the contribution of the other diagrams will give a weight factor of $1 + 1 - 1/2 - 1/2 = 1$. This is

\[
x_{3} = -\frac{\beta^{4}}{4} \left( \sum_{i < k_{1},k_{2}}^{q} \frac{1}{q\beta^{2}} \sum_{j,l=1}^{q} \left( \Phi \left( \frac{n_{j}}{q}, \frac{n_{l}}{q} \right) + 2\pi - 4\pi \theta(j-l) \right) \right)^{2}.
\]  

(5.27)

Note that this is equal to our first expression for $x_{1}$ in Eq. (5.19) except for the $\{i, k\}$ integral which is multiplied by a factor of $2q/3$ and the matrix element which is multiplied by $-1/q$. Therefore

\[
x_{3} = -\frac{2x_{1}}{3}.
\]  

(5.28)

The next diagram is three by three

\[
\left( \begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{array} \right)
\]  

\[
= \langle \hat{f}_{1j}^{2} \rangle \langle \hat{f}_{k1,1} \hat{f}_{k2,l} \rangle = \frac{1}{q^{2}\beta^{2}}.
\]  

(5.29)

Again, corresponding to the three columns, there are three possible values of the $i$ and $k$, yielding the same four pairings as above. The integration factors are the same and so again the weights are $1, 1, -1/2$ and $-1/2$ and so it will suffice to consider the possibility $i_{1} = i_{2} = i$. As $i$ is at a double point, $j_{1} = j_{2}$. However, unlike the previous case, now there
are three rows and so \( l_1 \neq l_2 \). This leaves us with

\[
x_4 = \frac{\beta^4}{4} \left( \sum_{i<k_1,k_2}^q \frac{1}{q^2\beta^2} \right) \times \sum_{j,l_1,l_2=1}^q \left( \Phi \left( \frac{n}{q}j, \frac{n}{q}l_1 \right) + 2\pi - 4\pi \theta(j - l_1) \right) \left( \Phi \left( \frac{n}{q}j, \frac{n}{q}l_2 \right) + 2\pi - 4\pi \theta(j - l_2) \right).
\]

(5.30)

The first line yields \( n^2/(12q^3) \). As in the case of \( x_2 \), the \( l_1 \) and \( l_2 \) may be summed in the last line using the binned Bethe equation, leaving

\[
x_4 = \frac{n^2}{12q^3} \sum_{j=1}^q \left( 2qK \left( \frac{n}{q}j \right) + 4\pi(j - q) + 2\pi q - 4\pi j \right)^2.
\]

(5.31)

Comparing with Eq. (5.23) we see that

\[
x_4 = -\frac{x_2}{3}.
\]

(5.32)

The final diagram is

\[
\begin{pmatrix}
\bullet \\
\downarrow \\
\bullet
\end{pmatrix}
\begin{pmatrix}
\bullet \\
\bullet
\end{pmatrix} = \langle \tilde{r}_{ij}\tilde{r}_{ij} \rangle \langle \tilde{f}_{k1}\tilde{f}_{k1} \rangle = \frac{1}{q^2\beta^2}.
\]

(5.33)

Again there are three columns and so the same three values of \( i \) and \( k \), with the same weights and so we need only consider the first case \( i_1 = i_2 = i \). Unlike the case of \( x_4 \), now \( i_1 = i_2 \) implies that \( l_1 = l_2 = l \). Thus we find

\[
x_5 = \frac{\beta^4}{4} \left( \sum_{i<k_1,k_2}^q \frac{1}{q^2\beta^2} \right) \times \sum_{j_1,j_2,l=1}^q \left( \Phi \left( \frac{n}{q}j_1, \frac{n}{q}l \right) + 2\pi - 4\pi \theta(j_1 - l) \right) \left( \Phi \left( \frac{n}{q}j_2, \frac{n}{q}l \right) + 2\pi - 4\pi \theta(j_2 - l) \right).
\]

(5.34)

The first row is identical to that of \( x_4 \) in Eq. (5.31). What about the second row? If one exchanges \( j \) with \( l \) then the \( \Phi \) terms look the same, but with their indices reversed.
Transposing the indices gives a minus sign in each summand. However $2\pi - 4\pi \theta(j - l)$ is also antisymmetric under the exchange of $j$ and $l$, therefore both factors in the second line change sign, leaving the second line invariant as well. Thus we have found

$$x_5 = x_4 = -\frac{x_2}{3}. \quad (5.37)$$

Adding all of these terms together we find that the variance of the anchored $\alpha$, at $O(n^2)$, is

$$x = \sum_{i=1}^{5} x_i = \frac{x_1 + x_2}{3} = n^2 \left( \frac{\langle \Phi^2 \rangle}{24} - \frac{\pi \langle \Phi \rangle}{3} - \frac{\langle K^2 \rangle}{3} + \frac{\pi^2}{2} \right). \quad (5.38)$$

Is this zero? Numerically we have found

$$\langle \Phi^2 \rangle = 5.44, \quad \langle \Phi \rangle \sim -1.13, \quad \langle K^2 \rangle = 11.48 \quad (5.39)$$

and so

$$x \sim (0.23 - 1.18 - 3.83 + 4.93)n^2 = 0.15n^2. \quad (5.40)$$

Is this compatible with zero? It is nearly twice the best fit Gaussian variance found at $n = 11$ in Fig. 1, but this is not obviously a sign of incompatibility as the $O(n)$ term could easily drive it down, with a coefficient of order unity.

6 Conclusions

Our goal is to devise a method to calculate, to arbitrary accuracy, the ground state and first excited state wave functionals of the $\mathbb{CP}^1$ nonlinear sigma model. We would like to use these states to study the behavior of the wavefunctional in the potential well corresponding to a chain which circumnavigates the target space, to learn how the two sides of the equator are connected for the various states. We feel, by analogy with the double well potential in quantum mechanics, that this will teach us how instantons generate the mass gap, and it will shed light on the role of instantons in Yang-Mills theory.

This model is equivalent to a high spin Heisenberg XXX spin chain, for which the states are in principle known, but in a rather unwieldy form which would be difficult to map to the sigma model. Our goal is to find a prescription to calculate the Heisenberg chain matrix elements which is sufficiently simple so that it can be mapped to the sigma mode. We begin, for sanity’s sake, with spin 1/2. To cast our problem in a way which is close to continuum field theory, we collected the lattice sites into bins. We believe that it is the bins, and not individual pairs of sites, which will eventually correspond to points in the continuum field theory. In the binning approximation, the richness of this system is smoothed away. We
found that standard combinatorial arguments in terms of partitions describe the behavior of the bins. Thus instead of complexities which are polynomial in \( N \), the chain length \( N \) essentially disappears from the problem.

Our strategy is to encode the information about a matrix element in a single function, \( \rho(\alpha) \), which is the density of phases \( \alpha \) in the CBA. The Fourier transform of \( \rho(\alpha) \) gives a matrix element.

Our initial hope was that \( \rho(\alpha) \) would be a Gaussian, and so this would be straightforward. However it turned it that the variance was of order \( O(N^{3/2}) \). In the Gaussian approximation this would lead to matrix elements of order \( e^{-N^{3}} \), which is inconsistent with the fact that there are only \( 2^N \) states. However we found that this Gaussian approximation is quite poor because \( \rho(\alpha) \) is rich in substructure which in fact dominates the Fourier transform. To fix this, we modified \( \alpha \) by introducing an anchor which leaves the matrix elements invariant. The modified \( \alpha \) appears to be free of substructure, at least in the cases investigated. Thus our proposal is to calculate the moments of \( \alpha \) using the combinatorial methods described and use these to reconstruct \( \rho(\alpha) \), whose Fourier transform gives the matrix elements.

What about the Gaussian approximation? If indeed \( \rho(\alpha) \) is a Gaussian, then matrix elements of \( O(e^{-N}) \) are only obtained if the variance of our anchored \( \alpha \) is \( O(N) \). The anchor eliminates the \( O(N^3) \) part and we have calculated here the \( O(N^2) \) contribution. We found that the \( O(N^2) \) coefficient is quite small and in the last step our approach was numerical. However it appears to be inconsistent with zero. If indeed it is nonzero, then what has gone wrong? Is our method doomed?

If the variance contains a term of \( O(N^2) \), then that term will dominate the variance at large \( N \), which is the limit of interest. But the question is whether it will dominate the matrix elements. If it does, we arrive at an inconsistency and so it means that we have made a mistake, for example perhaps our binning approximation is invalid. Whether it dominates depends on the distribution.

Consider the following three distributions \( \rho(\alpha) \). The first is a Gaussian with variance that scales as \( O(N^2) \) at large \( N \). The second is the weighted sum of two Gaussians with \( N \)-independent weights, one with a variance of \( O(N) \) and the other with a variance of \( O(N^2) \). The third, which generalizes the second, is of the form

\[
\rho(\alpha) \sim \exp \left[ -f(\alpha, N) \right], \quad f(\alpha, N) \sim \begin{cases} 
\frac{\alpha^2}{N} & \text{if } \alpha^2 << N \\
\frac{\alpha^2}{N^2} & \text{if } \alpha^2 >> N^2
\end{cases} \quad (6.1)
\]

In the first case, the matrix elements will be \( O(e^{-N^2}) \) and so we will have an inconsistency. In the second, at large \( N \) the broader Gaussian simply ceases to contribute the matrix elements.
elements, and so the matrix elements are of $O(e^{-N})$ as desired, determined entirely from the thin Gaussian. In the third case, for the first few standard deviations the probability falls rapidly as the distribution seems to be a thin Gaussian. So long as the cross over to the $O(N^2)$ is at sufficiently high $\alpha^2$ that the area of the thin region does not tend to zero at large $N$, then the matrix elements will again be determined by the thin region and so have the correct behavior.

So which is the case at hand? The calculation of higher moments can distinguish these cases, although at any finite moment, assumptions about the form of $\rho(\alpha)$ will be necessary to determine the potential. In other words, a concrete statement of the absence of substructure is needed. Fig. 1 shows that, at least at finite $N$, $\rho(\alpha)$ is leptokurtic. If this persists at infinite $N$, it would be inconsistent with the first case but consistent with the others.

In fact, it is possible for us to go beyond simply calculating moments. Eq. (4.49) is the entire joint probability density function (PDF) for $f_{ij}, f_{il}, f_{jk}$ and $f_{kl}$. It can be put in a useful form with the expansion (4.31) and $c$ can be found by imposing that $\langle 1 \rangle = 1$ as was done in Subsec. 4.3. Summing over $p_2$ and $p_3$ one is left with the joint PDF for $p_1$ and $p_4$ which are $f_{ij}$ and $f_{kl}$. This is easily converted into a joint PDF for $\tilde{f}_{ij}$ and $\tilde{f}_{kl}$, which via Eq. (4.23) yields $\rho(\alpha)$, which is the PDF for $\alpha$. If this can be calculated directly, at some order in $N$, the answer may be inserted into Eq. (1.1) to determine the matrix element. In this way, no assumptions regarding substructure are needed.

In most of this note we considered matrix elements with respect to the classical vacuum. However we believe that similar calculations may be performed in a more general case in which each bin consists of repetitions of the same pair, corresponding to a fixed point in the target space. The dependence of the matrix elements on this bin by bin choice is, in our opinion, the content of the vacuum state in the continuum theory. Therefore we will need to understand it, and understand how to calculate it.

And if this all works, how do we get to higher spin? After all, there is no CBA in these cases? The algebraic Bethe Ansatz provides a much more complicated construction of these states. However on the bright side they are still constructed from $N$ commuting copies of the creation operators $B(\lambda_i)$, and so there is still a permutation symmetry on the $\lambda_i$. This lends hope that it may be possible to write a state in some basic form, analogous to a single summand in CBA, which upon symmetrization gives the true state. Then the technology from $s = 1/2$ to handle binnings of permutations could be imported to this more complicated setting.
Acknowledgement

JE is supported by the CAS Key Research Program of Frontier Sciences grant QYZDY-SSW-SLH006. JE also thanks the Recruitment Program of High-end Foreign Experts for support.

References

[1] V. A. Fateev, I. V. Frolov and A. S. Shvarts, “Quantum Fluctuations of Instantons in the Nonlinear Sigma Model,” Nucl. Phys. B 154 (1979) 1. doi:10.1016/0550-3213(79)90367-5

[2] A. B. Zamolodchikov and A. B. Zamolodchikov, “Factorized s Matrices in Two-Dimensions as the Exact Solutions of Certain Relativistic Quantum Field Models,” Annals Phys. 120 (1979) 253. doi:10.1016/0003-4916(79)90391-9

[3] F. D. M. Haldane, “Continuum dynamics of the 1-D Heisenberg antiferromagnetic identification with the O(3) nonlinear sigma model,” Phys. Lett. A 93 (1983) 464. doi:10.1016/0375-9601(83)90631-X

[4] F. D. M. Haldane, “Nonlinear field theory of large spin Heisenberg antiferromagnets. Semiclassically quantized solitons of the one-dimensional easy Axis Neel state,” Phys. Rev. Lett. 50 (1983) 1153. doi:10.1103/PhysRevLett.50.1153

[5] I. Affleck, “The Quantum Hall Effect, σ Models at θ = π and Quantum Spin Chains,” Nucl. Phys. B 257 (1985) 397. doi:10.1016/0550-3213(85)90353-0

[6] L. D. Faddeev, “How algebraic Bethe ansatz works for integrable model,” [hep-th/9605187]

[7] H. Bethe, “On the theory of metals. 1. Eigenvalues and eigenfunctions for the linear atomic chain,” Z. Phys. 71 (1931) 205. doi:10.1007/BF01341708

[8] L. D. Faddeev and L. A. Takhtajan, “Spectrum and scattering of excitations in the one-dimensional isotropic Heisenberg model,” J. Sov. Math. 24 (1984) 241 [Zap. Nauchn. Semin. 109 (1981) 134]. doi:10.1007/BF01087245

[9] F. C. Alcaraz and M. J. Lazo, “The Bethe ansatz as a matrix product ansatz,” J. Phys. A 37 (2004) L1 doi:10.1088/0305-4470/37/1/L01 [cond-mat/0304170].

39
[10] H. Boos, M. Jimbo, T. Miwa, F. Smirnov and Y. Takeyama, “Density matrix of a finite sub-chain of the Heisenberg anti-ferromagnet,” Lett. Math. Phys. 75 (2006) 201 doi:10.1007/s11005-006-0054-x [hep-th/0506171].

[11] H. Katsura and I. Maruyama, “Derivation of Matrix Product Ansatz for the Heisenberg Chain from Algebraic Bethe Ansatz,” J. Phys. A 43 (2010) 175003 doi:10.1088/1751-8113/43/17/175003 [arXiv:0911.4215 [cond-mat.stat-mech]].

[12] N. Gromov, F. Levkovich-Maslyuk and G. Sizov, “New Construction of Eigenstates and Separation of Variables for SU(N) Quantum Spin Chains,” JHEP 1709 (2017) 111 doi:10.1007/JHEP09(2017)111 [arXiv:1610.08032 [hep-th]].

[13] J. Von Neumann, “On Infinite Direct Products,” Compositio Mathematica, Tome 6 (1939) 1.