

# Connectivity patterns in loop percolation I: the rationality phenomenon and constant term identities

Dan Romik

September 20, 2013

## Abstract

Loop percolation, also known as the dense  $O(1)$  loop model, is a variant of critical bond percolation in the square lattice  $\mathbb{Z}^2$  whose graph structure consists of a disjoint union of cycles. We study its connectivity pattern, which is a random noncrossing matching associated with a loop percolation configuration. These connectivity patterns exhibit a striking rationality property whereby probabilities of naturally-occurring events are dyadic rational numbers or rational functions of a size parameter  $n$ , but the reasons for this are not completely understood. We prove the rationality phenomenon in a few cases and prove an explicit formula expressing the probabilities in the “cylindrical geometry” as coefficients in certain multivariate polynomials. This reduces the rationality problem in the general case to that of proving a family of conjectural constant term identities generalizing an identity due to Di Francesco and Zinn-Justin. Our results make use of, and extend, algebraic techniques related to the quantum Knizhnik-Zamolodchikov equation.

Key words: loop percolation,  $O(1)$  loop model, XXZ spin chain, noncrossing matching, connectivity pattern, constant term identity, quantum Knizhnik-Zamolodchikov equation, wheel polynomials

2010 Mathematics Subject Classification: 60K35, 82B20, 82B23.

# 1 Introduction

## 1.1 Critical bond percolation and loop percolation

Critical bond percolation on  $\mathbb{Z}^2$  is the most natural model of a random subgraph of the square lattice: each edge of the lattice is included with probability  $1/2$ , independently of all other edges. Despite the simplicity of its definition, it is difficult to analyze rigorously. Though the spectacular recent results of Smirnov [29] and Lawler-Schramm-Werner [19] related to critical exponents, conformal invariance and the Schramm-Loewner Evolution (SLE) have resolved many of the outstanding open problems for the related model of critical site percolation on the triangular lattice, the same problems are still unsolved in the case of bond percolation on  $\mathbb{Z}^2$ .

In this paper we will study another percolation model, which is related to critical bond percolation on  $\mathbb{Z}^2$  and can be thought of as a variant of it. We will refer to the model as **loop percolation on  $\mathbb{Z}^2$** ; an equivalent model has been studied in the statistical physics literature under different names such as the **dense  $O(1)$  loop model** [21] or **completely packed loops** [36]. Another closely related model is the special case of the **XXZ spin chain** in which the so-called anisotropy parameter  $\Delta$  takes the value  $-1/2$ ; see [3, 6].

Our motivation in studying this model was twofold: first, to point out, and attempt to systematically exploit, the inherent interest and approachability of the model from the point of view of probability theory—qualities which have not been emphasized in previous studies. Second, to use and enhance some of the deep algebraic tools (referred to under the broad heading of the **quantum Knizhnik-Zamolodchikov equation**) that were developed in recent years to compute explicit formulas for several quantities of interest in the model. In this paper we focus on algebraic techniques. The follow-up paper [27] will contain additional results of a more probabilistic flavor.

Let us start with the definition of the model. We define loop percolation in the entire plane, but later we will also consider the model on a half-plane and on a semi-infinite cylinder.

**Definition 1.1** (Loop percolation). Consider the square lattice  $\mathbb{Z}^2$  equipped with a checkerboard coloring of the squares of the dual lattice, such that the square whose bottom-left corner is  $(0,0)$  is white. **Loop percolation** is the random subgraph  $LP = (\mathbb{Z}^2, V(LP))$  of  $\mathbb{Z}^2$  obtained by tossing a fair coin for each white square  $\square$  in the dual lattice, independently of all other coins,

and including in  $V(\text{LP})$  either the bottom and top edges of  $\square$  or the left and right edges of  $\square$  according to the result of the coin toss (Fig. 1).



Figure 1: The two possible edge configurations around a white lattice square.

Fig. 2 shows portions of a critical bond percolation configuration and an (unrelated) loop percolation configuration. In fact, the two models are equivalent, in the following sense. The black squares form the vertices of a graph where two black squares are adjacent if their centers differ by one of the vectors  $(0, \pm 2), (\pm 2, 0)$ . Identifying each black square with its center, this graph decomposes into disjoint components of “blue” and “red” sites according to parity, each component being isomorphic to  $\mathbb{Z}^2$ . Define a subgraph  $G$  of the blue component whose edges are precisely those that do not cross an edge of the loop percolation graph. Then it is easy to see that  $G$  is a critical bond percolation graph; see Fig. 3. Conversely, it is clear that starting from the critical bond percolation graph one can reconstruct the associated loop percolation graph.

As an immediate consequence of the above discussion, we see that the loop percolation graph almost surely decomposes into a disjoint union of cycles, or “loops.” Indeed, since each site of  $\mathbb{Z}^2$  is incident to precisely two white squares, it has degree 2, so the connected components of the graph are either loops or infinite paths; however, the presence of an infinite path would imply the existence of an infinite connected component in the “blue” critical bond percolation graph containing those blue sites adjacent to an edge in the infinite path, in contradiction to the well-known fact [14, Lemma 11.12] that critical bond percolation on  $\mathbb{Z}^2$  almost surely has no such infinite component.

Given the equivalence between loop percolation and critical bond percolation described above, one may ask why it is necessary to study loop percolation separately from critical bond percolation. We believe that there is much to be gained in doing so. In particular, the existing research on loop percolation and connections with other natural statistical physics models such as the XXZ spin chain and Fully Packed Loops indicate that studying this type of percolation as a model in its own right suggests a variety of

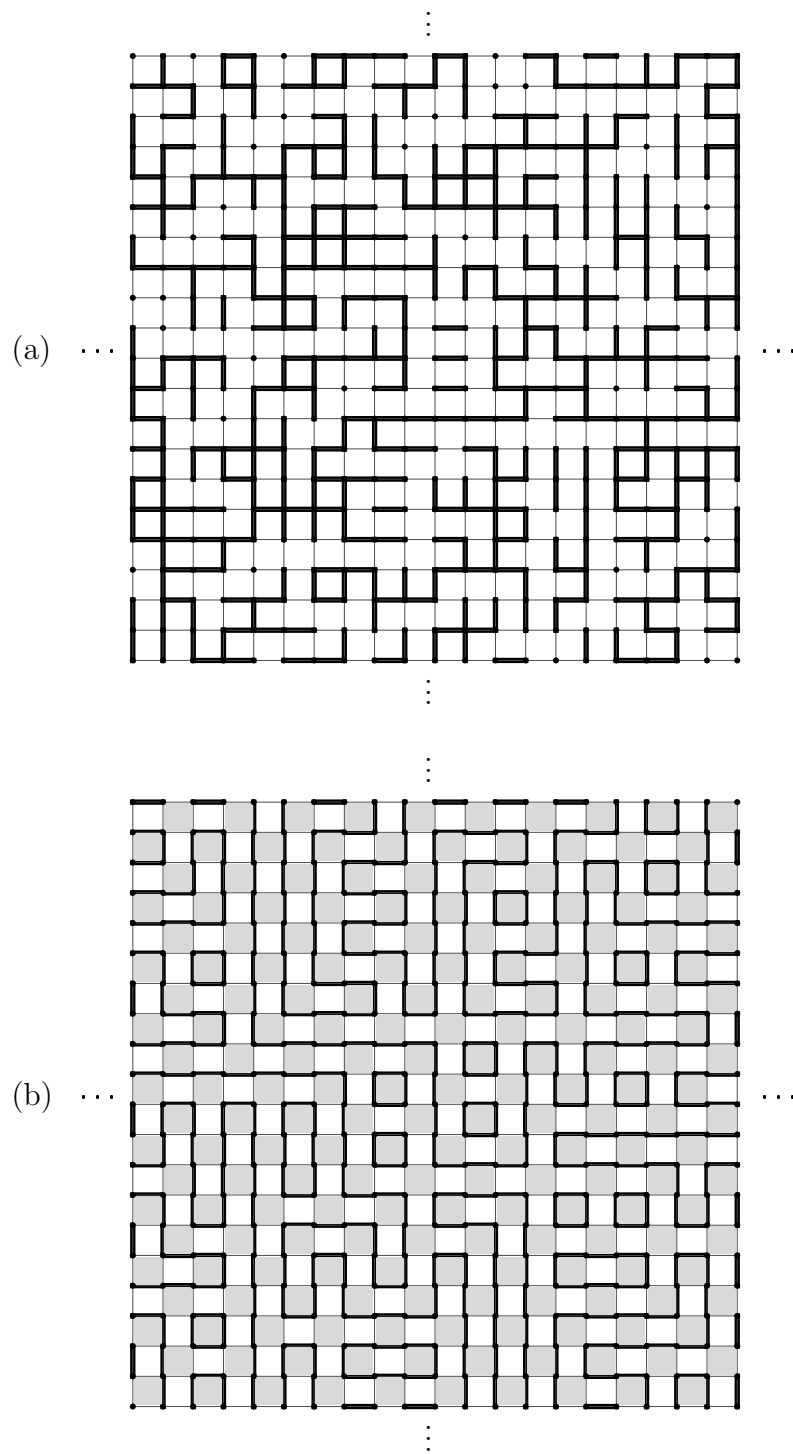


Figure 2: (a) Critical bond percolation. (b) Loop percolation.

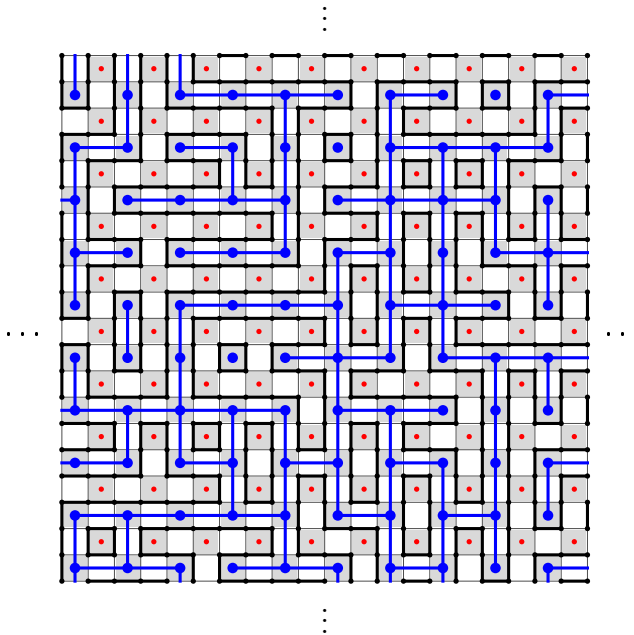


Figure 3: The critical bond percolation graph associated with a loop percolation graph.

mathematically interesting questions that one might not otherwise be led to consider by thinking directly of critical bond percolation.

Furthermore, the techniques developed for studying loop percolation have shown that it belongs to the (loosely defined) class of so-called “exactly solvable” models, for which certain remarkable algebraic properties can be used to get precise formulas for various quantities associated with the model. The new results presented in this paper will provide another strong illustration of this phenomenon and give further credence to the notion that loop percolation is quite worthy of independent study. Ultimately, we hope that this line of investigation may lead to new insights that could be used to attack some of the important open problems concerning critical bond percolation.

## 1.2 Noncrossing matchings

Our discussion will focus on certain combinatorial objects known as **non-crossing matchings** that are associated with loop percolation configura-

tions. Let us recall the relevant definitions. For two integers  $a < b$ , let  $[a, b]$  denote the discrete interval  $\{a, a + 1, \dots, b\}$ . Recall that a **noncrossing matching of order  $n$**  is a perfect matching of the numbers  $1, \dots, 2n$  (which we encode formally as a function  $\pi : [1, 2n] \rightarrow [1, 2n]$  such that  $\pi \circ \pi = \text{id}$  and  $\pi(k) \neq k$  for all  $k \in [1, 2n]$ ;  $\pi(k)$  represents the number matched to  $k$ ), which has the additional property that there do not exist numbers  $a < b < c < d$  in  $[1, 2n]$  such that  $\pi(a) = c, \pi(b) = d$ . If  $\pi(j) = k$  we say that  $j$  **and  $k$  are matched under  $\pi$**  and denote  $j \xleftrightarrow{\pi} k$ . Denote by  $\text{NC}_n$  the set of noncrossing matchings of order  $n$ . It is well-known that the number  $|\text{NC}_n|$  of elements in  $\text{NC}_n$  is  $\text{Cat}(n) = \frac{1}{n+1} \binom{2n}{n}$ , the  $n$ th Catalan number.

Similarly, an **infinite noncrossing matching** is a one-to-one and onto function  $\pi : \mathbb{Z} \rightarrow \mathbb{Z}$  that satisfies  $\pi \circ \pi = \text{id}$  and  $\pi(k) \neq k$  for all  $k \in \mathbb{Z}$ , such that there do not exist integers  $a < b < c < d$  for which  $\pi$  matches  $a$  to  $c$  and  $b$  to  $d$ . Denote by  $\text{NC}_{\mathbb{Z}}$  the set of noncrossing matchings on  $\mathbb{Z}$ .

Both finite and infinite noncrossing matchings can be represented graphically in a diagram in which noncrossing edges, or **arcs**, are drawn between the elements of any matched pair  $j \xleftrightarrow{\pi} k$ ; see Fig. 4. As illustrated in the figure, in the case of a finite noncrossing matchings there are two equivalent representations, as a matching of  $2n$  points arranged on a line or on a circle.

### 1.3 The connectivity pattern of loop percolation: the infinite case

We now associate a random noncrossing matching called the **connectivity pattern** with loop percolation configurations. There will be two variants of the problem according to the type of region being considered, resulting in infinite or finite noncrossing matchings. We start the discussion with the infinite case. To the best of our knowledge, this variant of the problem has not been previously considered in the literature, but nonetheless the rationality phenomenon we will discuss appears most striking in this setting.

**Definition 1.2** (Half-planar loop percolation). Denote by  $\mathbb{Z}_{\text{NE}}^2$  the half-lattice  $\mathbb{Z}_{\text{NE}}^2 = \{(m, n) : m + n \geq 0\}$ , and consider a loop percolation graph  $\text{LP}_{\text{NE}}$  defined as before but only on  $\mathbb{Z}_{\text{NE}}^2$  instead of on the entire plane. The connectivity pattern associated with it is an infinite noncrossing matching, which we denote by  $\Pi_*$ . It is defined as follows: for any  $j \in \mathbb{Z}$ , to compute  $\Pi_*(j)$ , start at the vertex  $(j, -j)$ . According to our convention regarding the coloring of the squares of the dual lattice,  $(j, -j)$  is at the bottom-left corner

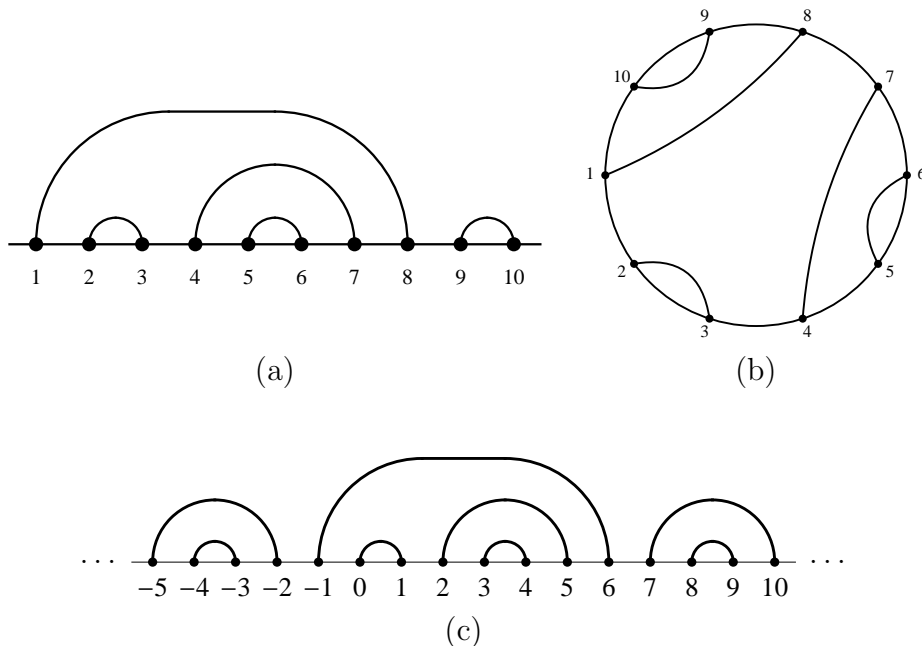
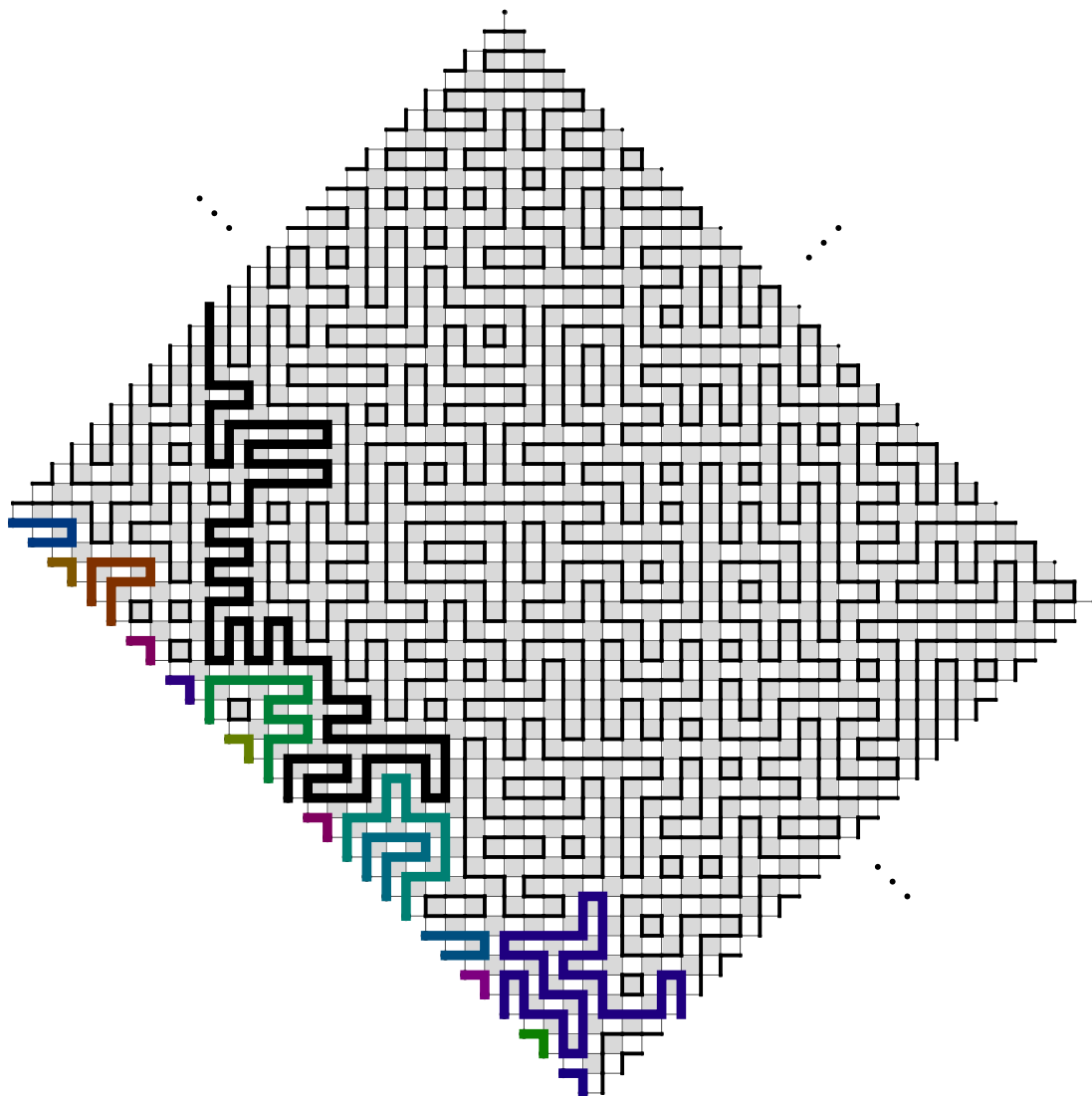


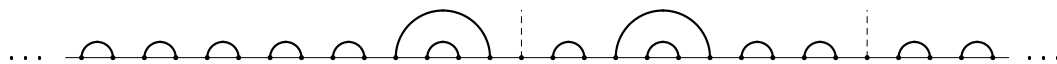
Figure 4: (a–b) a finite noncrossing matching shown as a matching of points on a line or on a circle; (c) an infinite noncrossing matching.

of a white square, so  $(j, -j)$  is incident to precisely one edge in the configuration. Follow that edge, and, continuing along the path of loop percolation edges leading out from  $(j, -j)$ , one eventually ends up at a vertex  $(k, -k)$  on the boundary of the half-lattice (since if we consider the configuration as part of a configuration on the entire plane, the path must be a loop as noted above). In this case, we say that  $j$  and  $k$  are matched, and denote  $j \xleftrightarrow{\Pi_*} k$ ; see Fig. 5 for an illustration.

It is clear from elementary topological considerations that  $\Pi_*$  is a (random) noncrossing matching. It turns out to have some remarkable distributional properties. In particular, a main motivation for the current work was the observation that the probabilities of many events associated with  $\Pi_*$  turn out rather unexpectedly to be rational numbers which can be computed explicitly. We refer to this as the **rationality phenomenon**. We will prove it rigorously in a few cases, and conjecture it for a large family of events. Although we have not been able to prove the conjecture in this generality, we also derive several results that establish strong empirical and theoretical



(a)

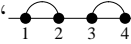


(b)

Figure 5: (a) A loop percolation configuration (the paths leading out of the vertices  $(n, -n)$  are highlighted and colored to emphasize the connectivities); (b) the associated connectivity pattern.



evidence for the correctness of the conjecture and reduce it to a much more explicit algebraic conjecture concerning the Taylor coefficients of a certain family of multivariate polynomials.

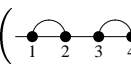
Given a finite noncrossing matching  $\pi_0 \in \text{NC}_n$  and an infinite noncrossing matching  $\pi \in \text{NC}_{\mathbb{Z}}$ , we say that  $\pi_0$  is a **submatching of  $\pi$**  if  $\pi|_{[1,2n]} \equiv \pi_0$ , and in this case denote  $\pi_0 \triangleleft \pi$ . We refer to an event of the form  $\{\pi_0 \triangleleft \Pi_*\}$  as a **submatching event**. We will often represent a submatching event schematically by drawing the diagram associated with the submatching  $\pi_0$ ; for example, the diagram “” corresponds to the submatching event  $\{(\Pi_*)_{[1,4]} = (2, 1, 4, 3)\}$ .

**Conjecture 1.3** (The rationality phenomenon for submatching events). *For any  $\pi_0 \in \text{NC}_n$ , the probability  $\mathbb{P}(\pi_0 \triangleleft \Pi_*)$  of the associated submatching event is a dyadic rational number that is computable by an explicit algorithm (Algorithm C described in Appendix A).*

Table 1 lists some of the simplest submatching events and their conjectured probabilities. We can prove the first two cases.

**Theorem 1.4.** *We have*

$$\mathbb{P}\left(\text{ \triangleleft \Pi_*\right) = \frac{3}{8}, \quad (1)$$

$$\mathbb{P}\left(\text{ \triangleleft \Pi_*\right) = \frac{97}{512}. \quad (2)$$

As we explain in Subsection 1.7, the relation (1) will follow as an immediate corollary of a result due to Fonseca and Zinn-Justin (Theorem 1.11 in Subsection 1.5). The second relation (2) can also be derived from the same theorem but in a slightly less trivial manner; see Theorem 1.16 in Subsection 1.6.

We note that the rationality phenomenon is not restricted to submatching events; a similar statement to Conjecture 1.3 also appears to hold for a larger family of **finite connectivity events**, which are events of the form  $\bigcap_{(j,k) \in A} \{j \xleftrightarrow{\Pi_*} k\}$  for some finite set  $A \subset \mathbb{Z} \times \mathbb{Z}$ . For example, in Theorem 1.18 in Subsection 1.7 we derive rigorously the value  $135/1024$  as the probability of a certain finite connectivity event that is not a submatching event. However, we mostly focus on submatching events since our main results (Theorems 1.13 and 2.24) pertain to them and as a result provide strong theoretical evidence for a rationality phenomenon in this setting.



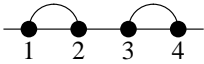
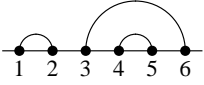
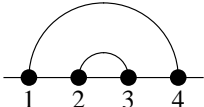
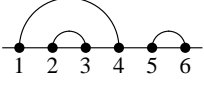
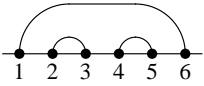
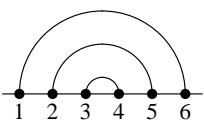
| Event   | Probability       | Event  | Probability             |
|---|-------------------|--|-------------------------|
|  | $\frac{3}{8}$     |  | $\frac{214093}{2^{21}}$ |
|  | $\frac{97}{512}$  |  | $\frac{69693}{2^{21}}$  |
|  | $\frac{59}{1024}$ |  | $\frac{69693}{2^{21}}$  |
|   |                   |  | $\frac{37893}{2^{21}}$  |
|   |                   |  | $\frac{7737}{2^{21}}$   |

Table 1: The probabilities of some submatching events in the noncrossing matching  $\Pi_*$ . The first two are established rigorously (Theorem 1.4), others are conjectured.

## 1.4 Loop percolation on a cylinder: background

Our study of half-planar loop percolation and its connectivity pattern is based on the fact that the model can be approached in a fairly straightforward manner as a limit of a model with a different geometry of a “semi-infinite cylinder.” This corresponds to loop percolation in a semi-infinite diagonal strip whose two infinite boundary edges are identified. The precise definition is as follows.

**Definition 1.5** (Cylindrical loop percolation). For any  $n \geq 1$ , the cylindrical loop percolation graph  $\text{LP}_n$  has vertex set

$$V(\text{LP}_n) = \{(x, y) \in \mathbb{Z}^2 : x + y \geq 0, -n + 1 \leq x - y \leq n\}, \quad (3)$$

where for any  $k \geq 0$  we identify the two vertices  $(-n + 1 + k, n - 1 + k)$  and  $(n + k, -n + k)$ . The edges are sampled randomly in the same manner as the usual loop percolation.

There is a convenient way of representing the cylindrical model as a tiling of a strip of the form  $[0, 2n] \times [0, \infty)$  with two kinds of square tiles, known as **plaquettes**, which are shown in Fig. 6. To see the correspondence, first rotate the diagonal strip in the original loop percolation configuration counter-clockwise by 45 degrees, then replace each (rotated) white lattice square with a plaquette whose connectivities correspond to the percolation edges in the white square. Once the plaquette tiling representation is drawn, one can wrap it around a cylinder to obtain a three-dimensional picture; see Fig. 7. The representation using plaquettes is the one traditionally used in most of the existing literature.



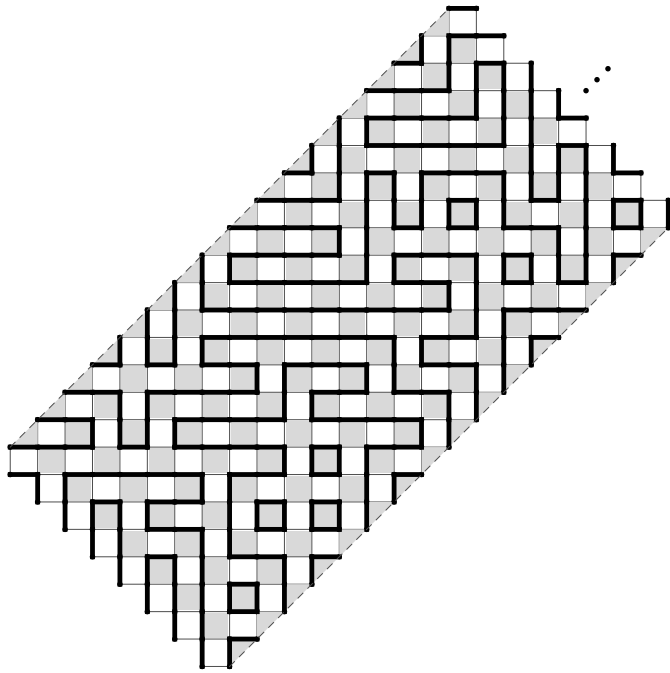
Figure 6: The two types of plaquettes.

Let  $\Pi_*^{(n)}$  denote the connectivity pattern associated with the cylindrical loop percolation graph  $\text{LP}_n$ , defined analogously to  $\Pi_*$  by following each of the paths originating at the vertices  $(-j, j)$ ,  $-n + 1 \leq j \leq n$  until it re-emerges in a vertex  $(-k, k)$ . It is easy to see that  $\Pi_*^{(n)}$  is a random *finite* noncrossing matching of order  $n$ ; the only justification needed is the following trivial lemma that shows that it is well-defined.

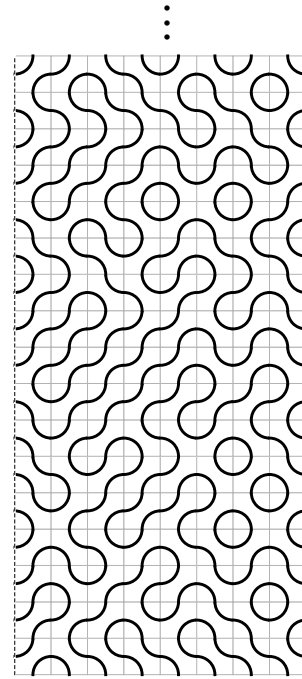
**Lemma 1.6.** *In the graph  $\text{LP}_n$ , almost surely all paths are finite.*

*Proof.* Consider the tiling representation of the model as in Fig. 7(b). It is easy to see that a row in the tiling in which the plaquettes alternate between the two types of plaquettes forces all paths below it to be finite by “bouncing back” any path attempting to cross the row. Any row has probability  $1/2^{2n-1}$  to have such structure, independently of other rows, so almost surely there will be infinitely many such rows.  $\square$

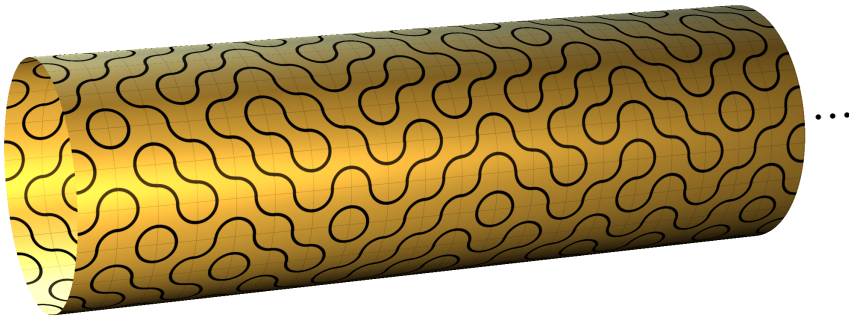
The cylindrical connectivity pattern  $\Pi_*^{(n)}$  has been the subject of extensive research in recent years, and appears in a surprising number of ways that do not seem immediately related to each other. Before presenting our new results on the behavior of  $\Pi_*^{(n)}$ , let us survey some of the known theory.



(a)



(b)



(c)

Figure 7: (a) Cylindrical loop percolation: the north-west and south-east boundary edges of the infinite diagonal strip are identified along the dashed lines; (b) representation of the same configuration as a tiling of two kinds of plaquettes; (c) wrapping a tiling of plaquettes around a cylinder.

For a matching  $\pi \in \text{NC}_n$  denote  $\mu_\pi = \mathbb{P}(\Pi_*^{(n)} = \pi)$ . A natural question is how to compute the probability vector  $\boldsymbol{\mu}_n = (\mu_\pi)_{\pi \in \text{NC}_n}$ . It is easy to see that  $\boldsymbol{\mu}_n$  is the stationary distribution of a Markov chain on  $\text{NC}_n$  where the transitions  $\pi \rightarrow \pi'$  correspond to the operation of extending the cylinder by one additional row of random plaquettes. This operation leaves the distribution  $\boldsymbol{\mu}_n$  invariant since it results in a plaquette tiling equal in distribution to the original one. Formally, there is a Markov transition matrix  $T_n^{(1/2)} = (t_{\pi, \pi'}^{(1/2)})_{\pi, \pi' \in \text{NC}_n}$  such that

$$\boldsymbol{\mu}_n T_n^{(1/2)} = \boldsymbol{\mu}_n,$$

where  $\boldsymbol{\mu}_n$  is considered as a row vector. (In the statistical physics literature  $T_n^{(1/2)}$  is usually referred to as a transfer matrix.)

The reason for the notation  $T_n^{(1/2)}$  is that we can generalize the model and define a matrix  $T_n^{(p)}$  for any  $0 < p < 1$ , corresponding to a tiling of independently sampled random plaquettes in which each choice between the two types of plaquette is made by tossing a coin with bias  $p$ . Remarkably, the value of  $p$  does not affect the distribution of the connectivity pattern.

**Theorem 1.7.** *The transition matrices  $(T_n^{(p)})_{0 < p < 1}$  are a commuting family of matrices. Consequently, since they are all stochastic, they all share the same row eigenvector  $\boldsymbol{\mu}_n$  associated with the eigenvalue 1.*

Note that using the transition matrix  $T_n^{(1/2)}$ , or  $T_n^{(p)}$  for any fixed  $p$ , is not necessarily the easiest way to compute  $\boldsymbol{\mu}_n$ , since to compute the entries of  $T_n^{(1/2)}$  one has to count the number of possible rows of  $n$  plaquettes (out of the  $2^{2n}$  possibilities) that would cause a given state transition  $\pi \rightarrow \pi'$ . It turns out that there is a more convenient (and more theoretically tractable) system of linear equations for computing  $\boldsymbol{\mu}_n$ , which involves a simpler matrix  $H_n$  that arises as a limiting case of the  $T_n^{(p)}$  as  $p \rightarrow 0$ , or symmetrically as  $p \rightarrow 1$ .

To define  $H_n$ , first define for each  $k \in [1, 2n]$  a mapping  $e_k : \text{NC}_n \rightarrow \text{NC}_n$  given by

$$e_k(\pi)(m) = \begin{cases} \pi(m) & \text{if } m \notin \{k, k+1, \pi(k), \pi(k+1)\}, \\ k+1 & \text{if } m = k, \\ k & \text{if } m = k+1, \\ \pi(k+1) & \text{if } m = \pi(k), \\ \pi(k) & \text{if } m = \pi(k+1), \end{cases} \quad (4)$$

where  $k + 1$  is interpreted as 1 if  $k = 2n$ . In words,  $e_k(\pi)$  is the matching  $\pi'$  obtained from  $\pi$  by unmatching the pairs  $k \xleftrightarrow{\pi} \pi(k)$  and  $k + 1 \xleftrightarrow{\pi} \pi(k + 1)$  and replacing them with the matched pairs  $k \xleftrightarrow{\pi'} k + 1$  and  $\pi(k) \xleftrightarrow{\pi'} \pi(k + 1)$ . In the case when  $k$  and  $k + 1$  are already matched under  $\pi$ , nothing happens and  $e_k(\pi) = \pi$ . It is easy to see that in general  $e_k(\pi)$  is a noncrossing matching. We refer to the  $e_k$  as the **Temperley-Lieb operators**. (They generate an algebra known as the **Temperley-Lieb algebra** [9, 31], but this has no bearing on the present discussion.)

**Theorem 1.8.** *Define a square matrix  $H_n$  with rows and columns indexed by elements of  $\text{NC}_n$  by*

$$(H_n)_{\pi, \pi'} = 2n - \# \{1 \leq k \leq 2n : e_k(\pi) = \pi'\}. \quad (5)$$

*Then  $\mu_n$  satisfies*

$$\mu_n H_n = 0.$$

Theorems 1.7 and 1.8 seem to be well-known to experts in the field, but their precise attribution is unclear to us. As explained in [36] (Sections 2.2.2 and 3.3.2), Theorem 1.7 follows from the Yang-Baxter equation. A fact equivalent to Theorem 1.8 is mentioned without proof in [21, Section 3]. We provide a simple proof of this result in Appendix A.

It is easy to see that the matrix  $M_n = I - \frac{1}{2n} H_n$  (where  $I$  is the identity matrix) is nonnegative and stochastic, i.e., it is a Markov transition matrix, associated with yet another Markov chain that has  $\mu_n$  as its stationary vector. That is,  $\mu_n$  can be computed by solving the vector equation  $\mu_n M_n = \mu_n$ , which can be written more explicitly as the linear system

$$\mu_\pi = \frac{1}{2n} \sum_{k=1}^{2n} \sum_{\substack{\pi' \in \text{NC}_n, \\ e_k(\pi') = \pi}} \mu_{\pi'} \quad (\pi \in \text{NC}_n). \quad (6)$$

The Markov chain  $(\pi_m)_{m \geq 0}$  associated with the transition matrix  $M_n$  has the following simple description as a random walk on  $\text{NC}_n$ , known as the **Temperley-Lieb random walk** or **Temperley-Lieb stochastic process** [24]: start with some initial matching  $\pi_0$ ; at each step, to obtain  $\pi_{m+1}$  from  $\pi_m$ , choose a uniformly random integer  $k \in \{1, 2, \dots, 2n\}$  (independently of all other random choices), and set  $\pi_{m+1} = e_k(\pi_m)$ .

The application of the maps  $e_k$  can be represented graphically by associating with each  $e_k$  a “connection diagram” of the form

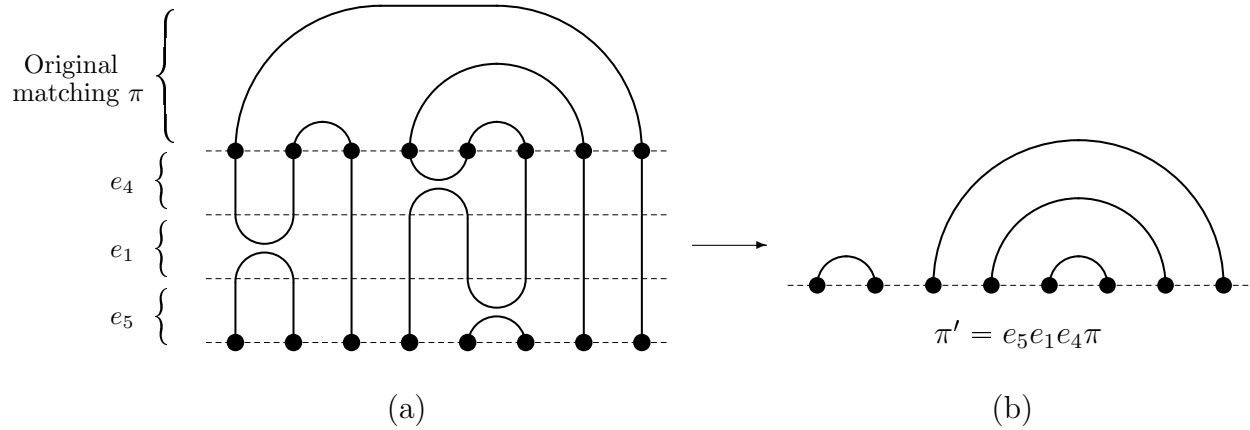
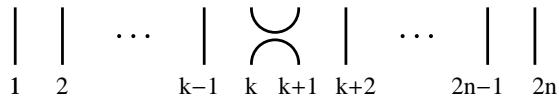


Figure 8: Graphical representation of the application of a sequence of operators  $e_k$ ,  $1 \leq k \leq 2n$  as a “composition of diagrams”: (a) the diagrams associated with operators  $e_4, e_1, e_5$  are attached to the diagram of the original matching; (b) the lines are “pulled” (and any loops are discarded) to arrive at the diagram for the transformed matching.



which will be “composed” with the diagram of the matching  $\pi$  to which  $e_k$  is applied by drawing one diagram below the other. (Note that to get the correct picture in the case  $k = 2n$  one should interpret the diagram as being drawn around a cylinder.) By composing a sequence of such diagrams with a noncrossing matching diagram one can compute the result of the application of the corresponding maps to the matching; see Fig. 8. Using this graphical interpretation, it can be seen easily that the stationary distribution  $\mu_n$  is realized as the distribution of the connectivity pattern of endpoints in an infinite composition of  $e_k$  connection diagrams (where the values of  $k$  are i.i.d. discrete uniform random variables in  $\{1, \dots, 2n\}$ ), drawn on a semi-infinite cylinder. This is illustrated in Fig. 9.

Yet another interpretation of  $\mu_n$  is as the ground state eigenvector associated with a certain quantum many-body system, the XXZ spin chain. The Hamiltonian of this spin chain is an operator acting on the space  $V = (\mathbb{C}^2)^{\otimes 2n}$ , and it has been shown that in the case of “twisted” periodic bound-

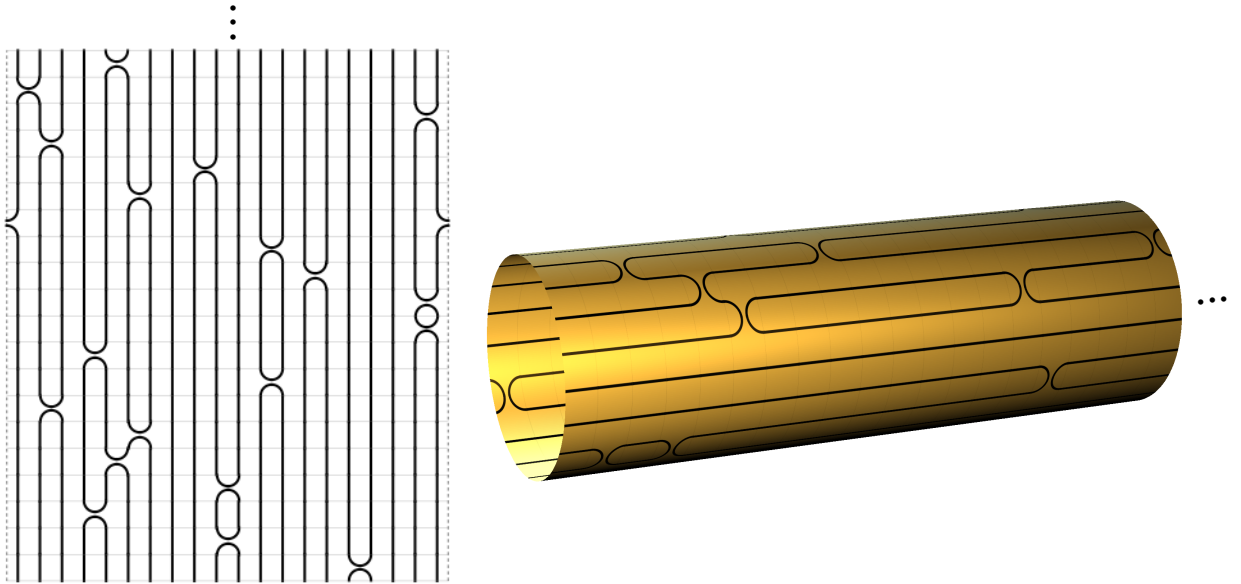


Figure 9: The connectivity pattern of endpoints in a semi-infinite arrangement of uniformly random i.i.d. Temperley-Lieb operator diagrams is invariant under the addition of another operator, hence has  $\mu_n$  as its distribution.

ary conditions and when a parameter  $\Delta$  of the chain, known as the anisotropy parameter, is set to the value  $\Delta = -1/2$ , the space  $V$  will possess a subspace  $U$  invariant under the action of the Hamiltonian, such that the restriction of the Hamiltonian to  $U$  coincides (under an appropriate choice of basis) with the operator  $H_n$  defined in (5); see [21, Section 8] and [36, Section 3.2.4].

One additional way in which the probability vector  $\mu_n$  makes an appearance is in the study of connectivity patterns associated with a different type of loop model known as the **fully packed loops** (FPLs). An FPL configuration of order  $n$  is a subset of the edges of an  $(n - 1) \times (n - 1)$  square lattice  $[0, n - 1] \times [0, n - 1]$ , to which are added  $2n$  of the  $4n$  “boundary” edges connecting the square to the rest of the lattice  $\mathbb{Z}^2$ , by starting with the edge from  $(0, 0)$  to  $(0, -1)$  and then taking alternating boundary edges as one goes around the boundary in a counter-clockwise direction (e.g., one would take edges incident to  $(2, 0)$ ,  $(4, 0)$ , etc.); these extra edges are referred to as **stubs**. The configuration is subject to the condition that any lattice vertex in the square is incident to exactly two of the configuration edges; see Fig. 10.



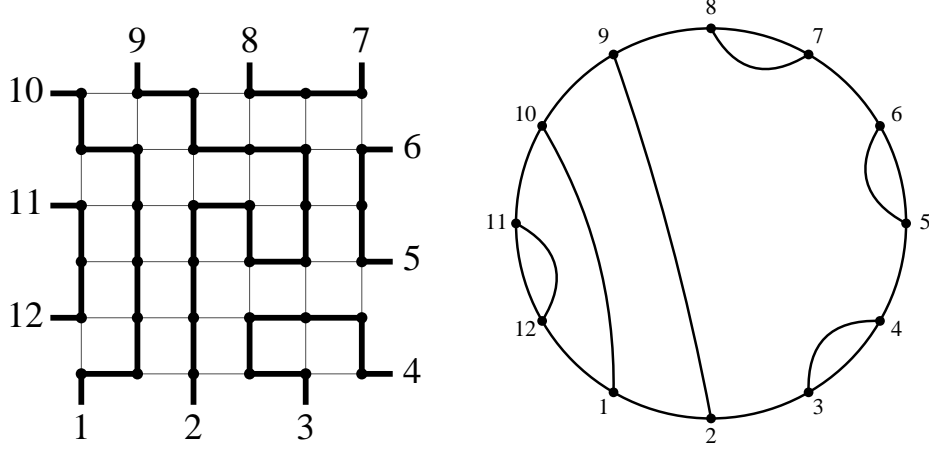


Figure 10: A fully packed loop configuration of order 6 and its associated noncrossing matching.

Let  $\text{FPL}_n$  denote the set of fully packed loop configurations of order  $n$ . A well-known bijection [25] shows that  $\text{FPL}_n$  is in correspondence with the set of **alternating sign matrices** (ASMs) of order  $n$ . These important combinatorial objects also have an interpretation as configurations of the **six-vertex model** (a.k.a. **square ice**) on an  $n \times n$  lattice with prescribed boundary behavior known as the **domain wall boundary condition**. It was conjectured by Mills, Robbins and Rumsey [20] and proved by Zeilberger [34] (see also [4], [18]) that the number  $|\text{FPL}_n|$  of ASMs of order  $n$  is given by the famous sequence of numbers 1, 2, 7, 42, 429,  $\dots$ , defined by

$$\text{ASM}(n) = \frac{1!4!7!\dots(3n-2)!}{n!(n+1)!\dots(2n-1)!}.$$

As the reader will see below, the function  $\text{ASM}(n)$  will play a central role in our current investigation of connectivity patterns in loop percolation on a cylinder.

Label the  $2n$  stubs around the square  $[0, n-1] \times [0, n-1]$  by the numbers 1 through  $2n$ , starting with the edge pointing down from  $(0, 0)$ . From the definition of fully packed loop configurations we see that any such configuration induces a connectivity pattern on the stubs, which is a noncrossing matching in  $\text{NC}_n$ , in an analogous manner to the way a loop percolation configuration on the semi-infinite cylinder does. For  $\pi \in \text{NC}_n$ , denote by

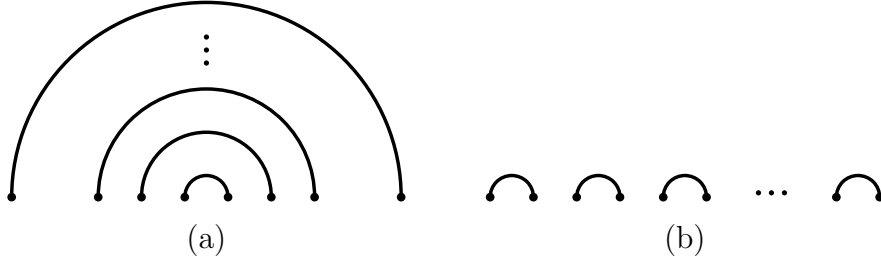


Figure 11: (a) The minimal matching  $\pi_{\min}^n$ ; (b) the maximal matching  $\pi_{\max}^n$ .

$A_n(\pi)$  the number of configurations in  $\text{FPL}_n$  whose connectivity pattern is equal to  $\pi$ .

It was first observed by Batchelor, de Gier and Nienhuis [3] that the coordinates of the vector  $\boldsymbol{\mu}_n$  are related to the enumeration of alternating sign matrices. They conjectured the following result, which was later proved by Zinn-Justin and Di Francesco [37, 38].

**Theorem 1.9** (Di Francesco-Zinn-Justin). *1. The numbers  $\mu_\pi$ , ( $\pi \in \text{NC}_n$ ) are all fractions of the form  $\alpha_n(\pi)/\text{ASM}(n)$  where  $\alpha_n(\pi)$  is an integer.*

*2. The minimal value of  $\alpha_n(\pi)$  is 1 and is attained for  $\pi = \pi_{\min}^n$ , the “minimal”<sup>1</sup> noncrossing matching consisting of  $n$  nested arcs (Fig. 11(a)).*

*3. The maximal value of  $\alpha_n(\pi)$  is  $\text{ASM}(n-1)$  and is attained for  $\pi = \pi_{\max}^n$ , the “maximal” noncrossing matching consisting of  $n$  nearest-neighbor arcs (Fig. 11(b)).*

Shortly after this discovery, Razumov and Stroganov discovered a much more precise conjecture [26] about the connection between alternating sign matrices and the vector  $\boldsymbol{\mu}_n$ ; it turns out that the correct thing to do is to look at the ASMs as fully packed loop configurations. Their conjecture, known for several years as the Razumov-Stroganov conjecture, was proved in 2010 by Cantini and Sportiello [5].

**Theorem 1.10** (Cantini-Sportiello-Razumov-Stroganov theorem). *For any*

---

<sup>1</sup>according to a certain partial ordering of noncrossing matchings that is discussed in Section 2.

$\pi \in \text{NC}_n$  we have  $\alpha_n(\pi) = A_n(\pi)$ . That is,

$$\mu_\pi = \frac{A_n(\pi)}{\text{ASM}(n)} \quad (\pi \in \text{NC}_n).$$

Probabilistically, this means that the vector  $\boldsymbol{\mu}_n$  is realized as the distribution of the connectivity pattern of a *uniformly random* FPL configuration of order  $n$ . Note that this is a finite probability space, whereas the Markov chain realization would require a potentially unbounded amount of randomness to generate a sample from  $\boldsymbol{\mu}_n$ .

Cantini and Sportiello’s proof of the conjecture of Razumov and Stroganov is highly nontrivial and involves subtle combinatorial and linear algebraic arguments. Note that while the two (*a posteriori* equivalent) definitions of  $\boldsymbol{\mu}_n$  as the distribution of connectivity patterns in cylindrical loop percolation and uniformly random fully packed loop configurations are superficially similar, the two models are quite different. In particular, the cylindrical model has an obvious symmetry under rotations of the cylinder by an angle  $2\pi/2n$ , which induces a symmetry under rotation on  $\boldsymbol{\mu}_n$ . The fact that the distribution of the connectivity pattern of random fully packed loops—which are defined on a square, not circular, geometry—also has the same symmetry, is far from obvious, and its earlier proof by Wieland [32] played an important role in Cantini and Sportiello’s analysis.

In this subsection we surveyed several settings in which the probability vector  $\boldsymbol{\mu}_n$  appears: as the distribution of the connectivity pattern of cylindrical loop percolation; as the stationary distribution of the Temperley-Lieb random walk (or, equivalently, the distribution of the connectivity pattern of a semi-infinite arrangement of connection diagrams associated with independent, uniformly random Temperley-Lieb operators); as the ground state of the XXZ spin chain; and as the distribution of the connectivity pattern of uniformly random fully packed loop configurations of order  $n$ . To conclude this discussion, we note that one of the minor results of this paper is an additional characterization of the connectivity probabilities  $(\mu_\pi)_{\pi \in \text{NC}_n}$  as coefficients in a certain natural linear-algebraic expansion; see Theorem 2.12 in Subsection 2.1.

## 1.5 Loop percolation on a cylinder: connectivity events

Having described the origins of the investigations into the random noncrossing matching  $\Pi_*^{(n)}$  and its distribution  $\boldsymbol{\mu}_n$ , we are ready to discuss the problem

of computing explicitly the probabilities of various events. Our main interest will be with submatching events of the type  $\{\pi_0 \triangleleft \Pi_*^{(n)}\}$ , where for  $\pi_0 \in \text{NC}_k$  and  $\pi \in \text{NC}_n$  ( $n \geq k$ ) the notation  $\pi_0 \triangleleft \pi$  means (as defined earlier when  $\pi$  is an infinite matching) that  $\pi_0$  is a submatching of  $\pi$ , i.e., that  $\pi|_{[1,2k]} \equiv \pi_0$ .

It was observed starting with numerical work of Mitra et al. [21], Zuber [39] and Wilson (unpublished work, cited in [39]) that the probabilities of certain events had nice formulas as rational functions in  $n$ . For example, one has empirically the relations

$$\mathbb{P} \left( \begin{array}{c} \bullet \quad \bullet \\ \text{---} \quad \text{---} \\ 1 \quad 2 \end{array} \triangleleft \Pi_*^{(n)} \right) = \frac{3}{2} \cdot \frac{n^2 + 1}{4n^2 - 1}, \quad (n \geq 1), \quad (7)$$

$$\mathbb{P} \left( \begin{array}{c} \bullet \quad \bullet \quad \bullet \quad \bullet \\ \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \\ 1 \quad 2 \quad 3 \quad 4 \end{array} \triangleleft \Pi_*^{(n)} \right) = \frac{1}{8} \cdot \frac{97n^6 + 82n^4 - 107n^2 - 792}{(4n^2 - 1)^2(4n^2 - 9)} \quad (n \geq 2), \quad (8)$$

$$\mathbb{P} \left( \begin{array}{c} \bullet \quad \bullet \quad \bullet \quad \bullet \\ \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \\ 1 \quad 2 \quad 3 \quad 4 \end{array} \triangleleft \Pi_*^{(n)} \right) = \frac{1}{16} \cdot \frac{59n^6 + 299n^4 + 866n^2 + 576}{(4n^2 - 1)^2(4n^2 - 9)} \quad (n \geq 2), \quad (9)$$

$$\begin{aligned} & \mathbb{P} \left( \begin{array}{c} \bullet \quad \bullet \quad \bullet \quad \bullet \quad \bullet \quad \bullet \\ \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \\ 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \end{array} \triangleleft \Pi_*^{(n)} \right) \\ &= \frac{1}{512} \cdot \frac{214093n^{12} - 980692n^{10} - 584436n^8 - 1887916n^6 + 1361443n^4 - 17432892n^2 - 316353600}{(4n^2 - 1)^3(4n^2 - 9)^2(4n^2 - 25)} \quad (n \geq 3), \end{aligned} \quad (10)$$

and several other such formulas, which can be discovered by a bit of experimentation after programming the linear equations (6) into a computer algebra system such as **Maple** or **Mathematica**.<sup>2</sup> The discovery of additional conjectural relations of this type is only difficult insofar as it taxes one's patience, programming skill and computational resources, but, as Zuber remarks at the end of [39], “*More conjectural expressions have been collected for other types of configurations ... but this seems a gratuitous game in the absence of a guiding principle.*” It therefore appeared sensible to wait for more theoretical developments before proceeding with the “gratuitous game.” Indeed, the first progress (and so far, to our knowledge, the only progress) on this front was made by Fonseca and Zinn-Justin [12], who, building on

---

<sup>2</sup>Unfortunately the size of the system is the Catalan number  $\text{Cat}(n)$  which grows exponentially with  $n$ , making it impractical to compute  $\mu_n$  for values of  $n$  much greater than  $n = 9$ . Zuber [39] managed the computation up to  $n = 11$  by using rotational symmetry to reduce the order of the system.

a theoretical framework developed earlier by Zinn-Justin and Di Francesco [37, 38] (see also [6]) managed to prove explicit formulas for two classes of events.

**Theorem 1.11** (Fonseca-Zinn-Justin [12]). *For each  $k \geq 1$ , Denote by  $\text{AC}_k^{(n)}$  the “anti-cluster” event that no two of the numbers  $1, \dots, k$  are matched under  $\Pi_*^{(n)}$ . Denote by  $B_k^{(n)}$  the event  $\{\pi_{\min}^k \triangleleft \Pi_*^{(n)}\}$  (the submatching event associated with the minimal matching  $\pi_{\min}^k$ ; see Fig. 11). Define a function*

$$R_k(n) = \begin{cases} \frac{\prod_{j=1}^{(k+1)/2} \prod_{m=j}^{2j-2} (n^2 - m^2)}{\prod_{j=0}^{(k-3)/2} (4n^2 - (2j+1)^2)^{(k-1)/2-j}} & k \text{ odd,} \\ \frac{\prod_{j=1}^{k/2} \prod_{m=j}^{2j-1} (n^2 - m^2)}{\prod_{j=0}^{k/2-1} (4n^2 - (2j+1)^2)^{k/2-j}} & k \text{ even.} \end{cases} \quad (11)$$

Then we have

$$\mathbb{P}(\text{AC}_k^{(n)}) = \frac{1}{\text{ASM}(k)} \frac{R_k(n)}{R_k(k)}, \quad (12)$$

$$\mathbb{P}(B_k^{(n)}) = \frac{1}{\text{ASM}(n)} \sum_{1 \leq a_1 < a_2 < \dots < a_{n-k}} \det \left( \binom{j+k}{a_i - j + k} \right)_{i,j=1}^{n-k}. \quad (13)$$

In particular, we have  $\mathbb{P}(\text{AC}_2^{(n)}) = \frac{5}{2} \frac{n^2-1}{4n^2-1}$ , which implies equation (7) since the event in that identity is the complement of  $\text{AC}_2^{(n)}$ .

Note that the approach of Fonseca and Zinn-Justin to the relation (12) started out, similarly to (13), by expressing the probabilities of these events as sums of determinants over a certain family of matrices with binomial coefficients. However, in the case of (12) a family of Pfaffian evaluations due to Krattenthaler [17] made it possible to evaluate the sum in closed form. As the authors of [12] point out, both the formulas (12) and (13) have an interesting interpretation in terms of the enumeration of certain families of totally symmetric self-complementary plane partitions.

## 1.6 Loop percolation on a cylinder: new results and the rationality phenomenon

The identities (7)–(10) generalize in a straightforward way to a conjecture on the form of the dependence on  $n$  of probabilities of submatching events.

**Conjecture 1.12** (The rationality phenomenon for submatching events; finite  $n$  case). *If  $\pi_0 \in \text{NC}_k$ , the probability of the submatching event  $\{\pi_0 \triangleleft \Pi_*^{(n)}\}$  has the form of a rational function in  $n$ , specifically*

$$\mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)}) = \frac{Q_{\pi_0}(n)}{\prod_{j=1}^k (4n^2 - j^2)^{k+1-j}} \quad (n \geq k), \quad (14)$$

where  $Q_{\pi_0}(n)$  is an even polynomial of degree  $k(k+1)$  with dyadic rational coefficients.  $Q_{\pi_0}$  can be computed by polynomial interpolation—see Algorithm B in Appendix A.

In pursuit of an approach that would lead to a proof of Conjecture 1.12, one of our goals has been to find a “guiding principle” of the type alluded to by Zuber, that would reveal the underlying structure behind the empirical phenomena described above. In Section 2 we will prove the following result, which can be thought of as one version of such a principle, and is our main result.

**Theorem 1.13** (Explicit formula for submatching event probabilities). *For any  $\pi_0 \in \text{NC}_k$ , there exists a multivariate polynomial  $F_{\pi_0}(w_1, \dots, w_k)$ , computable by an explicit algorithm (Algorithm E in Appendix A), with the following properties:*

1.  $F_{\pi_0}$  has integer coefficients.
2. All the monomials in  $F_{\pi_0}(w_1, \dots, w_k)$  are of the form  $\prod_{j=1}^k w_j^{2j-a_j}$  where  $1 \leq a_1 < \dots < a_k$  are integers satisfying  $a_j \leq 2j - 1$  for  $1 \leq j \leq k$ .<sup>3</sup>
3. The probability  $\mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)})$  is given for any  $n \geq k + 1$  by

$$\begin{aligned} \mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)}) &= \frac{1}{\text{ASM}(n)} [z_1^0 z_2^2 z_3^4 \dots z_n^{2n-2}] \left( F_{\pi_0}(z_2, \dots, z_{k+1}) \right. \\ &\quad \times \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j) \prod_{j=k+2}^n (1 + z_j) \Big), \end{aligned} \quad (15)$$

---

<sup>3</sup>Note that the number of different sequences  $(a_1, \dots, a_k)$  satisfying these conditions is  $\text{Cat}(k)$ , the  $k$ th Catalan number, and indeed the sequence  $(a_1, \dots, a_k)$  can be thought of as encoding a noncrossing matching in  $\text{NC}_k$ , a fact that will have a role to play later on—see Section 2.

where  $[z_1^{m_1} \dots z_n^{m_n}]g(z_1, \dots, z_n)$  denotes the coefficient of the monomial  $z_1^{m_1} \dots z_n^{m_n}$  in a polynomial  $g(z_1, \dots, z_n)$ .

The formula (15) can be recast in two equivalent forms which some readers may find more helpful: first, as a constant term identity

$$\mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)}) = \frac{1}{\text{ASM}(n)} \text{CT}_{z_1, \dots, z_n} \left( F_{\pi_0}(z_2, \dots, z_{k+1}) \right. \\ \left. \times \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j) \frac{\prod_{j=k+2}^n (1 + z_j)}{\prod_{j=1}^n z_j^{2j-2}} \right),$$

where  $\text{CT}_{z_1, \dots, z_n} g(z_1, \dots, z_n)$  denotes the constant term of a Laurent polynomial  $g(z_1, \dots, z_n)$ ; and second, as a multi-dimensional complex contour integral

$$\mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)}) = \frac{1}{\text{ASM}(n)} \oint \dots \oint \left( F_{\pi_0}(z_2, \dots, z_{k+1}) \right. \\ \left. \times \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j) \frac{\prod_{j=k+2}^n (1 + z_j)}{\prod_{j=1}^n z_j^{2j-1}} \right) \prod_{j=1}^n \frac{dz_j}{2\pi i},$$

where the contour for each of the  $z_j$ 's is a circle of arbitrary radius around 0.

Table 2 lists a few of the simplest submatching events and the multivariate polynomials associated to them. Note that the first case of the empty matching listed in the table (for which the probability of the submatching event is 1) corresponds to the known identity

$$\text{ASM}(n) = [z_1^0 z_2^2 \dots z_n^{2n-2}] \left( \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j) \prod_{j=2}^n (1 + z_j) \right). \quad (16)$$

This identity is already a difficult result. It was proved in [38] by Zinn-Justin and Di Francesco, relying conditionally on a conjectural anti-symmetrization identity that they discovered which itself was proved shortly afterwards by Zeilberger [35]. Their proof also relies on a highly nontrivial Pfaffian evaluation due to Andrews [1] (see also [2, 17]). The following conjecture can be thought of as a natural generalization or “deformation” of (16).

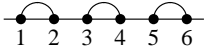

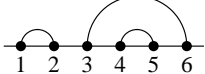
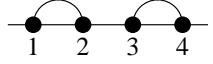
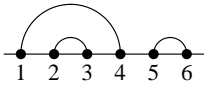
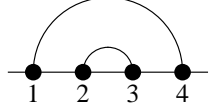
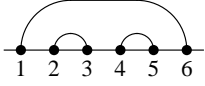
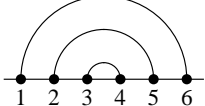
| k | $\pi_0$   | $\frac{F_{\pi_0}(w_1, \dots, w_k)}{w_1 \dots w_{2k}}$ | k | $\pi_0$  | $\frac{F_{\pi_0}(w_1, \dots, w_k)}{w_1 \dots w_{2k}}$ |
|---|---|---|---|--|---|
| 0 | empty   | 1   | 3 |  | 1   |
| 1 |  | 1   | 3 |  | $w_2 - w_1 w_2$                                       |
| 2 |  | 1   | 3 |  | $w_1$   |
| 2 |  | $w_1$   | 3 |  | $w_1 w_2$   |
|   |   |   | 3 |  | $w_1 w_2^2$   |

Table 2: The polynomials  $F_{\pi_0}$  associated with some submatching events, after factoring out the product  $w_1 \dots w_{2k}$  which always divides  $F_{\pi_0}$ . (The first case of the “empty” matching refers to the trivial matching of order  $k = 0$ , in which case the submatching event has probability 1.)

**Conjecture 1.14** (The algebraic rationality phenomenon). *Let  $k \geq 1$ , and let  $1 \leq a_1 < \dots < a_k$  be integers satisfying  $a_j \leq 2j - 1$  for all  $j$ . There exists a rational function of the form  $R(n) = P(n) / \prod_{j=1}^k (4n^2 - j^2)^{k+1-j}$ , where  $P(n)$  is an even polynomial of degree at most  $k(k+1)/2$  with dyadic rational coefficients, such that for all  $n \geq k+1$  we have that*

$$\begin{aligned}
& [z_1^0 z_2^2 \dots z_n^{2n-2}] \left( \prod_{j=1}^k z_{j+1}^{2j-a_j} \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j) \prod_{j=k+2}^n (1 + z_j) \right) \\
& = \text{ASM}(n) R(n).
\end{aligned} \tag{17}$$

It is unclear whether the specific assumption about the sequence  $a_1, \dots, a_k$



that enters the form of the monomial  $\prod_{j=1}^k z_{j+1}^{2j-a_j}$  is necessary to imply the rationality of  $R(n)$  in (17). The main assumption, which may be sufficient to imply the result, is that we are looking at a Taylor coefficient that is “a fixed distance away” from the coefficient in (16).

An immediate consequence of Theorem 1.13 is that Conjecture 1.14 essentially implies Conjecture 1.12, although a small gap remains regarding the validity of (14) in the case  $n = k$ .

**Theorem 1.15.** *Conjecture 1.14 implies a weaker version of Conjecture 1.12 in which (14) is only claimed to hold for  $n \geq k + 1$ .*

From the above discussion we see that, while we have been unable to get a complete understanding of the probabilities of submatching events, we have reduced the problem to the essentially algebraic question of understanding the form of the dependence of  $n$  of the polynomial coefficients appearing on the left-hand side of (17). Moreover, the algorithm for finding the polynomial  $F_{\pi_0}$  associated with a submatching event, which will be explained in Section 2, stems from a fairly detailed theoretical understanding of the model, and therefore already helps eliminate much of the mystery surrounding identities such as (7)–(10).

It should be noted as well that the relations (16) and (17) belong to a large family of identities known as **constant term identities**. The study of such identities became popular following the discovery by Dyson of the identity

$$\text{CT}_{z_1, \dots, z_n} \left( \prod_{1 \leq i \neq j \leq n} \left( 1 - \frac{z_j}{z_i} \right)^{a_j} \right) = \frac{(a_1 + \dots + a_n)!}{a_1! \dots a_n!} \quad (a_1, \dots, a_n \geq 0), \quad (18)$$

which became known as the Dyson conjecture [7]. (Dyson’s conjecture was proved by Gunson [15] and Wilson [33], and a particularly simple proof was later found by Good [13].) Research on such identities has been an active area that involves a mixture of techniques from algebraic combinatorics and the theory of special functions. In particular, Sills and Zeilberger [28] proved a deformation of Dyson’s identity in which a coefficient near the constant term is shown to equal the multinomial coefficient on the right-hand side of (18) times a rational function in the exponents  $a_1, \dots, a_n$ ; this is quite similar in spirit to the claim of Conjecture 1.14.

To conclude this section, we show how the relation (8) can be derived in a simple manner from the results of Fonseca and Zinn-Justin, and in addition

derive another identity concerning a finite connectivity event which is not a submatching event.

**Theorem 1.16.** *The identity (8) holds for  $n \geq 2$ ,<sup>4</sup> and we have the additional identity*

$$\mathbb{P} \left( \text{---} \overset{\curvearrowright}{\underset{1}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{4}{\bullet}} \text{---} \triangleleft \Pi_*^{(n)} \right) = \frac{15}{16} \cdot \frac{(n^2 - 4)(9n^4 + 38n^2 - 63)}{(4n^2 - 1)^2(4n^2 - 9)} \quad (n \geq 3) \quad (19)$$

(we use the notation for submatching events for convenience, but note that this refers to the event that  $\Pi_*^{(n)}$  matches the two pairs  $1 \overset{\Pi_*^{(n)}}{\longleftrightarrow} 2$  and  $4 \overset{\Pi_*^{(n)}}{\longleftrightarrow} 5$ ).

*Proof.* By (12), we know that

$$\mathbb{P}(\text{AC}_4^{(n)}) = \frac{33}{8} \cdot \frac{(n^2 - 1)(n^2 - 4)(n^2 - 9)}{(4n^2 - 1)^2(4n^2 - 9)}. \quad (20)$$

On the other hand, by the inclusion-exclusion principle, we have that

$$\begin{aligned} \mathbb{P}(\text{AC}_4^{(n)}) &= 1 - \mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{1}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \right) - \mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \right) - \mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{4}{\bullet}} \text{---} \right) \\ &\quad + \mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{1}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{4}{\bullet}} \text{---} \right) \\ &= 1 - 3 \cdot \frac{3}{2} \cdot \frac{n^2 + 1}{4n^2 - 1} + \mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{1}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{4}{\bullet}} \text{---} \right) \end{aligned}$$

(using (7) and the rotation-invariance of  $\Pi_*^{(n)}$ ). Substituting the probability from (20) and solving for  $\mathbb{P} \left( \Pi_*^{(n)} \in \text{---} \overset{\curvearrowright}{\underset{1}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{2}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{3}{\bullet}} \text{---} \overset{\curvearrowright}{\underset{4}{\bullet}} \text{---} \right)$  gives (8). For the second identity (19), perform a similar inclusion-exclusion computation for the event  $\text{AC}_5^{(n)}$ , whose probability is given according to (12) by

$$\mathbb{P}(\text{AC}_5^{(n)}) = \frac{11}{16} \cdot \frac{(n^2 - 1)(n^2 - 4)(n^2 - 9)}{(4n^2 - 1)^2(4n^2 - 9)}.$$

The details of the computation are easy and left to the reader. □

---

<sup>4</sup>This proves part of Conjecture 9 in Zuber's paper [39].

## 1.7 Consequences for loop percolation on a half-plane

Let us return to the original setting of loop percolation on a half-plane discussed earlier. The following result allows us to deduce exact results on probabilities of local connectivity events in the half-plane from corresponding results in the cylindrical model.

**Lemma 1.17.** *Let  $A \subset \mathbb{Z} \times \mathbb{Z}$  be a finite set, and let  $E_n$  denote the finite connectivity events*

$$E_n = \bigcap_{(j,k) \in A} \left\{ j \overset{\Pi^{(n)}}{\longleftrightarrow} k \right\}$$

*associated with  $A$  (which are defined for large enough  $n$ ). Then we have  $\mathbb{P}(E_n) \rightarrow \mathbb{P}(E)$  as  $n \rightarrow \infty$ , where  $E = \bigcap_{(j,k) \in A} \left\{ j \overset{\Pi_*}{\longleftrightarrow} k \right\}$ .*

*Proof.* Couple the cylindrical and half-planar loop percolation by using the same random bits to select the edge configuration incident to the vertices in the region  $V(\text{LP}_n)$  (defined in (3)). With this coupling, it is easy to see that the symmetric difference  $E_n \triangle E$  of the events  $E_n$  and  $E$  is contained in the event  $B_n$  that one of the half-planar loop percolation paths starting at the point  $(j, -j)$  for some  $j$  belonging to one of the pairs  $(j, k) \in A$  reaches the complement  $\mathbb{Z}^2 \setminus V(\text{LP}_n)$ . Since the loop percolation paths are almost surely finite, we have  $\lim_{n \rightarrow \infty} \mathbb{P}(B_n) = \mathbb{P}(\bigcap_{n=1}^{\infty} B_n) = 0$ , and therefore we get that  $|\mathbb{P}(E_n) - \mathbb{P}(E)| \leq \mathbb{P}(E_n \triangle E) \leq \mathbb{P}(B_n) \rightarrow 0$  as  $n \rightarrow \infty$ .  $\square$

As immediate corollaries from the lemma we get the following facts: first, Conjecture 1.12 (or the weaker version of it mentioned in Theorem 1.15) implies Conjecture 1.3. Second, Theorem 1.4 follows as a limiting case of (7) and (8). Third, we have the following explicit formulas for events in the half-plane model corresponding to (12) and (19).

**Theorem 1.18.** *For  $k \geq 1$ , let  $\text{AC}_k$  denote the anti-cluster event that no two of the numbers  $1, \dots, k$  are matched under  $\Pi_*$ . We have the formulas*

$$\mathbb{P}(\text{AC}_k) = \frac{1}{2^{\lfloor k/2 \rfloor \cdot \lfloor k/2 + 1 \rfloor} \text{ASM}(k) R_k(k)},$$

*where  $R_k(\cdot)$  is defined in (11) (see Table 3), and*

$$\mathbb{P}\left(\begin{array}{c} \text{---} \bullet_1 \text{---} \bullet_2 \text{---} \bullet_3 \text{---} \bullet_4 \text{---} \bullet_5 \text{---} \\ \text{---} \text{---} \end{array} \triangleleft \Pi_* \right) = \frac{135}{1024}.$$

| $k$                                 | 1             | 2             | 3                | 4                 | 5                     | 6                   | 7                       |
|-------------------------------------|---------------|---------------|------------------|-------------------|-----------------------|---------------------|-------------------------|
| $\mathbb{P}(\Pi_* \in \text{AC}_k)$ | $\frac{5}{8}$ | $\frac{1}{4}$ | $\frac{33}{512}$ | $\frac{11}{1024}$ | $\frac{2431}{2^{21}}$ | $\frac{85}{2^{20}}$ | $\frac{126293}{2^{35}}$ |

Table 3: Probabilities of the anti-cluster event  $\text{AC}_k$  for  $k = 1, \dots, 7$ .

## 2 The theory of wheel polynomials and the qKZ equation

Our goal in this section is to prove Theorem 1.13. The proof will be based on an extension of an algebraic theory that was developed in a recent series of papers [12, 36, 37, 38], where it is shown that a tool from the statistical physics literature known as the **quantum Knizhnik-Zamolodchikov equation** (or **qKZ equation**) can be applied to the study of the connectivity pattern of cylindrical loop percolation.

Mathematically, the qKZ equation as applied to the present setting reduces to the analysis of a vector space of multivariate polynomials satisfying a condition known as the **wheel condition**. (Similar and more general wheel conditions are also discussed in the papers [8, 16, 23].) We call such polynomials **wheel polynomials**. In Subsections 2.1–2.3 we will survey the known results regarding the theory of wheel polynomials that are needed for our purposes. In Subsection 2.4 we present a new result (Theorem 2.24) on expansions of families of wheel polynomials associated with submatching events, that is another main result of the paper. In Subsection 2.5 we show how to derive Theorem 1.13 from Theorem 2.24.

### 2.1 Background

Let  $\text{NC}_n$  denote as before the set of noncrossing matchings of  $1, \dots, 2n$ . It is well-known that elements of  $\text{NC}_n$  are in canonical bijection with the set of **Dyck paths** of length  $2n$ , and with the set of Young diagrams contained in the staircase shape  $(n-1, n-2, \dots, 1)$ . These bijections will play an important role, and are illustrated in Fig. 12; see [12, Section 2] for a detailed explanation.

We adopt the following notation related to these bijections. If  $\pi \in \text{NC}_n$ , let  $\pi_k$  be 1 if  $\pi(k) > k$  (“ $k$  is matched to the right”) or  $-1$  if  $\pi(k) < k$  (“ $k$  is matched to the left”). The vector  $(\pi_1, \dots, \pi_{2n})$  is the encoding of  $\pi$  as the

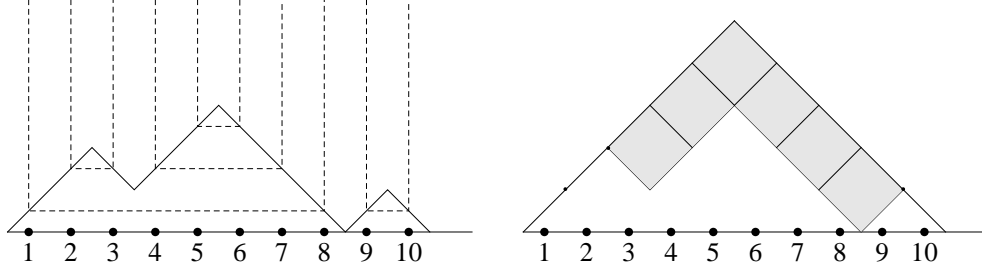


Figure 12: The noncrossing matching from Fig. 4(a) represented as a Dyck path and as a Young diagram.

sequence of steps in the associated Dyck path. Let  $\pi^+$  denote the sequence  $(a_1, \dots, a_n)$  of positions where  $\pi_k = 1$  (the positions of increase of the Dyck path). Encoding  $\pi$  in this way maps  $\text{NC}_n$  bijectively onto the set

$$\text{NC}_n^+ = \left\{ \mathbf{a} = (a_1, \dots, a_n) : 1 \leq a_1 < \dots < a_n \leq 2n-1 \right. \\ \left. \text{and } a_j \leq 2j-1 \text{ for all } j \right\}. \quad (21)$$

The Young diagram associated with a noncrossing matching  $\pi \in \text{NC}_n$  is denoted  $\lambda_\pi$ . For  $\pi, \sigma \in \text{NC}_n$ , denote  $\sigma \nearrow \pi$  if  $\lambda_\sigma$  is obtained from  $\lambda_\pi$  by the addition of a single box, and  $\pi \leq \sigma$  if  $\lambda_\pi$  is contained in  $\lambda_\sigma$ . (This partial order on  $\text{NC}_n$  is precisely the order with respect to which the matchings  $\pi_{\min}^n$  and  $\pi_{\max}^n$  defined in Subsection 1.4 are minimal and maximal, respectively.) Denote  $\pi \nearrow_j \sigma$  if  $\lambda_\sigma$  is obtained from  $\lambda_\pi$  by adding a box in a position that, in the coordinate system of Fig. 12, lies vertically above the positions  $j$  and  $j+1$  on the horizontal axis.

**Definition 2.1** (Wheel polynomials). Let  $q = e^{2\pi i/3}$ . A **wheel polynomial of order  $n$**  is a polynomial  $p(\mathbf{z}) = p(z_1, \dots, z_{2n})$  having the following properties:

1.  $p$  is a homogeneous polynomial of total degree  $n(n-1)$ ;
2.  $p$  satisfies the **wheel condition**

$$p(z_1, \dots, z_{2n}) \Big|_{z_k = q^2 z_j = q^4 z_i} = 0 \quad (1 \leq i < j < k \leq 2n). \quad (22)$$

Denote by  $W_n[\mathbf{z}]$  the vector space of wheel polynomials of order  $n$ .

It is easy to see that the condition (22) depends only on the cyclical ordering of the variables  $z_1, \dots, z_{2n}$  (in other words, the space  $W_n[\mathbf{z}]$  is invariant under the rotation action  $p(z_1, \dots, z_{2n}) \mapsto p(z_2, \dots, z_{2n}, z_1)$ ); hence the name “wheel polynomials.”

The algebraic properties of wheel polynomials are quite elegant and of direct relevance to our study of connectivity patterns in loop percolation. Below we survey some of the known theory.

**Example 2.2.** Let  $p(\mathbf{z}) = \prod_{1 \leq i < j \leq n} (qz_i - q^{-1}z_j) \prod_{n+1 \leq i < j \leq 2n} (qz_i - q^{-1}z_j)$ . It is easy to check that  $p$  is a wheel polynomial. Indeed, it is homogeneous of the correct degree, and if  $1 \leq i < j < k \leq 2n$  and we make the substitution  $z_k = q^2 z_j = q^4 z_i$ , then, if  $j \leq n$  there will be a zero factor in the first product  $\prod_{1 \leq i < j \leq n} (qz_i - q^{-1}z_j)$ , and similarly if  $j > n$  then there will be a zero factor in the second product. In either case the wheel condition (22) is satisfied.

Two types of pointwise evaluations of a wheel polynomial will be of particular interest: at the point  $\mathbf{z} = (1, 1, \dots, 1)$  and at a point  $\mathbf{z} = (q^{-\pi_1}, \dots, q^{-\pi_{2n}})$  associated with the steps of a Dyck path, where  $\pi \in \text{NC}_n$ . As a shorthand, we denote  $p(\mathbf{1}) = p(1, \dots, 1)$  and  $p(\pi) = p(q^{-\pi_1}, \dots, q^{-\pi_{2n}})$  and refer to these values as the **1-evaluation** and  **$\pi$ -evaluation** of  $p$ , respectively.

**Theorem 2.3.** *A polynomial  $p \in W_n[\mathbf{z}]$  is determined uniquely by its values  $p(\pi)$  as  $\pi$  ranges over the noncrossing matchings in  $\text{NC}_n$ . Equivalently, the  $\pi$ -evaluation linear functionals  $(\text{ev}_\pi)_{\pi \in \text{NC}_n}$  defined by  $\text{ev}_\pi(p) = p(\pi)$  span the dual vector space  $(W_n[\mathbf{z}])^*$ .*

*Proof.* See [11, Appendix C]. □

In particular, it follows that  $\dim W_n[\mathbf{z}] \leq |\text{NC}_n| = \text{Cat}(n)$ . We shall soon see that this is an equality, that is, the evaluation functionals  $(\text{ev}_\pi)_{\pi \in \text{NC}_n}$  are in fact a basis of  $(W_n[\mathbf{z}])^*$ . The proof of this fact involves a remarkable explicit construction of a family of wheel polynomials that will turn out to be a basis of  $W_n[\mathbf{z}]$  dual to  $(\text{ev}_\pi)_{\pi \in \text{NC}_n}$ , and will play a central role in our analysis.

For a multivariate polynomial  $p \in \mathbb{C}[z_1, \dots, z_m]$  and  $1 \leq j < m$ , define the divided difference operator  $\partial_j : \mathbb{C}[z_1, \dots, z_m] \rightarrow \mathbb{C}[z_1, \dots, z_m]$  by

$$(\partial_j p)(z_1, \dots, z_m) = \frac{p(z_1, \dots, z_{j+1}, z_j, \dots, z_{2n}) - p(z_1, \dots, z_j, z_{j+1}, \dots, z_{2n})}{z_{j+1} - z_j}.$$

The next two lemmas are easy to check; see [36, Sec. 4.2.1] for a related discussion.

**Lemma 2.4.** *If  $p \in W_n[\mathbf{z}]$  and  $1 \leq j \leq 2n - 1$ , then  $(qz_j - q^{-1}z_{j+1})\partial_j p$  is also in  $W_n[\mathbf{z}]$ .*

**Lemma 2.5.** *If for some  $1 \leq j \leq 2n - 1$ ,  $\pi \in \text{NC}_n$  satisfies  $j \xleftrightarrow{\pi} j + 1$ , then the set of preimages  $e_j^{-1}(\pi)$  of  $\pi$  under  $e_j$  (the Temperley-Lieb operator defined in (4)) consists of:*

1. *a noncrossing matching  $\hat{\pi} \in \text{NC}_n$  satisfying  $\pi \nearrow_j \hat{\pi}$ , if such a matching exists (i.e., if the associated Young diagram is still contained in the staircase shape  $(n - 1, \dots, 1)$ ); and*
2. *a set of noncrossing matchings  $\sigma \in \text{NC}_n$  satisfying  $\sigma \leq \pi$ .*

The last two lemmas make it possible to construct a family of wheel polynomials indexed by noncrossing matchings in  $\text{NC}_n$  (or, equivalently, by Young diagrams contained in the staircase shape  $(n - 1, \dots, 1)$ ) recursively. We start with an explicit polynomial known to be in  $W_n[\mathbf{z}]$ —a scalar multiple of the polynomial from Example 2.2 above—which we associate with the minimal matching  $\pi_{\min}^n$  (which corresponds to the empty Young diagram). We then define for each noncrossing matching  $\pi$  a polynomial obtained from the polynomials of matchings preceding  $\pi$  in the order  $\leq$  using linear combinations and the operation of Lemma 2.4. The precise definition is as follows.

**Definition 2.6** (qKZ basis). The **qKZ polynomials** are a family of polynomials  $(\Psi_\pi)_{\pi \in \text{NC}_n}$  defined using the following recursion on Young diagrams:

$$\Psi_{\pi_{\min}^n}(\mathbf{z}) = (-3)^{-\binom{n}{2}} \prod_{1 \leq i < j \leq n} (qz_i - q^{-1}z_j) \prod_{n+1 \leq i < j \leq 2n} (qz_i - q^{-1}z_j), \quad (23)$$

$$\Psi_\pi(\mathbf{z}) = (qz_j - q^{-1}z_{j+1}) \partial_j \Psi_\sigma - \sum_{\nu \in e_j^{-1}(\sigma) \setminus \{\pi, \sigma\}} \Psi_\nu \quad \text{if } \sigma \nearrow_j \pi. \quad (24)$$

Since for a given noncrossing matching  $\pi$ , the choices of  $\sigma$  and  $j$  such that  $\sigma \nearrow_j \pi$  are in general not unique, it is not clear that the above definition makes sense. The fact that it does is a nontrivial statement, which we include as part of the next result.

**Theorem 2.7.** *The qKZ polynomials satisfy:*

1.  *$\Psi_\pi$  is well-defined, i.e., the result of the recursive computation does not depend on the order in which boxes are added to the Young diagram.*

2.  $\Psi_\pi \in W_n[\mathbf{z}]$ .

3. Let  $\rho(\pi)$  denote the rotation operator acting on  $\pi$ , defined by  $(\rho(\pi))(k) = \pi(k+1)$  for  $1 \leq k < 2n$  and  $(\rho(\pi))(2n) = \pi(1)$ . Then we have

$$\Psi_{\rho(\pi)}(z_1, \dots, z_{2n}) = \Psi_\pi(z_2, \dots, z_{2n}, z_1). \quad (25)$$

*Proof.* See [36, Section 4.2].  $\square$

As explained in [36, Sections 4.1–4.2], equations (23)–(25) together are equivalent to a different system of equations which forms (a special case of) the qKZ equation. A different way of solving the same system, which we will not use here, is described in [10].

If  $\pi \in \text{NC}_n$  and  $1 \leq j \leq 2n-1$  is a number such that  $j \xleftrightarrow{\pi} j+1$ , denote by  $\hat{\pi}_j$  the noncrossing matching in  $\text{NC}_{n-1}$  obtained by deleting the arc connecting  $j$  and  $j+1$  from the diagram of the matching and relabelling the remaining elements. This operation makes it possible to perform many computations recursively. The next result provides an important example.

**Theorem 2.8.** *If  $j \xleftrightarrow{\pi} j+1 \in \text{NC}_n$  then for any  $\sigma \in \text{NC}_n$  we have*

$$\Psi_\pi(\sigma) = \begin{cases} 3^{n-1} \Psi_{\hat{\pi}_j}(\hat{\sigma}_j) & \text{if } j \xleftrightarrow{\sigma} j+1, \\ 0 & \text{otherwise.} \end{cases} \quad (26)$$

*Proof.* See [36, Section 4.2].  $\square$

**Corollary 2.9.** *1. For all  $\pi, \sigma \in \text{NC}_n$ , we have*

$$\Psi_\pi(\sigma) = \delta_{\pi, \sigma} = \begin{cases} 1 & \text{if } \pi = \sigma, \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

2. The qKZ polynomials  $(\Psi_\pi)_{\pi \in \text{NC}_n}$  are a basis for  $W_n[\mathbf{z}]$ .

3. The  $\pi$ -evaluation linear functionals  $(\text{ev}_\pi)_{\pi \in \text{NC}_n}$  are the basis of  $(W_n[\mathbf{z}])^*$  dual to  $(\Psi_\pi)_{\pi \in \text{NC}_n}$ .

4.  $\dim W_n[\mathbf{z}] = \text{Cat}(n)$ .

*Proof.* Part 1 follows by induction from (26), and parts 2–4 are immediate from part 1.  $\square$



**Example 2.10** ([36, Section 4.3.1]). Let  $p(\mathbf{z}) = s_\lambda(z_1, \dots, z_{2n})$ , the Schur polynomial associated with the Young diagram  $\lambda = (n-1, n-1, \dots, 2, 2, 1, 1)$ . That is, we have explicitly

$$p(\mathbf{z}) = \left( \prod_{1 \leq i < j \leq 2n} (z_j - z_i) \right)^{-1} \det (z_i^j)_{\substack{1 \leq i \leq 2n, \\ 0 \leq j \leq 3n-2, j \equiv 0, 1 \pmod{3}}} \quad (28)$$

Then  $p$  is homogeneous of degree  $n(n-1)$ , and the matrix whose determinant appears in (28) has the property that if we make the substitution  $z_k = q^2 z_j = q^4 z_i$  for some  $1 \leq i < j < k \leq 2n$ , the columns with index  $i, j, k$  of the matrix become linearly dependent (take the linear combination with coefficients  $1, q^2, q^4$ ), causing  $p$  to be 0. Thus,  $p$  is a wheel polynomial.

Since  $p$  is also a symmetric polynomial, it satisfies  $p(\pi) = p(\pi')$  for any  $\pi, \pi' \in \text{NC}_n$ . By Theorem 2.3 it follows that  $p$  is the unique symmetric wheel polynomial up to scalar multiplication. Note that this Schur function is also known to be the partition function of the six-vertex model with domain wall boundary condition at the “combinatorial point,” and played an important role in the study of the enumeration of alternating sign matrices; see [22, 30] and [36, Section 2.5.6].

Denote  $\psi_\pi = \Psi_\pi(\mathbf{1})$  and  $\boldsymbol{\psi}_n = (\psi_\pi)_{\pi \in \text{NC}_n}$ . The vector  $\boldsymbol{\psi}_n$  is related to the connectivity pattern  $\Pi_*^{(n)}$  in cylindrical loop percolation, as explained in the following result.

**Theorem 2.11.** *The vector  $\boldsymbol{\psi}_n$  is a solution of the linear system (6). Consequently it is a scalar multiple of the probability distribution  $\boldsymbol{\mu}_n$  of  $\Pi_*^{(n)}$ . That is, for any  $\pi \in \text{NC}_n$  we have  $\psi_\pi = a_n \mu_\pi$ , where  $a_n = \sum_{\pi \in \text{NC}_n} \psi_\pi$ .*

*Proof.* See [36, Sec. 4.3]. □

We will see later (see Theorem 2.22 below) that the normalization constant  $a_n$  is equal to  $\text{ASM}(n)$ . Assuming this fact temporarily, Theorem 2.11 implies the following interesting and purely algebraic characterization of the numbers  $\mu_\pi = \mathbb{P}(\Pi_*^{(n)} = \pi)$ , which does not seem to have been noted before.

**Theorem 2.12.** *For any wheel polynomial  $p \in W_n[\mathbf{z}]$ , we have that  $p(\mathbf{1}) = \text{ASM}(n) \sum_{\pi \in \text{NC}_n} \mu_\pi p(\pi)$ . In other words, the  $\mathbf{1}$ -evaluation functional  $\text{ev}_\mathbf{1}$  has the expansion*

$$\text{ev}_\mathbf{1} = \text{ASM}(n) \sum_{\pi \in \text{NC}_n} \mu_\pi \text{ev}_\pi$$

as a linear combination of the  $\pi$ -evaluation functionals.

*Proof.* If  $p = \sum_{\pi \in \text{NC}_n} b_\pi \Psi_\pi$  gives the expansion of  $p$  in terms of the qKZ basis polynomials, then

$$p(\mathbf{1}) = \sum_{\pi \in \text{NC}_n} b_\pi \Psi_\pi(\mathbf{1}) = \sum_{\pi \in \text{NC}_n} b_\pi \psi_\pi = \text{ASM}(n) \sum_{\pi \in \text{NC}_n} \mu_\pi b_\pi.$$

On the other hand, taking the  $\pi$ -evaluation of the expansion  $p = \sum_{\sigma \in \text{NC}_n} b_\sigma \Psi_\sigma$  easily implies using (27) that  $b_\pi = p(\pi)$ .  $\square$

A referee has pointed out to us that Theorem 2.12 can also be deduced from some of the results (specifically, Theorem 2 and equations (7) and (8)) of [10].

Next, we construct a second important family of wheel polynomials defined using contour integration.

**Definition 2.13** (Integral wheel polynomials). Denote

$$\mathcal{A}_n = \left\{ \mathbf{a} = (a_1, \dots, a_n) : 1 \leq a_1 \leq \dots \leq a_n \leq 2n-1 \right. \\ \left. \text{and } a_j \leq 2j-1 \text{ for all } j \right\}.$$

(These sequences generalize the steps of increase of Dyck paths—compare with (21).) For a sequence  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}_n$ , denote

$$\Phi_{\mathbf{a}}(z_1, \dots, z_{2n}) = (-1)^{\binom{n}{2}} \prod_{1 \leq j < k \leq 2n} (qz_j - q^{-1}z_k) \\ \times \oint \dots \oint \frac{\prod_{1 \leq j < k \leq n} (w_k - w_j)(qw_j - q^{-1}w_k)}{\prod_{j=1}^n \left( \prod_{k=1}^{a_j} (w_j - z_k) \prod_{k=a_j+1}^{2n} (qw_j - q^{-1}z_k) \right)} \prod_{j=1}^n \frac{dw_j}{2\pi i}, \quad (29)$$

where the contour of integration for each variable  $w_j$  surrounds in a counterclockwise direction the singularities at  $z_k$  ( $1 \leq k \leq a_j$ ), but does not surround the singularities at  $q^{-2}z_k$  ( $a_j + 1 \leq k \leq 2n$ ).

We will be mostly interested in  $\Phi_{\mathbf{a}}$  for  $\mathbf{a} \in \text{NC}_n^+$ —the set of such polynomials will turn out to be a basis for  $W_n[\mathbf{z}]$ —but the case when  $\mathbf{a}$  is in the larger set  $\mathcal{A}_n$  will play a useful intermediate role.

The above definition of  $\Phi_{\mathbf{a}}$  is somewhat difficult to work with. In particular, it is not even obvious that the  $\Phi_{\mathbf{a}}$ 's are polynomials. However, the contour integrals can be evaluated using the residue formula, which gives a more concrete representation of the  $\Phi_{\mathbf{a}}$ 's that makes it possible to show that they are in fact wheel polynomials and prove additional properties.

**Lemma 2.14.** 1. *The functions  $\Phi_{\mathbf{a}}$  can be written explicitly as follows.*

$$\begin{aligned}
\Phi_{\mathbf{a}}(\mathbf{z}) &= (-1)^{\binom{n}{2}} \prod_{1 \leq j < k \leq 2n} (qz_j - q^{-1}z_k) \\
&\quad \times \sum_{\substack{(m_1, \dots, m_n) \\ 1 \leq m_j \leq a_j, m_j \neq m_k}} \frac{\prod_{1 \leq j < k \leq n} (z_{m_k} - z_{m_j})(qz_{m_j} - q^{-1}z_{m_k})}{\prod_{j=1}^n \left( \prod_{\substack{k=1 \\ k \neq a_j}}^{a_j} (z_{m_j} - z_k) \prod_{k=a_j+1}^{2n} (qz_{m_j} - q^{-1}z_k) \right)} \\
&= (-1)^{\binom{n}{2}} \\
&\quad \times \sum_{\substack{M = (m_1, \dots, m_n) \\ 1 \leq m_j \leq a_j, m_j \neq m_k}} (-1)^{\text{inv}(M)} \frac{\prod_{1 \leq j < k \leq n} (qz_{m_j} - q^{-1}z_{m_k}) \prod_{\substack{1 \leq j < k \leq 2n \\ j \notin M \text{ or } j = m_p, k \leq a_p}} (qz_j - q^{-1}z_k)}{\prod_{j=1}^n \prod_{\substack{1 \leq k \leq a_j \\ k \notin M \text{ or } k > m_j}} (z_{m_j} - z_k)},
\end{aligned}$$

where for a sequence  $M = (m_1, \dots, m_n)$ ,  $\text{inv}(M)$  denotes the number of inversions of  $M$ , given by  $\text{inv}(M) = \#\{1 \leq i < j \leq n : m_i > m_j\}$ .

2.  $\Phi_{\mathbf{a}}$  is a wheel polynomial of order  $n$ .

*Proof.* See [38, Section 3]. □

For each sequence  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}_n$ , let  $(C_{\mathbf{a}, \sigma})_{\sigma \in \text{NC}_n}$  be coefficients such that

$$\Phi_{\mathbf{a}} = \sum_{\sigma \in \text{NC}_n} C_{\mathbf{a}, \sigma} \Psi_{\sigma}. \quad (30)$$

For  $\pi \in \text{NC}_n$  denote  $C_{\pi, \sigma} = C_{\mathbf{a}, \sigma}$  with  $\mathbf{a} = \pi^+$ . The coefficients  $(C_{\pi, \sigma})_{\sigma \in \text{NC}_n}$  have the property that

$$\Phi_{\pi} = \sum_{\sigma \in \text{NC}_n} C_{\pi, \sigma} \Psi_{\sigma}. \quad (31)$$

**Lemma 2.15.** *We have the relations*

$$\begin{aligned}\Phi_{\mathbf{a}}(\sigma) &= C_{\mathbf{a},\sigma} & (\mathbf{a} \in \mathcal{A}_n), \\ \Phi_{\pi}(\sigma) &= C_{\pi,\sigma} & (\pi \in \text{NC}_n).\end{aligned}$$

*Proof.* This follows immediately from (27) by considering the  $\pi$ -evaluation of both sides of (30)–(31).  $\square$

The next result is a recurrence relation that makes it possible to recursively compute the quantities  $\Phi_{\mathbf{a}}(\sigma) = C_{\mathbf{a},\sigma}$ , analogously to Theorem 2.8.

**Theorem 2.16.** *If  $\sigma \in \text{NC}_n$  has a “little arc”  $j \xrightarrow{\sigma} j+1$  and  $\mathbf{a} \in \mathcal{A}_n$ , let  $p$  be the number of times  $j$  appears in  $\mathbf{a}$ , so that we can write*

$$\mathbf{a} = (a_1, \dots, a_k, \overbrace{j, \dots, j}^{p \text{ times}}, a_{k+p+1}, \dots, a_n)$$

where  $a_k < j < a_{k+p+1}$ . In the case when  $p > 0$ , denote

$$\hat{\mathbf{a}}_j = (a_1, \dots, a_k, \overbrace{j, \dots, j}^{p-1 \text{ times}}, a_{k+p+1} - 2, \dots, a_n - 2). \quad (32)$$

Then we have

$$\Phi_{\mathbf{a}}(\sigma) = \begin{cases} 3^{n-1} \chi(p) \Phi_{\hat{\mathbf{a}}_j}(\hat{\sigma}_j) & \text{if } p > 0 \text{ (i.e., if } j \in \{a_1, \dots, a_n\}), \\ 0 & \text{if } p = 0, \end{cases} \quad (33)$$

where

$$\chi(p) = \frac{q^p - q^{-p}}{q - q^{-1}} = \begin{cases} 0 & \text{if } p \equiv 0 \pmod{3}, \\ 1 & \text{if } p \equiv 1 \pmod{3}, \\ -1 & \text{if } p \equiv 2 \pmod{3}. \end{cases} \quad (34)$$

*Proof.* See [38, Appendix A].  $\square$

Note that if  $\mathbf{a} = \pi^+ \in \text{NC}_n^+$  then  $\hat{\mathbf{a}}_j$  is in  $\mathcal{A}_{n-1}$  but not necessarily in  $\text{NC}_{n-1}^+$ , which is the reason why it was necessary to consider the family of polynomials indexed by the larger set  $\mathcal{A}_n$  even though we are mainly interested in the polynomials  $\Phi_{\pi}$ .

Like (26), the recurrence relation (33) can be solved to give an explicit formula for  $C_{\mathbf{a},\sigma}$ .

**Theorem 2.17.** For  $\mathbf{a} = (a_1, \dots, a_n) \in \text{NC}_n^+$  and  $\sigma \in \text{NC}_n$ , the coefficient  $C_{\mathbf{a},\sigma}$  is given by:

$$C_{\mathbf{a},\sigma} = \prod_{j < k, j \xleftrightarrow{\sigma} k} \chi(p_{j,k}(\mathbf{a})), \quad (35)$$

where we denote  $p_{j,k}(\mathbf{a}) = \#\{m : j \leq a_m < k\} - \frac{1}{2}(k - j - 1)$ .

*Proof.* See [38, Appendix A].  $\square$

**Theorem 2.18.** The matrix  $(C_{\pi,\sigma})_{\pi,\sigma \in \text{NC}_n}$  is triangular with respect to the partial order  $\leq$  on link patterns, and has 1s on the diagonal (i.e.,  $C_{\pi,\pi} = 1$  for all  $\pi \in \text{NC}_n$ ), hence it is invertible.

*Proof.* See [12, Sec. 3.3.2].  $\square$

**Corollary 2.19.**  $(\Phi_\pi)_{\pi \in \text{NC}_n}$  form a basis for  $W_n[\mathbf{z}]$ .

Denote the coefficients of the inverse matrix of  $(C_{\pi,\sigma})_{\pi,\sigma \in \text{NC}_n}$  by  $\tilde{C}_{\pi,\sigma}$ , or  $\tilde{C}_{\pi,\mathbf{a}}$  if  $\mathbf{a} = \sigma^+$ . In other words, the coefficients  $\tilde{C}_{\pi,\sigma}$  are defined by the relation

$$\Psi_\pi = \sum_{\sigma \in \text{NC}_n} \tilde{C}_{\pi,\sigma} \Phi_\sigma.$$

We have defined two square matrices  $\mathbf{C}_n = (C_{\pi,\sigma})_{\pi,\sigma \in \text{NC}_n}$  and  $\tilde{\mathbf{C}}_n = \mathbf{C}_n^{-1} = (\tilde{C}_{\pi,\sigma})_{\pi,\sigma \in \text{NC}_n}$ , which are the change of basis matrices transitioning between the two wheel polynomial bases we defined. Note that  $\tilde{C}_{\pi,\sigma}$ , like  $C_{\pi,\sigma}$ , are integers, but there does not seem to be a simple formula for them. The matrices  $\mathbf{C}_n$  and  $\tilde{\mathbf{C}}_n$  for  $n = 2, 3, 4$  are shown in Table 4. As we shall see, these matrices will enable us to translate certain sums of the qKZ polynomials (which are associated with certain connectivity events in loop percolation) to linear combinations of the  $\Phi_\pi$ 's. To get back to numerical results about the numbers  $\psi_\pi = \Psi_\pi(\mathbf{1})$  (which are related to the probabilities  $\mu_\pi$  according to Theorem 2.11), we will want to consider also the  $\mathbf{1}$ -evaluations of the  $\Phi_\pi$ 's. Denote  $\phi_{\mathbf{a}} = \Phi_{\mathbf{a}}(\mathbf{1})$ ,  $\phi_\pi = \Phi_\pi(\mathbf{1})$ . The coefficients  $\phi_\pi$  are related by

$$\begin{aligned} \phi_\pi &= \sum_{\sigma \in \text{NC}_n} C_{\pi,\sigma} \psi_\sigma, \\ \psi_\pi &= \sum_{\sigma \in \text{NC}_n} \tilde{C}_{\pi,\sigma} \phi_\sigma. \end{aligned}$$

| $n$ | $\mathbf{C}_n$  | $\tilde{\mathbf{C}}_n$   |
|-----|---|--|
| 2   | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  | $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$   |
| 3   | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$   | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$   |
| 4   | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ |

Table 4: The matrices  $\mathbf{C}_n, \tilde{\mathbf{C}}_n$  for  $n = 2, 3, 4$  (the ordering of the rows and columns corresponds to decreasing lexicographic order of the Young diagrams associated with noncrossing matchings). Note that  $\tilde{\mathbf{C}}_n$  can have entries larger in magnitude than 1; the smallest value of  $n$  for which this happens is  $n = 7$ .

**Lemma 2.20** ([38], Section 3.4).  $\phi_{\mathbf{a}}$  has the following contour integral formula:

$$\phi_{\mathbf{a}} = \oint \dots \oint \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=1}^n \frac{1}{u_j^{a_j}} \cdot \prod_{j=1}^n \frac{du_j}{2\pi i}, \quad (36)$$

where the contours are circles of arbitrary radius around 0. Equivalently,  $\phi_{\mathbf{a}}$  can be written as a constant term

$$\phi_{\mathbf{a}} = \text{CT}_{u_1, \dots, u_n} \left( \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=1}^n \frac{1}{u_j^{a_j-1}} \right) \quad (37)$$

or as a Taylor coefficient

$$\phi_{\mathbf{a}} = [u_1^{a_1-1} \dots u_n^{a_n-1}] \left( \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \right).$$

*Proof.* In the integral formula (29) set  $z_1 = \dots = z_{2n} = 1$  and then make the substitution  $w_j = \frac{1-q^{-1}u_j}{1-qu_j}$ . This gives (36) after a short computation.  $\square$

## 2.2 Product-type expansions and the sum rule

We are ready to start applying the theory developed above to the problem of computing probabilities of connectivity events for the random non-crossing matching  $\Pi_*^{(n)}$ . Our methods generalize those of Zinn-Justin and Di-Francesco [38], so let us recall briefly one of their beautiful discoveries. Define a set  $\mathcal{B}_n \subseteq \text{NC}_n^+$  consisting of  $2^{n-1}$  sequences, given by

$$\mathcal{B}_n = \{\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}_n : a_1 = 1, a_j \in \{2j-2, 2j-1\}, (2 \leq j \leq n)\}.$$

For each  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{B}_n$  define

$$\mathcal{L}(\mathbf{a}) = \{\pi \in \text{NC}_n : \text{for all } 1 \leq j \leq n, \pi_{2j-1} = 1 \text{ iff } a_j = 2j-1\}.$$

Note that  $\text{NC}_n = \bigsqcup_{\mathbf{a} \in \mathcal{B}_n} \mathcal{L}(\mathbf{a})$ .

**Theorem 2.21.** *For any  $\mathbf{a} \in \mathcal{B}_n$ , we have the expansion*

$$\Phi_{\mathbf{a}} = \sum_{\pi \in \mathcal{L}(\mathbf{a})} \Psi_{\pi}. \quad (38)$$

Theorem 2.21 and its elegant proof served as the inspiration for one of our main results (Theorem 2.24 in Subsection 2.4). As an aid in motivating and understanding our use of Zinn-Justin and Di Francesco's proof technique in a more complicated setting, we include their original proof, taken from [38, Section 3], below.

*Proof of Theorem 2.21.* By Theorem 2.3, it is enough to prove that for any  $\sigma \in \text{NC}_n$  we have

$$\Phi_{\mathbf{a}}(\sigma) = \sum_{\pi \in \mathcal{L}(\mathbf{a})} \Psi_{\pi}(\sigma). \quad (39)$$

The proof is by induction on  $n$ . Denote the right-hand side of (39) by  $\theta_{\mathbf{a}}(\sigma)$ . Let  $k \xrightarrow{\sigma} k+1$  be a little arc of  $\sigma$ . Let  $j$  be such that  $k \in \{2j-2, 2j-1\}$ . By the definition of  $\mathcal{B}_n$ , either  $k = a_j$  is the only occurrence of  $k$  in  $\mathbf{a}$ , or  $k$

does not occur in  $\mathbf{a}$  (in which case  $a_j$  is the number in  $\{2j-2, 2j-1\}$  of the opposite parity from  $k$ ). For the left-hand side of (39) we have by (33) that

$$\Phi_{\mathbf{a}}(\sigma) = \begin{cases} 3^{n-1}\Phi_{\hat{\mathbf{a}}_k}(\hat{\sigma}_k) & \text{if } a_j = k, \\ 0 & \text{if } a_j \neq k. \end{cases}$$

From the definition of  $\mathcal{B}_n$  it is easy to see that when  $k = a_j$ ,  $\hat{\mathbf{a}}_k \in \mathcal{B}_{n-1}$ . So, to complete the inductive proof it is enough to show that  $\theta_{\mathbf{a}}(\sigma)$  satisfies the same recurrence. First, assume that  $a_j \neq k$ . If  $k = 2j-1$  (so  $a_j = 2j-2$ ) then any  $\pi \in \mathcal{L}(\mathbf{a})$  will have the property that  $\pi_k = -1$ , that is,  $\pi$  matches  $k$  with a number to the left of  $k$ . In particular  $\pi$  does not contain the little arc  $(k, k+1)$ , and therefore  $\Psi_{\pi}(\sigma) = 0$ . This shows that  $\theta_{\mathbf{a}}(\sigma) = 0$ . Alternatively, if  $k = 2j-2, a_j = 2j-1$ , then any  $\pi \in \mathcal{L}(\mathbf{a})$  will have the property that  $\pi$  matches  $a_j = k+1$  with a number to the right of  $k+1$ , and again the little arc  $(k, k+1)$  cannot be in  $\pi$  and  $\Psi_{\pi}(\sigma) = 0$ , so the relation  $\theta_{\mathbf{a}}(\sigma) = 0$  also holds.

Finally, in the case  $a_j = k$ , the noncrossing matchings  $\pi \in \mathcal{L}(\mathbf{a})$  can be divided into those for which  $(k, k+1)$  is not a little arc of  $\pi$ , which satisfy  $\Psi_{\pi}(\sigma) = 0$ , and those for which  $(k, k+1)$  is a little arc of  $\pi$ . The ones in the latter class satisfy  $\Psi_{\pi}(\sigma) = 3^{n-1}\Psi_{\hat{\pi}_k}(\hat{\sigma}_k)$ , and furthermore the correspondence  $\pi \mapsto \hat{\pi}_k$  maps them bijectively onto  $\mathcal{L}(\hat{\mathbf{a}}_k)$ . Thus we get the relation

$$\theta_{\mathbf{a}}(\sigma) = 3^{n-1}\theta_{\hat{\mathbf{a}}_k}(\hat{\sigma}_k),$$

which is the same recurrence as the one for  $\Phi_{\mathbf{a}}(\sigma)$  and therefore completes the proof.  $\square$

Theorem 2.21 has several immediate corollaries that are directly relevant to the problem of deriving formulas for probabilities of connectivity events. First, summing (38) over the sequences  $\mathbf{a} \in \mathcal{B}_n$  gives that

$$\sum_{\pi \in \text{NC}_n} \Psi_{\pi} = \sum_{\mathbf{a} \in \mathcal{B}_n} \Phi_{\mathbf{a}}.$$

Taking the  $\mathbf{1}$ -evaluation of both sides then gives the relation

$$\sum_{\pi \in \text{NC}_n} \psi_{\pi} = \sum_{\mathbf{a} \in \mathcal{B}_n} \phi_{\mathbf{a}},$$



and the right-hand side can then be expanded as a constant term using (37), which gives that

$$\sum_{\pi \in \text{NC}_n} \psi_\pi = \text{CT}_{u_1, \dots, u_n} \left( \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=2}^n \frac{1 + u_j}{u_j^{2j-2}} \right). \quad (40)$$

Note that the key point that makes such an expansion possible is the fact that  $\mathcal{B}_n$  decomposes as a Cartesian product, which manifests itself as the product  $\prod_{j=2}^n (1 + u_j)/u_j^{2j-2}$  in the Laurent polynomial. It is precisely this product structure that our more general result seeks to emulate. Finally, applying the identity (16) gives the following result.

**Theorem 2.22.**  $\sum_{\pi \in \text{NC}_n} \psi_\pi = \text{ASM}(n)$ .

With this result, we are finally able to relate results about the vector  $\psi_n$  to statements about its normalized version, the probability vector  $\mu_n$ . As an example (also taken from [38]), letting  $\mathbf{a} = (1, 3, 5, \dots, 2n-1)$ , note that  $\mathcal{L}(\mathbf{a})$  is the singleton set  $\{\pi_{\max}^n\}$ , so, applying (38) in this case, then taking the  $\mathbf{1}$ -evaluation and using (37) as before, gives the identity

$$\psi_{\pi_{\max}^n} = \text{CT}_{u_1, \dots, u_n} \left( \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=1}^n \frac{1}{u_j^{2j-2}} \right).$$

This constant term can be evaluated iteratively as

$$\begin{aligned} & \text{CT}_{u_1, \dots, u_{n-1}} \text{CT}_{u_n} \left( \prod_{1 \leq j < k \leq n} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=1}^n \frac{1}{u_j^{2j-2}} \right) \\ &= \text{CT}_{u_1, \dots, u_{n-1}} \left( \prod_{1 \leq j < k \leq n-1} (u_k - u_j)(1 + u_k + u_j u_k) \prod_{j=1}^n \frac{1 + u_j}{u_j^{2j-2}} \right) \\ &= \text{ASM}(n-1) \end{aligned}$$

(use (16), noting that on the right-hand side of that identity, starting the product  $\prod_j (1 + z_j)$  at  $j = 1$  instead of  $j = 2$  does not affect the constant term), confirming part 3 of Theorem 1.9.

### 2.3 $p$ -nested matchings

If  $\pi \in \text{NC}_n$ , denote by  $\widehat{\pi}$  the noncrossing matching of order  $n+1$  obtained from  $\pi$  by relabeling the elements of  $[1, 2n]$  as  $2, \dots, 2n+1$  and adding a

“large arc” connecting 1 and  $2n + 2$ . We call this the **nesting of  $\pi$**  or “ $\pi$ -nested.” For an integer  $p \geq 0$  let  $\widehat{\pi}^p$  denote the  $p$ -nesting of  $\pi$ , defined as the result of  $p$  successive nesting operations applied to  $\pi$ . Similarly, for the associated sequence  $\mathbf{a} = \pi^+ \in \text{NC}_n^+$ , denote  $\widehat{\mathbf{a}} = (\widehat{\pi})^+$  and  $\widehat{\mathbf{a}}^p = (\widehat{\pi}^p)^+$ . If  $\mathbf{a} = (a_1, \dots, a_n)$ , it is easy to see that this can be written more explicitly as

$$\widehat{\mathbf{a}}^p = (1, \dots, p, p + a_1, \dots, p + a_n).$$

The following result is due to Fonseca and Zinn-Justin [12].

**Theorem 2.23.** *Let  $p \geq 0$ . Under the  $p$ -nesting operation, the expansions*

$$\begin{aligned} \Phi_\sigma &= \sum_{\pi \in \text{NC}_n} C_{\sigma, \pi} \Phi_\pi & (\pi \in \text{NC}_n), \\ \Psi_\pi &= \sum_{\sigma \in \text{NC}_n} \tilde{C}_{\pi, \sigma} \Phi_\sigma & (\sigma \in \text{NC}_n), \end{aligned}$$

*turn into*

$$\begin{aligned} \Phi_{\widehat{\sigma}^p} &= \sum_{\pi \in \text{NC}_n} C_{\sigma, \pi} \Psi_{\widehat{\pi}^p} & (\pi \in \text{NC}_n), \\ \Psi_{\widehat{\pi}^p} &= \sum_{\sigma \in \text{NC}_n} \tilde{C}_{\pi, \sigma} \Phi_{\widehat{\sigma}^p} & (\sigma \in \text{NC}_n). \end{aligned}$$

*Equivalently, we have the recursive relations*

$$C_{\widehat{\sigma}^p, \mu} = \begin{cases} C_{\sigma, \pi} & \text{if } \mu = \widehat{\pi}^p, \\ 0 & \text{otherwise,} \end{cases} \quad (41)$$

$$\tilde{C}_{\widehat{\pi}^p, \mu} = \begin{cases} \tilde{C}_{\pi, \sigma} & \text{if } \mu = \widehat{\sigma}^p, \\ 0 & \text{otherwise.} \end{cases} \quad (42)$$

*Proof.* The relation (42) is an immediate consequence of (41). To prove (41), note that the fact that  $C_{\widehat{\sigma}^p, \mu} = 0$  if  $\mu$  is not of the form  $\widehat{\sigma}^p$  for some  $\sigma \in \text{NC}_n$  follows from the triangularity of the basis change matrix with respect to the order  $\leq$ . The relation  $C_{\widehat{\sigma}^p, \widehat{\pi}^p} = C_{\sigma, \pi}$  is easy to prove by induction from (33).  $\square$

## 2.4 General submatching events

Fix  $k \geq 0$  and  $\pi_0 \in \text{NC}_k$ . For  $n \geq k+1$  define

$$\mathcal{E}_n(\pi_0) = \{\pi \in \text{NC}_n : \pi(j) = 1 + \pi_0(j-1) \text{ for } 2 \leq j \leq k+1\}.$$

In the terminology of Subsections 1.3 and 1.5, this set of matchings is related to the submatching event associated with the submatching  $\pi_0$ , except that we require the submatching to occur on the sites in the interval  $[2, k+1]$  instead of  $[1, k]$ . Surprisingly, this turns out to be the correct thing to do—see below.

Next, for any  $n \geq k+1$  and  $\mathbf{b} = (b_{k+2}, \dots, b_n)$  satisfying  $b_j \in \{2j-2, 2j-1\}$  for all  $j$ , define

$$\mathcal{E}_n(\pi_0, \mathbf{b}) = \left\{ \pi \in \mathcal{E}_n(\pi_0) : \pi_{2j-1} = 1 \text{ iff } b_j = 2j-1 \text{ } (k+2 \leq j \leq n) \right\}.$$

Note that

$$\mathcal{E}_n(\pi_0) = \bigsqcup_{\substack{\mathbf{b}=(b_{k+2}, \dots, b_n) \\ \forall j \ b_j \in \{2j-2, 2j-1\}}} \mathcal{E}_n(\pi_0, \mathbf{b}).$$

The following result will easily imply Theorem 1.13, and is one of the main results of this paper.

**Theorem 2.24.** *For any  $\pi_0 \in \text{NC}_k$ ,  $n \geq k+1$  and vector  $\mathbf{b} = (b_{k+2}, \dots, b_n)$  satisfying  $b_j \in \{2j-2, 2j-1\}$  for all  $j$ , we have*

$$\sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})} \Psi_\pi = \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \Phi_{(1, 1+a_1, \dots, 1+a_k, b_{k+2}, \dots, b_n)}. \quad (43)$$

It follows that for any  $\pi_0 \in \text{NC}_k$  and  $n \geq k+1$ ,

$$\sum_{\pi \in \mathcal{E}_n(\pi_0)} \Psi_\pi = \sum_{\substack{b_{k+2}, \dots, b_n \\ \forall j \ b_j \in \{2j-2, 2j-1\}}} \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \Phi_{(1, 1+a_1, \dots, 1+a_k, b_{k+2}, \dots, b_n)}. \quad (44)$$

*Proof.* By (42), the right-hand side of (43) is equal to

$$\sum_{\mathbf{a}=(a_1, \dots, a_{k+1}) \in \text{NC}_{k+1}^+} \tilde{C}_{\pi_0, \mathbf{a}} \Phi_{(a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n)}.$$

So, instead of proving (43) we will show that

$$\sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})} \Psi_\pi = \sum_{\mathbf{a}=(a_1, \dots, a_{k+1}) \in \text{NC}_{k+1}^+} \tilde{C}_{\overline{\pi_0}, \mathbf{a}} \Phi_{(a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n)}. \quad (45)$$

We will prove by induction on  $n$  the following slightly stronger claim: if we have an expansion of the form

$$\Psi_{\overline{\pi_0}} = \sum_{\mathbf{a} \in \mathcal{A}_{k+1}} c_{\mathbf{a}} \Phi_{\mathbf{a}} \quad (46)$$

for some constants  $c_{\mathbf{a}}$  ( $\mathbf{a} \in \mathcal{A}_{k+1}$ ), then for any vector  $\mathbf{b} = (b_{k+2}, \dots, b_n)$  satisfying  $b_j \in \{2j-2, 2j-1\}$  as in the theorem, the equality

$$\sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})} \Psi_\pi = \sum_{\mathbf{a} \in \mathcal{A}_{k+1}} c_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})} \quad (47)$$

holds. (Note that this claim implies (45) by taking  $c_{\mathbf{a}} = \tilde{C}_{\overline{\pi_0}, \mathbf{a}}$  for  $\mathbf{a} \in \text{NC}_n^+$  and  $c_{\mathbf{a}} = 0$  otherwise, but allowing the sum on the right-hand side of (46) to range over the bigger set  $\mathcal{A}_{k+1}$  helps make the induction work.) Denote the left-hand side of (47) by  $X_{\pi_0, \mathbf{b}}$  and the right-hand side by  $Y_{c, \mathbf{b}}$ . We now prove that  $X_{\pi_0, \mathbf{b}} = Y_{c, \mathbf{b}}$ .

**The induction base:** here  $k = 0$ ,  $n = 1$ ,  $\pi_0 = \emptyset$  (the empty matching),  $\mathcal{A}_{k+1} = \mathcal{A}_1 = \{1\}$  and the expansion necessarily takes the form

$$\Psi_{\overline{\emptyset}} = \Phi_1.$$

Also  $\mathcal{E}_n(\pi_0, \mathbf{b}) = \mathcal{L}((1, b_2, \dots, b_n))$ , so the claim reduces to the relation (38) from Theorem 2.21.

**The inductive step:** First, if  $n = k + 1$  then the claim (47) is the same as the assumption (46), since  $\mathbf{b}$  is a trivial vector of length 0 and  $\mathcal{E}_n(\pi_0, \mathbf{b})$  is the singleton set  $\{\overline{\pi_0}\}$ , so there is nothing to prove. From now on assume  $n > k + 1$ . We need to show that for all  $\sigma \in \text{NC}_n$ ,  $X_{\pi_0, \mathbf{b}}(\sigma) = Y_{c, \mathbf{b}}(\sigma)$ . Let  $1 \leq m \leq 2n - 1$  be a number such that  $m \xleftrightarrow{\sigma} m + 1$ .

**Case 1.** Assume  $m \geq 2k + 2$ .

**Subcase 1a.** Assume  $m \notin \mathbf{b}$  (that is,  $m \neq b_j$  for all  $j$ ). Then by (33) we have that  $\Phi_{a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n}(\sigma) = 0$  for any  $\mathbf{a} = (a_1, \dots, a_{k+1}) \in \mathcal{A}_{k+1}$ , so  $Y_{c, \mathbf{b}}(\sigma) = 0$ .

**Subsubcase 1a(i).** Assume  $m$  is odd, i.e.,  $m = 2j - 1$  for some  $k + 2 \leq j \leq n$ . If  $\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})$  then, since we assumed  $m \notin \mathbf{b}$ ,  $b_j = 2j - 2$  and

therefore by definition  $\pi_m = -1$ , that is,  $\pi(m) = \pi(2j - 1) < m$ , so in particular  $(m, m + 1)$  is not a little arc of  $\pi$  and therefore  $\Psi_\pi(\sigma) = 0$ . It follows that  $X_{\pi_0, \mathbf{b}}(\sigma) = 0 = Y_{c, \mathbf{b}}(\sigma)$ .

**Subsubcase 1a(ii).** Assume  $m$  is even, i.e.,  $m = 2j - 2$  for some  $k + 2 \leq j \leq n$ . If  $\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})$  then, since we assumed  $m \notin \mathbf{b}$ ,  $b_j = 2j - 1$  and therefore by definition  $\pi(m + 1) = \pi(2j - 1) > m + 1$  and in particular  $(m, m + 1)$  is not a little arc of  $\pi$ . Again it follows that  $\Psi_\pi(\sigma) = 0$  and therefore  $X_{\pi_0, \mathbf{b}}(\sigma) = 0 = Y_{c, \mathbf{b}}(\sigma)$ .

**Subcase 1b.** Assume that  $m \in \mathbf{b}$ , i.e.,  $m = b_j$  for some  $k + 2 \leq j \leq n$ . Then by (33), for any  $\mathbf{a} = (a_1, \dots, a_{k+1}) \in \mathcal{A}_{k+1}$  we have

$$\Phi_{(a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n)}(\sigma) = 3^{n-1} \Phi_{(\mathbf{a}, \hat{\mathbf{b}}_m)}(\hat{\sigma}_m),$$

where  $\hat{\mathbf{b}}_m$  is obtained from  $\mathbf{b}$  by deleting the  $m$  and shifting all the numbers to the right of  $m$  down by 2. It follows that

$$Y_{c, \mathbf{b}}(\sigma) = 3^{n-1} Y_{c, \hat{\mathbf{b}}_m}(\hat{\sigma}_m). \quad (48)$$

On the other hand, we have

$$\begin{aligned} X_{\pi_0, \mathbf{b}}(\sigma) &= \sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b}), \ m \xleftrightarrow{\pi} m+1} \Psi_\pi(\sigma) + \sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b}), \ m \xleftrightarrow{\pi} m+1} \Psi_\pi(\sigma) \\ &= 0 + \sum_{\pi \in \mathcal{E}_n(\pi_0, \mathbf{b}), \ m \xleftrightarrow{\pi} m+1} 3^{n-1} \Psi_{\hat{\pi}_m}(\hat{\sigma}_m). \end{aligned}$$

Relabeling  $\hat{\pi}_m$  in the last sum as  $\pi$  transforms it into

$$3^{n-1} \sum_{\pi \in \mathcal{E}_{n-1}(\pi_0, \hat{\mathbf{b}}_m)} \Psi_\pi(\hat{\sigma}_m) = 3^{n-1} X_{\pi_0, \hat{\mathbf{b}}_m}(\hat{\sigma}_m),$$

so by induction we get using (48) that  $X_{\pi_0, \mathbf{b}}(\sigma) = Y_{\pi_0, \mathbf{b}}(\sigma)$ .

**Case 2.** Assume  $m \leq 2k + 1$ .

**Subcase 2(a).** Assume  $m \xleftrightarrow{\pi_0} m + 1$  (when thinking of  $\pi_0$  as “living” on the interval  $[2, 2k + 1]$ ), and note that in this case actually  $m$  must be  $\leq 2k$ . Then for any  $\pi \in \mathcal{E}_n(\pi_0, \mathbf{b})$  we have

$$\Psi_\pi(\sigma) = 3^{n-1} \Psi_{\hat{\pi}_m}(\hat{\sigma}_m),$$

and consequently, denoting  $\mathbf{b}' = (b_{k+2} - 2, \dots, b_n - 2)$ , we see that

$$X_{\pi_0, \mathbf{b}}(\sigma) = 3^{n-1} X_{(\widehat{\pi_0})_m, \mathbf{b}'}(\hat{\sigma}_m).$$

To evaluate  $Y_{c,\mathbf{b}}(\sigma)$ , we use (33) to get that

$$\begin{aligned}
Y_{c,\mathbf{b}}(\sigma) &= \sum_{\mathbf{a} \in \mathcal{A}_{k+1}} c_{\mathbf{a}} \Phi_{(\mathbf{a},\mathbf{b})}(\sigma) = \sum_{\mathbf{a} \in \mathcal{A}_{k+1}, m \in \mathbf{a}} c_{\mathbf{a}} \Phi_{(\mathbf{a},\mathbf{b})}(\sigma) \\
&= \sum_{\mathbf{a} \in \mathcal{A}_{k+1}, m \in \mathbf{a}} 3^{n-1} \chi(p(\mathbf{a}, m)) c_{\mathbf{a}} \Phi_{(\hat{\mathbf{a}}_m, \mathbf{b}')}(\hat{\sigma}_m) \\
&= \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')}(\hat{\sigma}_m),
\end{aligned} \tag{49}$$

where  $p(\mathbf{a}, m)$  denotes the number of times  $m$  appears in  $\mathbf{a}$ ,  $\hat{\mathbf{a}}_m$  is as in (32), and in the last line the sum was reorganized as a linear combination with some new coefficients  $d_{\mathbf{a}'}$ . (The fact that this sum ranges over  $\mathcal{A}_k$  explains why we needed to strengthen the inductive hypothesis at the beginning of the proof.)

Note also that in a similar way, replacing the vector  $\mathbf{b}$  with the empty vector  $\emptyset$  of length 0, we have for any  $\mu \in \text{NC}_{k+1}$  such that  $m \xleftrightarrow{\mu} m+1$ ,

$$\begin{aligned}
Y_{c,\emptyset}(\mu) &= \sum_{\mathbf{a} \in \mathcal{A}_{k+1}} c_{\mathbf{a}} \Phi_{\mathbf{a}}(\mu) = \sum_{\mathbf{a} \in \mathcal{A}_{k+1}, m \in \mathbf{a}} c_{\mathbf{a}} \Phi_{\mathbf{a}}(\mu) \\
&= \sum_{\mathbf{a} \in \mathcal{A}_{k+1}, m \in \mathbf{a}} 3^{n-1} \chi(p(\mathbf{a}, m)) c_{\mathbf{a}} \Phi_{\hat{\mathbf{a}}_m}(\hat{\mu}_m) \\
&= \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{\mathbf{a}'}(\hat{\mu}_m),
\end{aligned}$$

with exactly the same coefficients  $d_{\mathbf{a}'}$  (the main thing to notice is that appending the  $\mathbf{b}$  at the end in the subscript of  $\Phi$  has no effect on the way the recurrence (33) proceeds, i.e., the  $\mathbf{b}$  “doesn’t interact” with the  $\mathbf{a}$ ). On the other hand, the case  $n = k+1$  discussed at the beginning of the proof gives that  $Y_{c,\emptyset}(\mu) = X_{\pi_0,\emptyset}(\mu) = 3^{n-1} X_{(\pi_0)_m,\emptyset}(\hat{\mu}_m)$ . Combining this with (49), we get that

$$X_{(\pi_0)_m,\emptyset}(\nu) = 3^{-(n-1)} \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{\mathbf{a}'}(\nu)$$

for any  $\nu \in \text{NC}_k$ , or in other words simply that

$$X_{(\pi_0)_m,\emptyset} = 3^{-(n-1)} \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{\mathbf{a}'}.$$

Invoking the inductive hypothesis again, we conclude that

$$X_{(\pi_0)_m,\mathbf{b}'} = 3^{-(n-1)} \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')}.$$

In particular,

$$3^{(n-1)} X_{(\pi_0)_m, \mathbf{b}'}(\hat{\sigma}_m) = \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')}(\hat{\sigma}_m).$$

But as we showed above, the LHS is equal to  $X_{\pi_0, \mathbf{b}}(\sigma)$  and the RHS is equal to  $Y_{c, \mathbf{b}}(\sigma)$ , so this proves the claim.

**Subcase 2(b).** Assume  $m \xleftrightarrow{\pi_0} m+1$ . Then for a similar reason as in subcase 2(a) above,  $X_{\pi_0, \mathbf{b}}(\sigma) = 0$ . We need to show that also  $Y_{c, \mathbf{b}}(\sigma) = 0$ . The computation from subcase 2(a) is still valid, so we can write as before  $Y_{c, \mathbf{b}}(\sigma) = \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')}(\hat{\sigma}_m)$ , and also

$$Y_{c, \emptyset}(\mu) = \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{\mathbf{a}'}(\hat{\mu}_m)$$

for any  $\mu \in \text{NC}_{k+1}$  for which  $m \xleftrightarrow{\mu} m+1$ . But we know (again from the case  $n = k+1$ ) that  $Y_{c, \emptyset}(\mu) = X_{\pi_0, \emptyset}(\mu) = 0$ , so we deduce that

$$\sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{\mathbf{a}'} = 0.$$

We claim that this implies also that

$$\sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')} = 0, \tag{50}$$

which, if true, would also give in particular that

$$Y_{c, \mathbf{b}}(\sigma) = \sum_{\mathbf{a}' \in \mathcal{A}_k} d_{\mathbf{a}'} \Phi_{(\mathbf{a}', \mathbf{b}')}(\hat{\sigma}_m) = 0,$$

and finish the proof.

It remains to prove (50). First, for convenience relabel  $\mathbf{a}'$  as  $\mathbf{a}$ ,  $\mathbf{b}'$  as  $\mathbf{b}$  and  $n-1$  as  $n$ . The claim then becomes that if  $\mathbf{b} = (b_{k+1}, \dots, b_n)$  satisfies  $b_j \in \{2j-2, 2j-1\}$  for all  $j$  and  $(d_{\mathbf{a}})_{\mathbf{a} \in \mathcal{A}_k}$  are coefficients such that  $\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{\mathbf{a}} = 0$  then  $\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})} = 0$ . The proof is by induction on  $n$ ; the idea as before is that the  $\mathbf{a}'$  and  $\mathbf{b}'$  components “don’t interact” when we perform recursive computations using (33), and the proof involves similar ideas to those used above. If  $n = k$  there is nothing to prove, so assume that  $n > k$ . Let  $\nu \in \text{NC}_n$ , and pick some  $g$  such that  $g \xleftrightarrow{\pi} g+1$ . Divide into two cases,

when  $g \geq 2k$  or when  $g \leq 2k - 1$ . We will show that in each case we have  $\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})}(\nu) = 0$ , which will imply the claim for the usual reason (Theorem 2.3).

In the first case, if  $g \notin \{b_{k+1}, \dots, b_n\}$  then  $\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})}(\nu) = 0$  by (33). Alternatively, if  $g \in \{b_{k+1}, \dots, b_n\}$  then, again by (33), we have that

$$\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})}(\nu) = 3^{n-1} \sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \bar{\mathbf{b}})}(\hat{\nu}_g), \quad (51)$$

where we denote  $\bar{\mathbf{b}} = (b_{k+1} - 2, \dots, b_n - 2)$  (note that the  $\chi(p)$  factors in (33) are all equal to 1 because of the structure of  $\mathbf{b}$ ). The right-hand side of (51) is equal to 0 by the inductive hypothesis (applied with the vector  $\bar{\mathbf{b}}$  replacing  $\mathbf{b}$ ).

Finally, in the case where  $g \leq 2k - 1$ , we have

$$\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{(\mathbf{a}, \mathbf{b})}(\nu) = \sum_{\mathbf{a} \in \mathcal{A}_k, g \in \{a_1, \dots, a_k\}} d_{\mathbf{a}} 3^{n-1} \chi(p(\mathbf{a}, g)) \Phi_{(\hat{\mathbf{a}}_g, \bar{\mathbf{b}})}(\hat{\nu}_g) \quad (52)$$

(where the notation  $p(\mathbf{a}, g)$  was defined immediately after (49)), and this can be rewritten as a sum of the form

$$\sum_{\mathbf{a}' \in \mathcal{A}_{k-1}} f_{\mathbf{a}'} \Phi_{(\mathbf{a}', \bar{\mathbf{b}})}(\hat{\nu}_g) \quad (53)$$

where  $(f_{\mathbf{a}'}_{\mathbf{a}' \in \mathcal{A}_{k-1}})$  are some coefficients. On the other hand, by repeating the same computation in the case  $n = k$  where  $\mathbf{b}$  is replaced by the empty vector of length 0, we see that for any  $\mu \in \text{NC}_k$  such that  $g \xleftrightarrow{\mu} g + 1$ , we have

$$\sum_{\mathbf{a} \in \mathcal{A}_k} d_{\mathbf{a}} \Phi_{\mathbf{a}}(\mu) = \sum_{\mathbf{a}' \in \mathcal{A}_{k-1}} f_{\mathbf{a}'} \Phi_{\mathbf{a}'}(\hat{\mu}_g), \quad (54)$$

where the key observation is that the coefficients are the same numbers  $f_{\mathbf{a}'}$  as in (53). But we know (by the assumption of the claim we are trying to prove) that the left-hand side of (54) is equal to 0. Since this is true for any  $\mu \in \text{NC}_k$  for which  $g \xleftrightarrow{\mu} g + 1$ , we conclude that  $\sum_{\mathbf{a}' \in \mathcal{A}_{k-1}} f_{\mathbf{a}'} \Phi_{\mathbf{a}'} = 0$ . The inductive hypothesis now implies that  $\sum_{\mathbf{a}' \in \mathcal{A}_{k-1}} f_{\mathbf{a}'} \Phi_{(\mathbf{a}', \bar{\mathbf{b}})} = 0$ . In particular, the expression in (53), which is equal to the left-hand side of (52), is 0.  $\square$



## 2.5 Proof of Theorem 1.13

We are now in a position to prove Theorem 1.13, with an explicit description of the multivariate polynomial  $F_{\pi_0}(w_1, \dots, w_k)$  whose existence is claimed in the theorem. Let  $\pi_0 \in \text{NC}_k$ . The polynomial  $F_{\pi_0}$  simply encodes the  $\pi_0$ th row of the matrix  $\tilde{C}_n$ ; more precisely, we define it as

$$F_{\pi_0}(w_1, \dots, w_k) = \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \prod_{j=1}^k w_j^{2j-a_j}. \quad (55)$$

Denote  $\Omega_n(\mathbf{z}) = \prod_{1 \leq i < j \leq n} (z_j - z_i)(1 + z_j + z_i z_j)$ . Taking the  $\mathbf{1}$ -evaluation of both sides of (44) and combining the resulting equation with Theorems 2.11 and 2.22, we get that

$$\begin{aligned} \text{ASM}(n) \mathbb{P}(\Pi_*^{(n)} \in \mathcal{E}_n(\pi_0)) &= \sum_{\pi \in \mathcal{E}_n(\pi_0)} \psi_\pi \\ &= \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \sum_{\substack{b_{k+2}, \dots, b_n \\ b_j \in \{2j-2, 2j-1\}}} \phi_{(1, 1+a_1, \dots, 1+a_k, b_{k+2}, \dots, b_n)}. \end{aligned}$$

The left-hand side is also equal to  $\text{ASM}(n) \mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)})$  by the rotational symmetry of  $\Pi_*^{(n)}$ . To evaluate the right-hand side, represent each summand  $\phi_{(1, 1+a_1, \dots, 1+a_k, b_{k+2}, \dots, b_n)}$  as a constant term using (37), to get that the last

expression can be written as

$$\begin{aligned}
& \text{CT}_{\mathbf{z}} \left[ \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \sum_{\substack{b_{k+2}, \dots, b_n \\ \forall j \ b_j \in \{2j-2, 2j-1\}}} \Omega_n(\mathbf{z}) \prod_{j=2}^{k+1} \frac{1}{z_j^{a_{j-1}}} \prod_{j=k+2}^n \frac{1}{z_j^{b_{j-1}}} \right] \\
&= \text{CT}_{\mathbf{z}} \left[ \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \Omega_n(\mathbf{z}) \prod_{j=2}^{k+1} \frac{1}{z_j^{a_{j-1}}} \prod_{j=k+2}^n \left( \frac{1}{z_j^{2j-3}} + \frac{1}{z_j^{2j-2}} \right) \right] \\
&= \text{CT}_{\mathbf{z}} \left[ \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \Omega_n(\mathbf{z}) \prod_{j=2}^{k+1} \frac{1}{z_j^{a_{j-1}}} \prod_{j=k+2}^n \frac{1+z_j}{z_j^{2j-2}} \right] \\
&= \text{CT}_{\mathbf{z}} \left[ \Omega_n(\mathbf{z}) \prod_{j=k+2}^n (1+z_j) \prod_{j=1}^n \frac{1}{z_j^{2j-2}} \left( \sum_{\mathbf{a}=(a_1, \dots, a_k) \in \text{NC}_k^+} \tilde{C}_{\pi_0, \mathbf{a}} \prod_{j=2}^{k+1} z_j^{2j-2-a_{j-1}} \right) \right] \\
&= \text{CT}_{\mathbf{z}} \left[ \Omega_n(\mathbf{z}) \prod_{j=k+2}^n (1+z_j) \prod_{j=1}^n \frac{1}{z_j^{2j-2}} \cdot F_{\pi_0}(z_2, \dots, z_{k+1}) \right],
\end{aligned}$$

which is clearly equal to  $\text{ASM}(n)$  times the right-hand side of (15).  $\square$

## Appendix A. Summary of algorithms

The research described in this paper has been significantly aided by computer-aided experimentation and numerical computations performed by the author. We believe use of such methods is likely to continue to play a role in the discovery of new results extending and building on this work. In the hope of stimulating such further research, we summarize the theory presented above in the form of explicit algorithms for computing quantities of interest related to connectivity patterns of loop percolation and the theory of wheel polynomials.

**Algorithm A** Computation of the probability vector  $\boldsymbol{\mu}_n = (\mu_\pi)_{\pi \in \text{NC}_n}$

**Step 1.** Compute the matrix  $S_n = (s_{\pi, \sigma})_{\pi, \sigma \in \text{NC}_n}$  where

$$s_{\pi, \sigma} = \#\{1 \leq j \leq 2n : e_j(\pi) = \sigma\}$$

(with  $e_j$  defined by the equation (4), interpreted modulo  $2n$ ).

**Step 2.** Compute  $\mu_n$  as the solution to the vector equation

$$\mu_n(S_n - 2n\mathbf{I}) = \mathbf{0}$$

(where  $\mathbf{I}$  is the identity matrix), normalized to be a probability vector.

**Algorithm B** Computation of the polynomials  $Q_{\pi_0}$  (assuming Conjecture 1.12)

**Step 1.** Use Algorithm A above to compute the submatching event probabilities  $p_n = \mathbb{P}(\pi_0 \triangleleft \Pi_*^{(n)})$  for  $n = k, k+1, \dots, k(k+3)/2$ .

**Step 2.** Compute the numbers  $q_n = p_n \prod_{j=1}^k (4n^2 - j^2)^{k+1-j}$ .

**Step 3.** Use the Lagrange interpolation formula to compute the unique polynomial  $G_{\pi_0}(m)$  of degree  $\leq k(k+1)/2$  satisfying

$$G_{\pi_0}(n^2) = q_n \quad (k \leq n \leq k(k+3)/2).$$

**Step 4.**  $Q_{\pi_0}$  is given by  $Q_{\pi_0}(n) = G_{\pi_0}(n^2)$ .

**Algorithm C** Computation of submatching event probabilities in the half-planar model (assuming Conjecture 1.12)

Use Algorithm B above to compute the polynomial  $Q_{\pi_0}$ . Let  $q_{\pi_0}^*$  be the leading coefficient of  $Q_{\pi_0}$ . The probability  $\mathbb{P}(\pi_0 \triangleleft \Pi_*)$  is  $q_{\pi_0}^*/2^{k(k+1)}$ .

**Algorithm D** Computation of the matrices  $\mathbf{C}_n, \tilde{\mathbf{C}}_n$

Translating (35) to a notation appropriate for the computation of  $C_{\pi, \sigma}$ , we denote  $p_{j,k}(\pi) = \#\{j \leq \ell < k : \pi_\ell = 1\} - \frac{1}{2}(k-j-1)$ . The matrix  $\mathbf{C}_n = (C_{\pi, \sigma})_{\pi, \sigma \in \text{NC}_n}$  is now computed using the explicit formula

$$C_{\pi, \sigma} = \prod_{j < k, j \xleftrightarrow{\sigma} k} \chi(p_{j,k}(\pi))$$

(with  $\chi(\cdot)$  defined in (34)), and  $\tilde{\mathbf{C}}_n = \mathbf{C}_n^{-1}$  is its inverse matrix.

**Algorithm E** Computation of the polynomials  $F_{\pi_0}$  (see Theorem 1.13)

**Step 1.** Compute  $\tilde{\mathbf{C}}_k$  using Algorithm D above.

**Step 2.** Use (55) to compute  $F_{\pi_0}$ .

## Appendix B. Proof of Theorem 1.8

We start by writing a more explicit formula for the entries of the transition matrices  $T_n^{(p)}$ . To do so, encode each row of plaquettes as a vector  $\mathbf{a} \in \{0, 1\}^{2n}$ , where a 0 coordinate corresponds to a “type 0” plaquette, defined as the plaquette shown on the left in Fig. 6, and a 1 coordinate corresponds to a “type 1” plaquette, which is the one on the right-hand side of the same figure. The parameter  $p$  corresponds to the probability of a type 1 plaquette. For a given plaquette-row vector  $\mathbf{a} \in \{0, 1\}^{2n}$ , denote by  $f_{\mathbf{a}}(\cdot)$  an operator that takes a noncrossing matching  $\pi \in \text{NC}_n$  and returns a new noncrossing matching  $\pi' = f_{\mathbf{a}}(\pi)$  which is the result of composing the diagram of  $\pi$  with the row of plaquettes, in an analogous manner to the composition of matching diagrams with Temperley-Lieb operators shown in Fig. 8.

With this notation, it is clear from the definition of the transition matrix  $T_n^{(p)}$  that its entries  $(T_n^{(p)})_{\pi, \pi'}$  are given by

$$(T_n^{(p)})_{\pi, \pi'} = \sum_{\substack{\mathbf{a} \in \{0, 1\}^{2n} \\ f_{\mathbf{a}}(\pi) = \pi'}} p^{\sum_j a_j} (1 - p)^{\sum_j (1 - a_j)} \quad (\pi, \pi' \in \text{NC}_n).$$

Differentiate this equation with respect to  $p$  at  $p = 0$ . On the right-hand side the nonzero contributions will come from the vector  $\mathbf{a} = \mathbf{0} = (0, \dots, 0)$  and from the vectors of the form  $\mathbf{a} = \mathbf{a}_k = (0, \dots, 0, 1, 0, \dots, 0)$  with only one coordinate in some position  $k$  equal to 1. This gives

$$\frac{d}{dp} \Big|_{p=0} (T_n^{(p)})_{\pi, \pi'} = -2n \delta_{f_{\mathbf{0}}(\pi), \pi'} + \#\{1 \leq k \leq 2n : f_{\mathbf{a}_k}(\pi) = \pi'\}.$$

It is now not difficult to check (see Fig. 13) that  $f_{\mathbf{0}}(\pi) = \rho(\pi)$ , where  $\rho$  is the rotation operator on matchings (defined in Theorem 2.7), and that the operator  $f_{\mathbf{a}_k}$  bears a simple relation to the Temperley-Lieb operator  $e_k$ , namely, we have  $f_{\mathbf{a}_k} = \rho \circ e_k = e_{k-1} \circ \rho$  (where  $e_{k-1}$  is interpreted with the usual mod  $2n$  convention).

Summarizing, since the vector  $\boldsymbol{\mu}_n$  is the stationary probability vector for the matrices  $T_n^{(p)}$ , that is, it satisfies  $\boldsymbol{\mu}_n T_n^{(p)} = \boldsymbol{\mu}_n$ , the matrix  $Q_n = \frac{d}{dp} \Big|_{p=0} T_n^{(p)}$  satisfies  $\boldsymbol{\mu}_n Q_n = 0$ , and by the above observations it is easy to see that it is related to the operator  $H_n$  from (5) by  $H_n = -R^{-1} Q_n$ , where  $R$  is a matrix version of the rotation operator  $\rho$ , whose entries are given by  $R_{\pi, \pi'} = \delta_{\rho(\pi), \pi'}$ . It remains to observe that  $\boldsymbol{\mu}_n$  is invariant also under the rotation action, i.e., we have  $\boldsymbol{\mu}_n R = \boldsymbol{\mu}_n$ , to conclude that  $\boldsymbol{\mu}_n H_n = 0$ , as claimed.  $\square$

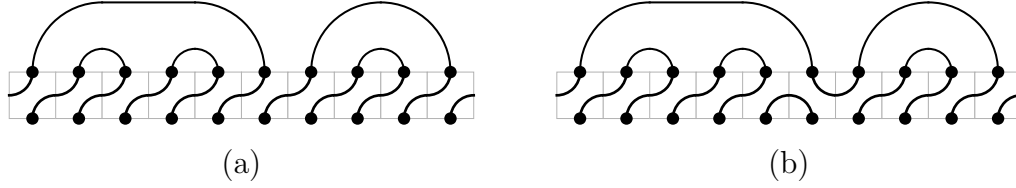


Figure 13: (a) Composing the diagram of a matching  $\pi$  with a row of type 0 plaquettes causes a simple rotation; (b) flipping the  $k$ th plaquette from type 0 to type 1 and leaving all other plaquettes as type 0 is equivalent to the application of a Temperley-Lieb operator  $e_k$  prior to the rotation.

## Acknowledgements

The author thanks Omer Angel, Nathanaël Berestycki, Jan de Gier, Christina Goldschmidt, Alexander Holroyd, Rick Kenyon, James Martin, Bernard Nienhuis, Ron Peled, David Wilson, Doron Zeilberger, and Paul Zinn-Justin for comments and helpful discussions during the preparation of the paper, and an anonymous referee for helpful suggestions.

The author was supported by the National Science Foundation under grant DMS-0955584, and by grant #228524 from the Simons Foundation.

## References

- [1] G. E. Andrews. Plane partitions V: the TSSCPP conjecture. *J. Combin. Theory Ser. A* **66** (1994), 28–39.
- [2] G. E. Andrews, W. H. Burge. Determinant identities. *Pacific J. Math.* **158** (1993) 1–14.
- [3] M. Batchelor, J. de Gier, B. Nienhuis. The quantum symmetric XXZ chain at  $\Delta = -1/2$ , alternating-sign matrices and plane partitions. *J. Phys. A* **34** (2001), L265–L270.
- [4] D. M. Bressoud, Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture. Cambridge University Press, 1999.
- [5] L. Cantini, A. Sportiello. Proof of the Razumov-Stroganov conjecture. *J. Comb. Theory Ser. A* **118** (2011), 1549–1574.

- [6] P. Di Francesco, P. Zinn-Justin, J.-B. Zuber. Sum rules for the ground states of the  $O(1)$  loop model on a cylinder and the XXZ spin chain. *J. Stat. Mech.* (2006), P08011.
- [7] F. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Math. Phys.* **3** (1962), 140–156.
- [8] B. Feigin, M. Jimbo, T. Miwa, E. Mukhin. Symmetric polynomials vanishing on the shifted diagonals and Macdonald polynomials. *Int. Math. Res. Not.* **2003** (2003), 1015–1034.
- [9] J. de Gier. Loops, matchings and alternating-sign matrices. *Discrete Math.* **298** (2005), 365–388.
- [10] J. de Gier, A. Lascoux, M. Sorrell. Deformed Kazhdan-Lusztig elements and Macdonald polynomials. *J. Combin. Theory Ser. A* **119** (2012), 183–211.
- [11] T. Fonseca, P. Zinn-Justin. On the doubly refined enumeration of alternating sign matrices and totally symmetric self-complementary plane partitions. *Electron. J. Combin.* 15 (2008), #R81.
- [12] T. Fonseca, P. Zinn-Justin. On some ground state components of the  $O(1)$  loop model. *J. Stat. Mech. Theory Exp.* (2009), P03025.
- [13] I. J. Good. Short proof of a conjecture of Dyson. *J. Math. Phys.* **11** (1970), 1884.
- [14] G. Grimmett. Percolation, 2nd Ed. Springer, 1999.
- [15] J. Gunson. Proof of a conjecture of Dyson in the statistical theory of energy levels. *J. Math. Phys.* 3 (1962), 752–753.
- [16] M. Kasatani. Subrepresentations in the polynomial representation of the double affine Hecke algebra of type  $GL_n$  at  $t^{k+1}q^{r-1} = 1$ . *Int. Math. Res. Not.* **2005** (2005), 1717–1742.
- [17] C. Krattenthaler. Determinant identities and a generalization of the number of totally symmetric self-complementary plane partitions. *Electron. J. Combin.* 4 (1997), #R27.

- [18] G. Kuperberg. Another proof of the alternating sign matrix conjecture. *Internat. Math. Res. Notes* **1996** (1996), 139–150.
- [19] G. F. Lawler, O. Schramm, W. Werner. One-arm exponent for critical 2D percolation. *Electron. J. Probab.* **7**, paper no. 2 (2002).
- [20] W. H. Mills, D. P. Robbins, H. Rumsey, Alternating sign matrices and descending plane partitions. *J. Combin. Theory Ser. A* **34** (1983), 340–359.
- [21] S. Mitra, B. Nienhuis, J. de Gier, M. T. Batchelor. Exact expressions for correlations in the ground state of the dense  $O(1)$  loop model. *J. Stat. Mech. Theory Exp.* (2004), P09010.
- [22] S. Okada. Enumeration of symmetry classes of alternating sign matrices and characters of classical groups. *J. Algebraic Combin.* **23** (2006), 43–69.
- [23] V. Pasquier. Quantum incompressibility and Razumov-Stroganov type conjectures. *Ann. Henri Poincaré* **7** (2006), 397–421.
- [24] P. A. Pearce, V. Rittenberg, J. de Gier, B. Nienhuis. Temperley-Lieb stochastic processes. *J. Phys. A.* **35** (2002), L661–L668.
- [25] J. Propp, The many faces of alternating sign matrices. *Discr. Math Theor. Comp. Sci.* Proceedings of **DM-CCG**, Conference Volume **AA** (2001), 43–58.
- [26] A. V. Razumov, Yu. G. Stroganov. Combinatorial nature of ground state vector of  $O(1)$  loop model. *Theor. Math. Phys.* **138** (2004), 333–337.
- [27] D. Romik. Connectivity patterns in loop percolation II: pipe percolation. In preparation.
- [28] A. V. Sills, D. Zeilberger. Disturbing the Dyson conjecture (in a GOOD way). *Exp. Math.* **15** (2006), 187–191.
- [29] S. Smirnov. Critical percolation in the plane: conformal invariance, Cardy’s formula, scaling limits. *C. R. Acad. Sci. Paris Sér. I Math.* **333**, 239–244 (2001).

- [30] Yu. Stroganov. Izergin-Korepin determinant at a third root of unity. *Theor. Math. Physics* **146** (2006), 53–62.
- [31] N. Temperley, E. Lieb. Relations between the percolation and colouring problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the percolation problem. *Proc. R. Soc. A* **322** (1971), 251–280.
- [32] B. Wieland. A large dihedral symmetry of the set of alternating sign matrices. *Electron. J. Combin.* **7** (2000), #R37.
- [33] K. Wilson. Proof of a conjecture by Dyson. *J. Math. Phys.* **3** (1962), 1040–1043.
- [34] D. Zeilberger. Proof of the alternating sign matrix conjecture. *Electronic J. Combin* **3** (1996), R13.
- [35] D. Zeilberger. Proof of a conjecture of Philippe Di Francesco and Paul Zinn-Justin related to the qKZ equations and to Dave Robbins’ two favorite combinatorial objects. Preprint (2007), available at <http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimPDF/diFrancesco.pdf>
- [36] P. Zinn-Justin. Six-vertex, loop and tiling models: integrability and combinatorics. Habilitation thesis, [arXiv:0901.0665](https://arxiv.org/abs/0901.0665).
- [37] P. Zinn-Justin, P. Di Francesco. Around the Razumov-Stroganov conjecture: proof of a multi-parameter sum rule. *Electron. J. Combin.* **12** (2005), #R6.
- [38] P. Zinn-Justin, P. Di Francesco. Quantum Knizhnik-Zamolodchikov equation, totally symmetric self-complementary plane partitions and alternating sign matrices. *Theor. Math. Phys.* **154** (2008), 331–348.
- [39] J.-B. Zuber. On the counting of fully packed loop configurations: some new conjectures. *Elec. J. Combin.* **11** (2004), #R13.

Dan Romik  
 Department of Mathematics  
 University of California, Davis  
 One Shields Avenue, Davis, CA 95616  
 Email: [romik@math.ucdavis.edu](mailto:romik@math.ucdavis.edu)