

CORRELATIONS IN THE MULTISPECIES TASEP AND A CONJECTURE BY LAM

ARVIND AYYER AND SVANTE LINUSSON

ABSTRACT. We study correlations in the multispecies TASEP on a ring. Results on the correlation of two adjacent points prove two conjectures by Thomas Lam on

- (a) the limiting direction of a reduced random walk in \tilde{A}_{n-1} and
- (b) the asymptotic shape of a random integer partition with no hooks of length n , a so called n -core.

We further investigate two-point correlations far apart and three-point nearest neighbour correlations and prove explicit formulas in almost all cases. These results can be seen as a finite strengthening of correlations in the TASEP speed process by Amir, Angel and Valkó. We also give conjectures for certain higher order nearest neighbour correlations. We find an unexplained independence property (provably for two points, conjecturally for more points) between points that are closer in position than in value that deserves more study.

1. INTRODUCTION

Probabilistic processes on combinatorial structures are currently of much interest. Thomas Lam initiated the study of infinite random reduced words in affine Weyl groups in [12]. This can equivalently be formulated as a random walk on the affine Coxeter arrangement conditioned never to cross the same hyperplane twice. He first states a remarkable formula for the limiting direction of the random walk on the Weyl alcoves in terms of a certain finite Markov chain on the underlying finite Weyl group; see Theorem 2.7 below. He further specializes to the affine Weyl group \tilde{A}_{n-1} and conjectures a closed formula for the limiting direction. Via known bijections (see, for example, [13]) between affine Grassmannians and so-called n -cores, a special class of integer partitions, he also conjectures a limit shape for the natural growth process of these partitions; see Theorem 3.2. This limit shape is a natural finite version of the famous limit shape by Rost, [11, 17]. The first purpose of this paper is to prove these two conjectures.

Lam states his conjectures in terms of a certain Markov chain on the set of permutations. This chain turns out to be equivalent (see [5]) to a Markov chain, which is a multispecies variant of the Totally Asymmetric Simple Exclusion Process (or TASEP) on a ring. Unknown to Lam, that Markov chain was already studied by probabilists and statistical physicists, and the stationary distribution was given an explicit interpretation in terms of queueing theory by Ferrari and Martin [8, 9].

We will prove Lam's conjectures by studying certain two-point correlations in the multispecies TASEP on a ring. In another direction, it turns out that Amir,

Received by the editors May 6, 2014 and, in revised form, February 10, 2015.
2010 *Mathematics Subject Classification*. Primary 05A05, 05A17, 20F55, 60K35, 82C22.

Angel and Valkó [4] have studied the correlations in an infinite volume limit of the multispecies TASEP, called the *TASEP speed process*. Motivated by the study of correlations in the three-species exclusion process on \mathbb{Z} [7, 16], they gave exact formulas for various marginals in the TASEP speed process. In particular, they studied two-point and three-point nearest-neighbour correlations in great detail as well as two-point correlations a distance apart.

This naturally leads us to the second purpose of this paper, namely to further explore correlations in the multispecies TASEP. Apart from the two-point correlations above, we find a closed formula, Theorem 5.1 for the correlation of two points further apart. In important special cases, the formula simplifies to extremely simple expressions. We also give closed formulas for correlation of three adjacent particles in five of the six possible cases. The correlation formulas show a remarkable independence property, provably for two particles and conjecturally for more particles; see Remark 4.3, Corollary 5.2 and the discussion in Section 8. Our proofs are of a somewhat intricate combinatorial nature, relying on the multiline queues of Ferrari and Martin. We would be very interested to see a more conceptual proof that could give some understanding and hopefully prove some of our conjectures. In a certain sense, several of the formulas by Amir, Angel and Valkó [4] are limits of our formulas on correlations, which thus are finite strengthenings of theirs.

The structure of the paper is as follows. Section 2 contains background information: Section 2.1 defines the multispecies TASEP and the multiline process, Section 2.2 defines infinite reduced words in \tilde{A}_n and Section 2.3 defines n -cores. In Section 3 we present the conjectures by Lam about the limiting objects, first the limiting direction of the random walk in Section 3.1 and the limiting shape of n -cores in Section 3.2. We also prove there that Lam's conjectures follow from our computations about the nearest neighbour two-point correlations in Section 4, where we also explain the connection to the TASEP speed process. We continue our study of two-point correlations far apart in Section 5. Proofs of the formulas there will require enumerative formulas for certain constrained semistandard Young tableaux with two columns, which we relegate to Section 6. We finally study nearest neighbour three-point correlations in Section 7. We end with some conjectures and open problems in Section 8 that are deserving of future study.

Note added in proof. Conjectures 7.4, 8.2 and 8.3 have been proved in [1], and Conjecture 8.1 has been proved in [2]. \square

2. BACKGROUND

2.1. Multispecies TASEP and multiline queues. In this section we describe the multispecies TASEP and the multiline queue process defined by Ferrari and Martin [8] that we will use. We will borrow notation from [5], but we will describe the homogeneous model that we are interested in rather than the more general inhomogeneous model. The interested reader should consult [5] for more details.

The multispecies version of the totally asymmetric simple exclusion process (or TASEP for short) is a stochastic process which can be defined on an arbitrary directed graph, but we will only define it on directed path graphs of length N . The physical model is as follows. There are N locations arranged in the shape of a ring, and each location is occupied by a particle. There are n species of particles, labelled 1 through n and m_i particles of species i . A multispecies TASEP model is thus determined by the n -tuple $m \equiv (m_1, \dots, m_n)$ such that $m_1 + \dots + m_n = N$.

The space of configurations will be denoted Ω_m . It is easy to see that $|\Omega_m| = \binom{N}{m_1, \dots, m_n}$ and can be described naturally by multipermutations S_m , with m_i repetitions of i , for $i \in [n]$. If each $m_i = 1$, then the state space becomes that of ordinary permutations S_n .

Each particle carries an independent exponential clock which rings with rate 1. Whenever the clock for particle j rings, it tries to exchange positions with the particle k to its immediate left. The exchange is only successful if $k > j$ and fails otherwise. This Markov process is completely described by its generator, which is an $|\Omega_m| \times |\Omega_m|$ matrix M_m labelled by multipermutations whose off-diagonal (σ, τ) -entry describe the transition rate from $\tau \rightarrow \sigma$,

$$M_m(\sigma, \tau) = 1, \quad \text{if } \sigma_i = \tau_{i+1} > \tau_i = \sigma_{i+1} \text{ and } \sigma_j = \tau_j \text{ for } j \neq i, i + 1,$$

for $\sigma \neq \tau$ and whose diagonal entries are negatives of the total incoming transitions,

$$M_m(\sigma, \sigma) = -\#\{\rho \in \Omega_m, \rho \neq \sigma : M_m(\rho, \sigma) = 1\}.$$

Let $\mathbb{P}_m(\sigma)$ denote the stationary probability of the multipermutation σ of type m . The following facts about the stationary probability distribution of the multispecies TASEP are well known.

- Proposition 2.1.** (i) Let $\sigma' = (\sigma_2, \dots, \sigma_N, \sigma_1)$ denote the rotated version of σ . Then $\mathbb{P}_m(\sigma') = \mathbb{P}_m(\sigma)$.
 (ii) Let $m^{rev} = (m_n, m_{n-1}, \dots, m_1)$ and $\sigma^{rev} = (n + 1 - \sigma_N, \dots, n + 1 - \sigma_1)$. Then $\sigma^{rev} \in \Omega_{m^{rev}}$ and $\mathbb{P}_{m^{rev}}(\sigma^{rev}) = \mathbb{P}_m(\sigma)$.
 (iii) The stationary probability that the first site is occupied by species i is given by $\mathbb{P}_m(w_1 = i) = m_i/N$.

The first property in Proposition 2.1 is called *rotational symmetry* and follows simply because the Markovian dynamics is independent of the position. The second is a generalized version of what is sometimes called *particle-hole symmetry*, although notice that the direction has also been switched. The last property follows from the first.

The following amazing result about stationary probability distribution was proved by Ferrari and Martin.

Theorem 2.2 (P. Ferrari and J. Martin, [9], Theorem 5.1). *The stationary probability of multipermutations in Ω_m are integer multiples of*

$$\left(\prod_{i=1}^n \binom{N}{m_1 + \dots + m_i} \right)^{-1}$$

with the integer being 1 for the reverse permutation

$$\underbrace{n \cdots n}_{m_n} \underbrace{(n-1) \cdots (n-1)}_{m_{n-1}} \cdots \underbrace{1 \cdots 1}_{m_1}.$$

Theorem 2.2 was proved by Ferrari and Martin by considering a larger Markov chain and projecting to the multispecies TASEP Markov chain by the procedure known as *lumping*. We now describe this larger model, which is called the (*Ferrari-Martin*) *multiline process*. Just as the multispecies TASEP is defined on a ring of circumference N , the multiline process is defined on a discrete cylinder of circumference N and height $n - 1$ called a *multiline queue*, which can be thought of as a

stack of $n - 1$ such rings containing a total of $(n - 1)N$ sites. Each site contains exactly one of two symbols, \circ and \bullet , called vacant and occupied respectively.

Given an n -tuple $m \equiv (m_1, \dots, m_n)$ of positive integers summing to N , the set of multiline queues of type m , denoted Ω_m^{FM} , is the set of all configurations of \circ 's and

\bullet 's such that the number of \bullet 's at row i , enumerated from the top, is $M_i = \sum_{j=1}^i m_j$.

Since the positions of \bullet 's at each row are independent of one another, the total number of configurations is

$$|\Omega_m^{FM}| = \prod_{i=1}^{n-1} \binom{N}{M_i},$$

which is the same as the denominator in Theorem 2.2 since $M_n = N$. See Figure 1 for a multiline queue corresponding to $n = 5$, $N = 9$ and $m = (2, 1, 2, 2, 2)$ (ignore the numbers and the lines for now).

Ferrari and Martin [9] defined transitions between multiline queues that turned Ω_m^{FM} into a Markov chain, which was called the (homogeneous) Ferrari-Martin multiline process. For our purposes, we will not need the description of the multiline process and we skip it. See [5] for a combinatorial description of the process.

Theorem 2.3 (P. Ferrari and J. Martin, [9, Theorem 3.1]). *The stationary distribution of the Ferrari-Martin multiline process is uniform on Ω_m^{FM} .*

We do not need the precise definition of the transitions in the Ferrari-Martin multiline queues. The interested reader is referred to [8]. See [5] for illustrative examples of both the multispecies TASEP and the multiline queueing Markov chains. Several different sets of transitions rules that also give uniform stationary distribution on Ω_m^{FM} are given in [15].

The connection between the multiline process and the TASEP is expressed through the procedure formally known as lumping of Markov chains [14, Lemma 2.5]. We will now describe the procedure called a *bully path projection* which relates the stationary distribution of the multispecies TASEP to the stationary distribution of the multiline process, which is the uniform distribution on multiline queues by Theorem 2.3. As expected from a projection procedure, this will be a surjective map $B : \Omega_m^{FM} \rightarrow \Omega_m$.

Let $C \in \Omega_m^{FM}$. The projection is defined recursively by bully paths. A *bully path* is a path going through the multiline queue which always moves rightwards or downwards and which contains exactly one \bullet from each row. Moreover, it moves downwards from a given row j if and only if it has encountered a \bullet in row j that has not already been part of another bully path.

See Figure 1 for the running example constructing the bully path for a particular multiline queue. We start by defining bully paths starting at locations $(1, i)$ (where we are using matrix notation for positions in multiline queues) such that $C_{1,i} = \bullet$ as described above. (The order in which we run these paths among these \bullet 's turns out not to matter.) All along these paths, we mark by a 1 all the \bullet 's in each row where the bully path moves downwards. Therefore, there will be m_1 positions in each row containing \bullet marked by a 1. In Figure 1, bully paths for 1 are coloured red.

Next, we start bully paths at locations $(2, j)$ such that $C_{2,j} = \bullet$ and moreover which were not part of the bully paths from the first row. By the definition of

multiline queues and the nature of the bully paths, there will be exactly m_2 of these. All along these paths, we mark by a 2 all the \bullet 's in each row where the bully path moves downwards. In Figure 1, bully paths for 2 are coloured blue. We continue in this way for all the \bullet 's in all the rows. There will be exactly m_j bully paths starting in row j leading to a j on row n for $j \in [n - 1]$. Finally, we mark the symbol n on row n below all k 's such that the $C_{n-1,k} = \circ$. In Figure 1, bully paths for 3 and 4 are coloured green and black respectively.

At the end, the configuration on row n of entries from $\{1, \dots, n\}$ is precisely a multipermutation in Ω_m . This is the required configuration in Ω_m . We call this projection B . We remark that B is well defined in the sense that the order of the bully paths starting at a given row do not matter. The following theorem essentially states that the bully path projection is a lumping of Markov chains.

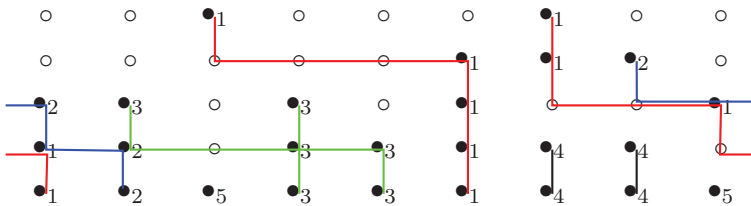


FIGURE 1. A multiline queue for $N = 9, m = (2, 1, 2, 2, 2)$, with the bully-path projection to a multipermutation.

Theorem 2.4 (P. Ferrari and J. Martin [9, Theorem 4.1]). *The n th line of the multiline process on Ω_m^{FM} constructed by the bully path projection is exactly the multispecies TASEP on Ω_m .*

Therefore one way to understand correlations in the multispecies TASEP is by a combinatorial understanding of the multiline queues which give rise to those multipermutations which contribute to the correlations. In particular, the stationary probability of a multipermutation is given as follows.

Corollary 2.5. *The stationary probability of a multipermutation π in the TASEP is equal to*

$$\frac{\#\{q \in \Omega_m^{FM} : B(q) = \pi\}}{\prod_{i=1}^{n-1} \binom{N}{M_i}}.$$

2.2. Infinite reduced words in \tilde{A}_{n-1} . Weyl groups are special finite Coxeter groups of fundamental importance in both the theory of Lie algebras and in algebraic combinatorics; see e.g. [6, 10]. A Weyl group is generated by reflections in hyperplanes perpendicular to the roots in a so-called root system. An affine Weyl group is a special infinite group with one more simple reflection than the underlying (finite) Weyl group which causes the addition of an infinite number of roots and hyperplanes.

Any element of a general Coxeter group (finite or infinite) can be represented as a word in the generators. A reduced word is a word which cannot be reduced in length by using relations among the generators. An infinite reduced word in the affine Weyl group \tilde{A}_{n-1} is a word $\dots s_{i_5} s_{i_4} s_{i_3} s_{i_2} s_{i_1}$, where s_{i_j} 's, $0 \leq i_j \leq n$,

are generators and such that the word is infinite to the left and all finite initial sequences $s_{i_k} s_{i_{k-1}} \dots s_{i_2} s_{i_1}$ are reduced words of the affine Weyl group.

Lam [12] studied a specific kind of random infinite reduced word in a general affine Weyl group. Here we will focus on the group \tilde{A}_{n-1} for which Lam stated interesting conjectures. This random word can be seen as a walk on the alcoves of the arrangement corresponding to \tilde{A}_{n-1} , i.e. $\{x_i - x_j = d : 1 \leq i < j \leq n, d \in \mathbb{Z}\}$ such that it never crosses the same hyperplane twice (starting in the fundamental alcove $x_n + 1 > x_1 > x_2 > \dots > x_n$). Lam considered uniformly reduced walks, i.e. at each step choosing uniformly one of the legal hyperplanes to cross. See Figure 2 for an example of a reduced walk. Note that \tilde{A}_{n-1} lives in \mathbb{R}^n but it extends trivially in the direction of $(1, 1, \dots, 1)$. The essential dimension is $n - 1$, and in Figure 2 it is presented as a two-dimensional intersection.

From general Coxeter theory, it is not difficult to see that a uniform reduced random walk is a Markov process, i.e. it only matters where the walk is currently stationed.

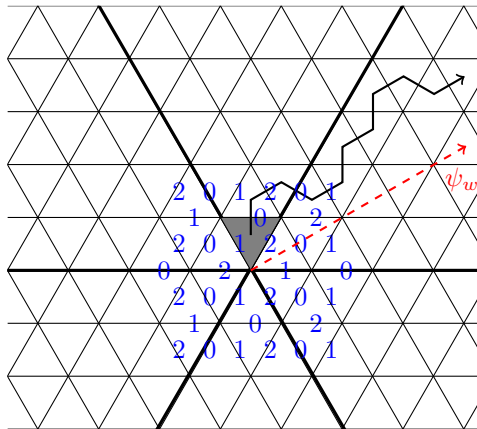


FIGURE 2. A reduced random walk on the alcoves of the \tilde{A}_2 arrangement. The shown walk has reduced word $\dots s_2 s_1 s_0 s_1 s_2 s_0 s_1 s_0 s_1 s_0 s_2 s_0$. The thick lines divide into Weyl chambers. A random walk trapped in chamber w will almost surely be asymptotically parallel to the red dashed line in the direction of ψ_w .

The study of reduced words of affine Weyl groups is part of a larger goal to try to lift results from the finite groups to the affine situation. A reduced random walk will with probability one ‘get trapped’ in one of the chambers (cones) of the underlying finite Weyl group A_{n-1} (which is the symmetric group) after a finite time. A walk trapped in chamber $w \in A_{n-1}$ will almost surely go in the direction of a vector ψ_w in that chamber. Reduced walks for general Weyl groups was studied by Lam [12]. We will state his results in the case of \tilde{A}_{n-1} .

Theorem 2.6 (Lam [12]). *Let (X_0, X_1, \dots) be a reduced random walk in \tilde{A}_{n-1} and let $v(X_i)$ denote the unit vector pointing towards the central point of the alcove X_i . There exists a unit vector $\psi \in \mathbf{R}^n$ so that almost surely we have*

$$(1) \quad \lim_{N \rightarrow \infty} v(X_i) \in A_{n-1} \cdot \psi.$$

Lam also proves, and this is the deepest theorem in the paper, that the Markov chain can be projected down to a certain Markov chain Ω on A_{n-1} , with the following properties. If we let $\zeta(w)$ be the stationary distribution for $w \in A_{n-1}$ for Ω , then $\zeta(w)$ is also the probability that the reduced random walk ends up in chamber w_0w^{-1} . The chain Ω is equivalent (it exchanges values instead of positions) to the TASEP we discuss in this paper. Furthermore he proved that the distribution ζ also determines the direction of the vector ψ . The standard representation of the root system for A_{n-1} uses simple roots $\alpha_i = e_i - e_{i+1}$ and highest root $\theta = e_1 - e_n$, where e_i are the standard basis. Let r_θ be the reflection in the hyperplane perpendicular to θ .

Theorem 2.7 (Lam [12]). *The vector ψ for \tilde{A}_{n-1} of Theorem 2.6 is given by*

$$\psi = \frac{1}{Z} \sum_{w \in A_{n-1} : r_\theta w > w} \zeta(w)w^{-1}(\theta),$$

where Z is a normalization factor. Furthermore, the probability that the reduced random walk ends up in the chamber corresponding to w is $\zeta(w^{-1}w_0)$.

Lam also stated the following conjecture, where ρ is the sum of all positive roots.

Conjecture 2.8 (Lam [12, Conjecture 2]). *For \tilde{A}_{n-1} , $\psi = \alpha\rho$ for some real number $\alpha > 0$.*

We will prove this conjecture in Section 3.

2.3. n -cores. The n -cores are special Young diagrams which come up in the study of the affine Grassmanian in algebraic combinatorics (see e.g. [13]) and can be defined in several equivalent ways. We will define these in terms of the hook length, but one can also define them in terms of ribbons.

Recall that a partition $\lambda = (\lambda_1, \dots, \lambda_r)$, whose entries are positive and weakly decreasing, can be represented as an array of left-justified boxes (called cells) such there are λ_1 cells at row 1, λ_2 cells at row 2, and so on. Such a representation is called a Young diagram, also denoted by λ . We will take the convention that the rows are arranged from top to bottom in increasing order. The *hook* of a cell c in a Young diagram is the set of cells directly to its right and directly below it, as well as c itself. The *hook length* of c is the number of elements in its hook. An n -core then is a partition that contains no cell whose hook length is divisible by n . Figure 3 shows an example of a 4-core.

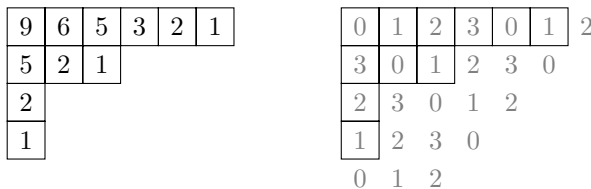


FIGURE 3. Two copies of the same 4-core. To the left is the hook length stated in each box. To the right is the content modulo 4.

The growth model of a random n -core is the following. Each position (i, j) (matrix indexing) in the quarter plane is marked by its content $j - i \pmod n$. A

growth corner of a Young diagram is a square just outside the diagram were a box could be added and it would still be a diagram. At each time step an integer t in $[0, n - 1]$ is chosen uniformly at random. For every position that is a growth corner and with content $t \pmod n$ we then add a box to the diagram. If, for example, $n = 4$ and the random sequence of integers $0, 2, 3, 1, 2, 3, 0, 1$, we would get the diagram in Figure 3. Note that the first time the integer 2 appears in the sequence no box is added. There is a well-known bijection between n -cores and affine Grassmanian elements of \tilde{A}_{n-1} . The sequence in the growth model corresponds to left-multiplication by the simple generators. For more information on this model of growing n -cores and the relation to algebraic combinatorics, see [13].

Note that this is the natural generalization of the growth model studied by Rost [17] and Johansson [11]. As we will see in Section 3.2 the limit shape also matches the limit shape in that situation.

3. LIMIT THEOREMS

In this section we will prove Conjecture 2.8, that is, we will determine the limiting direction of reduced random walks in the affine weyl groups \tilde{A}_{n-1} . This also implies the exact limiting shape of partitions with no hooks of length n ; see Section 3.2.

3.1. Reduced random walks in \tilde{A}_{n-1} . Let e_i denote the unit vector in the i th coordinate direction. The highest root in Theorem 2.7 is $\theta = e_1 - e_n$. In our setting Theorem 2.7 means summing over all permutations with $w_n > w_1$, where $w^{-1}(\theta) = e_{w_1} - e_{w_n}$ and $\zeta(w)$ is the stationary distribution for w in the TASEP over permutations on a ring.

The important quantity is thus $\mathbb{P}(w_1 = i, w_n = j) = \sum_{w:w_1=i,w_n=j} \zeta(w)$. By rotational symmetry, namely Proposition 2.1(i), we can switch to the first and second position of w . For $j > i$, let $E_{j,i}(n) = \mathbb{P}(w_1 = j, w_2 = i)$ and $E_{i,j}(n) = \mathbb{P}(w_1 = i, w_2 = j)$ in the TASEP Ω_n . We rewrite Lam’s formula in Theorem 2.7 as

$$(2) \quad \psi = \frac{1}{Z} \sum_{j=2}^n \sum_{i=1}^{j-1} E_{j,i}(e_i - e_j).$$

Using this, we will be able to prove Conjecture 2.8.

Theorem 3.1. *The unit vector $\psi = \psi_{id}$ for the limiting direction of a reduced random walk constrained to the fundamental chamber is*

$$\psi = \frac{1}{\sqrt{2 \binom{n+1}{3}}} \sum_{k=1}^n (n + 1 - 2k)e_k.$$

Proof. Given Theorem 4.2 the coefficient of e_k in (2) is

$$\frac{1}{n \binom{n}{2}} \left(\binom{n+1-k}{2} - \binom{k}{2} \right) = \frac{n+1-2k}{\binom{n}{2}}.$$

□

Note that this can also be stated (as Lam did) as ψ is the sum of all positive roots, i.e. $\psi = \alpha \sum_{1 \leq i < j \leq n} e_i - e_j$, for some constant α .

An intuitive way to understand (2) is that when particle i jumps over particle j , this corresponds to a step in the walk (constrained to the fundamental chamber) crossing a hyperplane of the type $x_i - x_j = d$, for some integer d .

3.2. Limit shape of random n -cores. Thanks to the work of Lam we deduce as a direct consequence of Theorem 3.1 the limit shape for a random n -core. Let C_n be the piecewise linear curve with vertices $v_i = \gamma(\binom{i}{2}, \binom{n-i+1}{2})$, for $i = 1, \dots, n$, where $\gamma = \frac{2\sqrt{6}}{n\sqrt{n^2-1}}$ is a scaling constant to make the area 1 between the curve and the axes; see Figure 4.

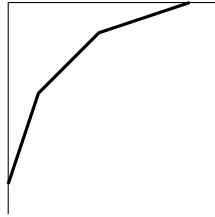


FIGURE 4. The limiting piecewise-linear curve C_4 for random 4-cores.

Let D_n^K be the boundary of a random n -core after K time steps scaled so the area of the n -core is 1. Now, the theorem is that the curve C_n is the limit shape of a random n -core.

Theorem 3.2. *For each $\epsilon, \delta > 0$ there is an L such that for all $K > L$ we have*

$$\mathbb{P}(|D_n^K - C_n| > \delta) < \epsilon,$$

where the distance between the curves is the supremum of the distances measured along the diagonals $y = -x + c$ for all c .

The proof of this theorem follows from Proposition 2 in [12, Section 5] and Theorem 3.1.

Remark 3.3. The classical limit shape for partitions, first proved by Rost [17] in studies of first passage percolation, is $\sqrt{x} + \sqrt{-y} = 6^{1/4}$, scaled so the area is 1. Note that the vertices of C_n converge to being on that shape, so the limit of C_n as $n \rightarrow \infty$ is the classical shape.

Remark 3.4. The slopes of the curve D_n^K are deterministically $i/(n-i)$, i.e. it will have a pattern of $n-i$ steps to the left and then i steps down for a long time until it changes to $n-i-1$ steps to the left and $i+1$ steps down. The position of this change corresponds to the vertex v_{n-i} of the limit shape. The positions where the slope changes are random and converge to the v_i 's.

4. CORRELATION OF TWO ADJACENT PARTICLES

For a particle i in the TASEP Ω_n all the particles of a higher class look the same and all the particles of a lower class look the same. If we want to study the correlation of the particle of class j and class $i, i < j$ in Ω_n we can study the correlation of 2 and 4 in the five species TASEP Ω_m , where $m = (i-1, 1, j-i-i, 1, n-j)$. We call this the **projection principle**, and it will be used repeatedly in this and the coming sections. Even though five species systems are easier to study

than arbitrary n , they are often still too complicated to make precise calculations. We will go one step further and project to many three species systems. In this section we will use $m_{s,t} = (s, t, n - s - t)$ to denote a system with s 1's, t 2's and the remaining 3's. We will focus on all possible projections where $i > s$ and $j > s + t$ so that the particle of class i (resp. j) will become a 2 (resp. 3). Studying the correlation between 2 and 3 in these systems will give us the correlation between particles i and j .

Let $T_{s,t} = \mathbb{P}(w_1 = 3, w_2 = 2)$ in $\Omega_{m_{s,t}}$. For $i < j$, recall that $E_{j,i} \equiv E_{j,i}(n) = \mathbb{P}(w_1 = j, w_2 = i)$ and $E_{i,j} \equiv E_{i,j}(n) = \mathbb{P}(w_1 = i, w_2 = j)$ in the TASEP Ω_n . We will omit the dependence on n when it is clear from the context. The following lemma is a consequence of the projection principle.

Lemma 4.1. *For all $0 \leq s, t < n$ with $s + t \leq n$,*

$$T_{s,t} = \sum_{j=t+s+1}^n \sum_{i=s+1}^{s+t} E_{j,i}(n).$$

By a standard inclusion-exclusion argument, we have

$$E_{j,i}(n) = T_{i-1,j-i} - T_{i,j-i-1} - (T_{i-1,j-i+1} + T_{i,j-i}),$$

and hence it is enough to compute $T_{s,t}$ to know $E_{j,i}$. Computing E_{w_1,w_2} for small values of n reveals a clear pattern. See Table 1 for the values of $n \binom{n}{2} E_{w_1,w_2}$ when $n = 5$. The pattern in the lower left triangle of this table is easy to spot, and for larger n also the pattern in the upper right triangle.

TABLE 1. Values of $n \binom{n}{2} E_{w_1,w_2}$ for $n = 5$.

$w_1 \setminus w_2$	1	2	3	4	5
1	0	4	2	2	2
2	1	0	5	2	2
3	2	1	0	5	2
4	3	2	1	0	4
5	4	3	2	1	0

Theorem 4.2. *For any $1 \leq i \leq j \leq n$, we have*

$$(3) \quad \begin{aligned} E_{j,i} &= \frac{j-i}{n \binom{n}{2}}, \\ E_{i,j} &= \begin{cases} \frac{1}{n^2} + \frac{i(n-i)}{n^2(n-1)}, & \text{if } i = j-1, \\ \frac{1}{n^2}, & \text{if } i < j-1. \end{cases} \end{aligned}$$

Remark 4.3. Note that the probability that i is followed by j is the same for all $i < j - 1$. Is there an easy explanation for this? The probabilities $\mathbb{P}(w_1 = j) = \mathbb{P}(w_2 = i) = \frac{1}{n}$ and hence the probability $E_{i,j} = 1/n^2$ could plausibly be interpreted as independence.

Before we go on to the proof, we recall some standard combinatorial definitions. A semistandard Young tableau (SSYT) [18] of shape λ is a partition of shape λ with positive integers in the boxes that are strictly increasing along columns and weakly increasing along rows. Set $\text{SSYT}_{r,l}(m)$ to be the number of semistandard Young tableaux on a shape of two columns of lengths $r \geq l$ with no number exceeding m . An application of the hook-content formula from standard combinatorial theory (see [18, Corollary 7.21.4]) leads to

$$(4) \quad \text{SSYT}_{r,k}(m) = \begin{cases} \frac{r-k+1}{r+1} \binom{m}{r} \binom{m+1}{k}, & \text{if } m \geq r \geq k \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proof of Theorem 4.2. We will start with the case $E_{j,i}$. As described above we will use the projection principle and study the three species system $\Omega_{m_s,t}$, where $m_{s,t} = (s, t, n - s - t)$. Since an easy calculation leads to

$$\sum_{j=t+s+1}^n \sum_{i=s+1}^{s+t} (j - i) = t(n - s)(n - t - s)/2,$$

it suffices to prove that

$$(5) \quad n \binom{n}{2} T_{s,t} = t(n - s)(n - t - s)/2.$$

We will compute $T_{s,t}$ using the theory of multiline queues by Ferrari and Martin explained in Section 2.1. In the chain $\Omega_{m_s,t}^{FM}$ we have s particles (occupied sites) on the first row and $s + t$ particles on the second row. Note that if $s + t = n$, then there are no 3's, so the formula (5) is trivially satisfied. Assume therefore that $s + t < n$ and $s, t \geq 1$. Any multiline queue that projects to a word starting with a 3 and a 2 in the second place must necessarily have the following structure:

$$(6) \quad \begin{array}{ccccccc} \circ & \circ & . & . & \dots & . & . \\ \circ & \bullet & . & . & \dots & . & . \\ \hline 3 & 2 & . & . & \dots & . & . \end{array}$$

That is, no first class particle can be standing in the queue to use the particle in the second position of the second row. This happens exactly if

$$\begin{array}{ccccccccccc} 3 & \leq & z_{1,3} & < & \dots & < & z_{1,t+2} & < & \dots & < & z_{1,n-s-1} & < & z_{1,n-s} & \leq & n \\ & & & & & & \vee & & & & & \vee & & \vee & & \\ & & 3 & \leq & z_{2,2} & < & \dots & < & z_{2,n-s-t-1} & < & z_{2,n-s-t}, & & & & \end{array}$$

where $z_{a,b}$ is the position of the b th vacant position in row a . To count these configurations, we use the connection with semistandard Young tableaux explained above.

The change of variables $\lambda_{1,b} = n + 1 - z_{1,n-s+1-b}$ and $\lambda_{2,b} = n + 1 - z_{2,n-s-t+1-b}$ shows that the number of multiline queues as shown in (6) is equivalent to $\text{SSYT}_{n-s-2,n-s-t-1}(n - 2)$ given in (4).

Since $|\Omega_{m_s,t}^{FM}| = \binom{n}{s} \binom{n}{s+t}$, we get

$$\begin{aligned} \binom{n}{s} \binom{n}{s+t} T_{s,t} &= \frac{t}{n - s - 1} \binom{n - 2}{n - s - 2} \binom{n - 1}{n - s - t - 1} \\ &= \binom{n}{s} \binom{n}{s+t} \frac{t(n - s)(n - s - 1)(n - s - t)}{(n - s - 1)n(n - 1)n}. \end{aligned}$$

This simplifies to

$$T_{s,t} = \frac{t(n-s)(n-s-t)}{n^2(n-1)},$$

as is wanted. There are two possibilities left. If $t = 0$, there are no 2's, and (5) is again trivially true. If $s = 0$ we have no 1's, t 2's and $n - t$ 3's. The first row is thus only vacant positions, and of all the $\binom{n}{t}$ ways of placing t occupied positions on row 2, there are $\binom{n-2}{t-1}$, which gives a word starting with a 3 and then a 2 as in (6). Thus

$$T_{0,t} = \frac{\binom{n-2}{t-1}}{\binom{n}{t}} = \frac{t(n-t)}{n(n-1)},$$

again satisfying (5).

The cases $j < i$ are derived analogously, which we describe briefly. We use the projection principle to the same Markov chain $\Omega_{m_s,t}$ as before. In this case we can sum up the $E_{j,i}$ to get

$$\sum_{j=t+s+1}^n \sum_{i=s+1}^{s+t} E_{j,i}(n) = \frac{(n-s-t)(s+tn)}{n^2(n-1)}.$$

We must thus prove that $\binom{n}{s} \binom{n}{s+t} T_{s,t}$, which is the number of multiline queues of the form

$$(7) \quad \begin{array}{ccccccc} & \circ & . & . & . & \dots & . \\ & \bullet & \circ & . & . & \dots & . \\ \hline & 2 & 3 & . & . & \dots & . \end{array}$$

is equal to $\binom{n}{s} \binom{n}{s+t} \frac{(n-s-t)(s+tn)}{n^2(n-1)}$. This splits into two cases:

- (i) There's no first class particle in queue above the 3,

$$\begin{array}{ccccccc} & \circ & \circ & . & . & \dots & . \\ & \bullet & \circ & . & . & \dots & . \\ \hline & 2 & 3 & . & . & \dots & . \end{array}$$

This is identical to the previous case and is enumerated by

$$\text{SSYT}_{n-s-2, n-s-t-1}(n-2) = \frac{t}{n-s-1} \binom{n-2}{n-s-2} \binom{n-1}{n-s-t-1}.$$

- (ii) There is a first class particle in queue above the 3,

$$\begin{array}{ccccccc} & \circ & \bullet & . & . & \dots & . \\ & \bullet & \circ & . & . & \dots & . \\ \hline & 2 & 3 & . & . & \dots & . \end{array}$$

Here we have one more vacant position to decide in the first row, which is the same as the first column of the SSYT being one longer. This case is therefore enumerated by

$$\text{SSYT}_{n-s-1, n-s-t-1}(n-2) = \frac{t+1}{n-s} \binom{n-2}{n-s-1} \binom{n-1}{n-s-t-1}.$$

Adding the result of these two cases gives the desired result $\binom{n}{s} \binom{n}{s+t} T_{s,t}$ as wanted. □

We now explain the connection to the TASEP speed process studied in [4]. To avoid complications, we do not explain the TASEP speed process itself, but an infinite species TASEP on \mathbb{Z} with the same stationary measure; see Theorem 1.5 in [4], which is defined as follows. There is a particle at each site, and each particle is assigned a species (or class) uniformly in $[-1, 1]$. Each particle carries an independent exponential clock with rate 1, and when the clock rings, the particle tries to exchange with the particle to its right. The exchange only succeeds when the particle to the right has a higher species value than itself. The authors computed various marginals in the stationary distribution of this process, e.g. they proved that the probability that a particle having higher, equal or lower class than its right neighbor is $1/3, 1/6$ and $1/2$ respectively; see [4, Theorem 1.7]. We have the following finite strengthening.

Corollary 4.4. *Let as before w_1, w_2 be the first two letters in the cyclic permutation in Ω_n . Then probability at stationarity is exactly*

$$\begin{aligned} \mathbb{P}[w_1 > w_2] &= \frac{1}{3} + \frac{1}{3n}, \\ \mathbb{P}[w_1 = w_2 - 1] &= \frac{1}{6} + \frac{7n - 6}{6n^2}, \\ \mathbb{P}[w_1 < w_2 - 1] &= \frac{1}{2} - \frac{3n - 2}{2n^2}. \end{aligned}$$

Proof. Sum over all appropriate j and i in Theorem 4.2. □

We believe that the n -species TASEP studied here, Ω_m with $m = (1, \dots, 1)$, converges as a Markov process to the infinite species TASEP explained above in the limit as n tends to infinity. In particular, the stationary distribution and all correlations should also converge. We now show that two-point nearest neighbour correlations do converge to their result starting from Theorem 4.2. This could alternatively be proven from the fact that the proofs in [4] also rely on projecting down to a three species model and studying the two row multiline queue. However, we choose to give a direct proof from the formulas.

Corollary 4.5 (Compare with Theorem 1.7 of [4]). *In the limit as $n \rightarrow \infty$, the probability distribution on $\{1, \dots, n\}^2$ given by E_{w_1, w_2} converges, after suitable rescaling, to the probability distribution on $[-1, 1]^2$ whose density is $f(x, y) + \mathbb{1}_{\{x=y\}}g(x)$, where $g(x) = \frac{1-x^2}{8}$ and*

$$f(x, y) = \begin{cases} \frac{1}{4}, & x > y, \\ \frac{y-x}{4}, & x < y. \end{cases}$$

Proof. We want to take the limit $w_1, w_2, n \rightarrow \infty$ at the same rate and rescale the resulting square so that it becomes $[-1, 1]^2$. Following the convention in Theorem 4.2, we will let i (resp. j) be the smaller (resp. larger) of w_1 and w_2 . We thus have to take the limit so that

$$\frac{i}{n} \rightarrow \frac{x+1}{2} \text{ and } \frac{j}{n} \rightarrow \frac{y+1}{2} \text{ as } n \rightarrow \infty.$$

A natural way of converting the probability mass function in (3) into a probability density function in $[-1, 1]^2$ is to divide the latter into n^2 smaller squares of area $4/n^2$ so that the (w_1, w_2) th square contributes E_{w_1, w_2} according to (3). This is naturally done by multiplying the values by $n^2/4$ and taking the limit $n \rightarrow \infty$.

When $w_1 = i < j - 1 = w_2 - 1$, $n^2 E_{i,j}/4$ directly gives $1/4$. On the other hand,

$$\frac{n^2}{4} E_{j,i} = \frac{j - i}{2(n - 1)} \xrightarrow{n \rightarrow \infty} \frac{y - x}{4},$$

again as desired. However, this procedure does not work for $E_{j-1,j}$ because that line carries a nontrivial fraction of the mass. (Notice that the scaling of $E_{j-1,j}$ in (3) goes like $1/n$ for j, n large.) The resulting continuous measure on $[-1, 1]^2$ is thus not absolutely continuous with respect to Lebesgue measure.

The correct way to take the limit is after multiplying by $\frac{n}{2}$ instead, leading to

$$\frac{n}{2} E_{j-1,j} = \frac{n - 1 + (j - 1)(n - j + 1)}{2n(n - 1)} \xrightarrow{n \rightarrow \infty} \frac{1 - y^2}{8},$$

which becomes the singular continuous part of the density on the diagonal, completing the proof. \square

5. TWO-POINT CORRELATIONS FURTHER APART

In this section we will study the correlation between two positions of the TASEP that are a fixed distance apart. Consistent with the previous section, let us denote

$$E_{j,i}(1, a) := \mathbb{P}[w_1 = j, w_a = i],$$

namely the probability that the first letter is j and the a th letter is i at stationarity in the multispecies TASEP. Although the notation is somewhat redundant, it avoids possible confusion between positions and labels. Without loss of generality, we can take $j > i$ because if $j < i$, the theorem gives a formula by the rotational symmetry in Proposition 2.1(i), namely $E_{j,i}(1, a) = E_{i,j}(1, n - a + 2)$. Our main result here is an explicit formula for $E_{j,i}(1, a)$ in Theorem 5.1.

An important ingredient in the study of two-point correlations is the following object, whose study we will undertake in Section 6:

$$Y_{r,l}^\beta(m) := \text{number of SSYTs on two columns } r \geq l \text{ with no entry exceeding } m \text{ such that the number } \beta \text{ appears somewhere in the second column.}$$

The proofs of the results in this section depend heavily on results which will be proven in Section 6.

Theorem 5.1. *For $j > i$,*

$$E_{j,i}(1, a) = \frac{Y_{n-i, n-j+1}^{a-1}(n-1)}{\binom{n}{i-1} \binom{n}{j-1}} - \frac{Y_{n-i, n-j}^{a-1}(n-1)}{\binom{n}{i-1} \binom{n}{j}} - \frac{Y_{n-i-1, n-j+1}^{a-1}(n-1)}{\binom{n}{i} \binom{n}{j-1}} + \frac{Y_{n-i-1, n-j}^{a-1}(n-1)}{\binom{n}{i} \binom{n}{j}}.$$

We postpone the proof of Theorem 5.1 for later but study some of its consequences for now. In the following situation, we get what is possibly the most striking result of this section.

Corollary 5.2. *For $1 \leq i < j \leq n$ and $a \leq j - i$ we have*

$$E_{i,j}(1, a) = \frac{1}{n^2}.$$

Remark 5.3. Note that $\mathbb{P}[w_1 = j] = \mathbb{P}[w_a = i] = \frac{1}{n}$, so this could be interpreted as an independence when the difference in value is larger than the difference in position.

We can follow the ideas of Corollary 4.5 and derive the probability density when $i, j, n \rightarrow \infty$ for a fixed position a . This directly leads to the proof of [4, Theorem 6.1(i)].

Corollary 5.4 (Theorem 6.1(i) of [4]). *The joint probability density function $f(x, y)$ for the two-point correlation $\mathbb{P}(w_1, w_a)$ when a is fixed, $w_1 < w_a$ and $w_1, w_a, n \rightarrow \infty$ is $1/4$, and consequently, $\mathbb{P}(w_1 < w_a) = 1/2$ in the limit.*

Corollary 5.5. *For $2 \leq a \leq n$ and $1 \leq i \leq n - 1$,*

$$E_{n,i}(1, a) = \frac{1}{n^2} + \frac{(n - i) \binom{i-1}{a-2} - \binom{i-1}{a-1}}{a n \binom{n}{a}}.$$

Remark 5.6. There are two simplifications of Corollary 5.5.

- (i) If $a > i + 1$, $E_{n,i}(1, a) = \frac{1}{n^2}$.
- (ii) If $i = n - 1$, $E_{n,n-1}(1, a) = \frac{a - 1}{n \binom{n}{2}}$.

We will assume that $j > i$ and follow similar arguments as in Section 4. This time we will project down to the three-species system with sector $m_{x,y} = (n - x - y, x, y)$ (note the slight difference with $m_{s,t}$ in Section 4). We will think of j as one of the 3's and i as one of the 2's. Let $D_{y,x}(1, a)$ be the probability that there is a 3 in the first position and a 2 in position a in this three-species system. Then by the projection principle we have

$$D_{y,x}(1, a) := \sum_{i=n-x-y+1}^{n-y} \sum_{j=n-y+1}^n E_{j,i}(1, a).$$

We can recover the $E_{j,i}(1, a)$'s from the $D_{y,x}(1, a)$'s by

$$(8) \quad \begin{aligned} E_{j,i}(1, a) &= D_{n-j+1, j-i}(1, a) - D_{n-j, j-i+1}(1, a) \\ &\quad - D_{n-j+1, j-i-1}(1, a) + D_{n-j, j-i}(1, a). \end{aligned}$$

We now claim the following.

Lemma 5.7. *If $x + y \leq n$, then*

$$D_{y,x}(1, a) = \frac{Y_{y+x-1, y}^{a-1} (n - 1)}{\binom{n}{y+x} \binom{n}{y}}.$$

In the special case $x + y = n$, $D_{y,x}(1, a)$ becomes independent of a ,

$$D_{y,x}(1, a) = \frac{y(n - y)}{n(n - 1)}.$$

Proof. The total number of multiline queues with sector $m_{y,x}$ is $\binom{n}{y+x} \binom{n}{y}$ since we can choose as vacant positions any $x + y$ positions in the first row and any y positions in the second row. Hence $\binom{n}{y+x} \binom{n}{y} \cdot D_{y,x}(1, a)$ is the number of multiline

we can replace a in both by $n - 1$. Therefore, using Proposition 6.3 below,

$$\begin{aligned} g_{j,i}(a) &= \frac{j Y_{j-1,i}^{n-1}(n-1) - (n+1-j) Y_{j-2,i}^{n-1}(n-1)}{n \binom{n-1}{j-1} \binom{n}{i}} \\ &= \frac{j \left[\binom{n-1}{i-1} \binom{n-1}{j-1} - \binom{n-2}{i-2} \binom{n}{j} \right] - (n+1-j) \left[\binom{n-1}{i-1} \binom{n-1}{j-2} - \binom{n-2}{i-2} \binom{n}{j-1} \right]}{n \binom{n-1}{j-1} \binom{n}{i}} \\ &= \frac{\binom{n-1}{i-1} \binom{n-1}{j-1}}{n \binom{n-1}{j-1} \binom{n}{i}} \\ &= \frac{i}{n^2}. \end{aligned}$$

Then $E_{j,i}(1, a) = g_{j,i}(a) - g_{j,i-1}(a)$ implies the result. □

Proof of Corollary 5.5. We use Theorem 5.1 in the case $j = n$. Notice that the second and fourth terms are zero because they correspond to SSYT with only one column. Therefore, we have

$$E_{n,i}(1, a) = \frac{Y_{n-i,1}^{a-1}(n-1)}{n \binom{n}{i-1}} - \frac{Y_{n-i-1,1}^{a-1}(n-1)}{n \binom{n}{i}}.$$

Let us look at the first term. The second term is similar. We have to count SSYT with two columns whose column lengths are $n - i$ and 1, whose maximum entry is $n - 1$ and where $a - 1$ sits in the second column. Among all $\binom{n-1}{n-i}$ choices for the entries in the first column, we can't accept the ones where the minimal entry is greater than $a - 1$. Thus,

$$Y_{n-i,1}^{a-1}(n-1) = \binom{n-1}{n-i} - \binom{n-a}{n-i},$$

which leads to

$$E_{n,i}(1, a) = \frac{\binom{n-1}{n-i} - \binom{n-a}{n-i}}{n \binom{n}{i-1}} - \frac{\binom{n-1}{n-i-1} - \binom{n-a}{n-i-1}}{n \binom{n}{i}}.$$

After a few manipulations, this leads to

$$E_{n,i}(1, a) = \frac{1}{n^2} + \frac{1}{na \binom{n}{a}} \left[(n-i) \binom{i}{a-1} - (n+1-i) \binom{i-1}{a-1} \right],$$

which is equivalent to what we wanted to prove. □

6. CONSTRAINED SEMISTANDARD YOUNG TABLEAUX

In this section we compute the number of semistandard Young tableaux with two columns of lengths $r \geq l$, with maximum entry m and such that a fixed β lies in the second column. We denoted this by $Y_{r,l}^\beta(m)$ in Section 5. To that end, we introduce the following quantities. Similarly, let $Z_{r,l}^{\alpha,\beta}(m)$ be the number of SSYTs of shape r, l , with no entry exceeding m such that the first row is (α, β) . Finally, let $X_r^{\alpha,\beta}$ be the number of SSYTs of shape r, r , with entries (α, β) in the last row. We illustrate these objects in Figure 5.

We first give enumerative results for both the latter quantities.

Lemma 6.1. *For all $r \leq \alpha \leq \beta$ we have $X_r^{\alpha,\beta} = \binom{\beta}{r-1} \binom{\alpha-1}{r-1} - \binom{\beta-1}{r-2} \binom{\alpha}{r}$.*

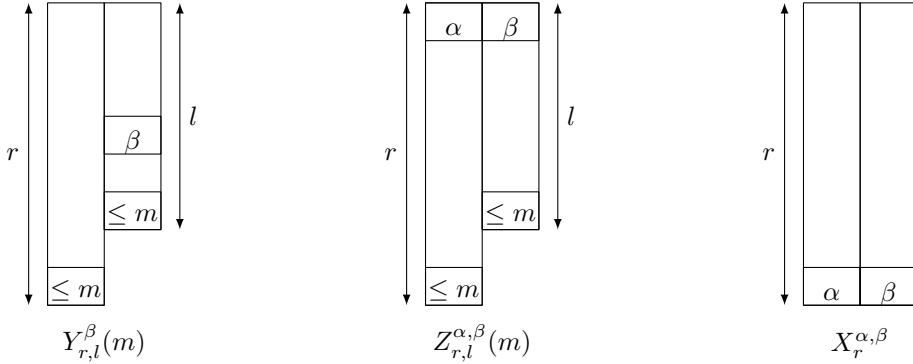


FIGURE 5. Illustration of the various constrained two-column tableaux.

Proof. The lemma is clearly true for $r = 1$ since $X_1^{\alpha,\beta} = 1$; as usual we define $\binom{n}{-1} = 0$. The lemma then follows by induction over r using the recursion $X_{r+1}^{\alpha,\beta} = \sum_{a=r}^{\alpha-1} \sum_{b=a}^{\beta-1} X_r^{a,b}$. \square

Using the obvious identity $Z_{r,l}^{\alpha,\beta}(m) = Z_{r,l}^{1,\beta-\alpha+1}(m - \alpha + 1)$, it suffices to give a formula for $Z_{r,l}^{1,\beta}(m)$.

Lemma 6.2. *For all $\beta \leq m, l \leq r \leq m$ we have*

$$Z_{r,l}^{1,\beta}(m) = \sum_{1 \leq i \leq j \leq \beta} (-1)^{i+1} \binom{j-2}{i-2} \binom{\beta-j+i-1}{i-1} \text{SSYT}_{r-i,l-i}(m-j).$$

Proof. We will keep r, l, m fixed throughout the proof and introduce the shorter notation $\text{SS}_i(j) := \text{SSYT}_{r-i,l-i}(m-j)$.

The recursion

$$(10) \quad Z_{r,l}^{1,\beta}(m) = \text{SS}_1(1) - \sum_{a=2}^{\beta} \sum_{b=a}^{\beta} Z_{r-1,l-1}^{a,b}(m)$$

follows easily by studying which pairs of numbers $a \leq b$ cannot be the entries in the second row and subtracting those cases. Now, let us use the ansatz

$$(11) \quad Z_{r,l}^{1,\beta}(m) = \sum_{1 \leq i \leq j \leq \beta} g_{i,j}^{\beta} \cdot \text{SS}_i(j)$$

for some coefficients $g_{i,j}^{\beta}$. Plugging the ansatz (11) into the recursion (10) translates, after a few steps, to

$$\sum_{1 \leq i \leq j \leq \beta} g_{i,j}^{\beta} \cdot \text{SS}_i(j) = \text{SS}_1(1) - \sum_{2 \leq a \leq b \leq \beta} \sum_{1 \leq i \leq j \leq b-a+1} g_{i,j}^{b-a+1} \cdot \text{SS}_{i+1}(j+a-1).$$

This forces $g_{1,1}^{\beta} = 1$ and gives the recursion

$$g_{s,t}^{\beta} = - \sum_{r=s-1}^{t-1} \sum_{u=r}^{r+\beta-t} g_{s-1,r}^u.$$

Induction over s now easily proves

$$g_{s,t}^\beta = (-1)^{s+1} \binom{t-2}{s-2} \binom{\beta-t+s-1}{s-1},$$

which is the desired coefficient. □

We now consider special values of β , where the formula for $Y_{r,l}^\beta(m)$ is simple.

Proposition 6.3. (i) $Y_{r,l}^1(m) = \text{SSYT}_{r-1,l-1}(m-1)$.

(ii)

$$Y_{r,l}^m(m) = \binom{m}{l-1} \binom{m}{r} - \binom{m-1}{l-2} \binom{m+1}{r+1}.$$

Proof. (i) The first row has to be $(1, 1)$, and hence the minimum entry is 2 in the remainder of the SSYT.

(ii) $\beta = m$ can only be present in the second column in the last row. Let us sum over all possible entries in the first column at the same row. Using the definition of $X_r^{\alpha,\beta}$ and $Z_{r,l}^{\alpha,\beta}(m)$, we get

$$Y_{r,l}^m(m) = \sum_{y=1}^m X_l^{y,m} Z_{r-l+1,1}^{y,m}(m).$$

Since $Z_{r,1}^{\alpha,m}(m)$ is just the binomial coefficient $\binom{m-y}{r-l}$, we get

$$\begin{aligned} Y_{r,l}^m(m) &= \sum_{y=1}^m \binom{m-y}{r-l} \left(\binom{m}{l-1} \binom{y-1}{l-1} - \binom{m-1}{l-2} \binom{y}{l} \right), \\ &= \binom{m}{l-1} \sum_{y=1}^m \binom{m-y}{r-l} \binom{y-1}{l-1} - \binom{m-1}{l-2} \sum_{y=1}^m \binom{m-y}{r-l} \binom{y}{l}, \\ &= \binom{m}{l-1} \binom{m}{r} - \binom{m-1}{l-2} \binom{m+1}{r+1}. \end{aligned}$$

□

Theorem 6.4. For $m \geq r \geq l \geq 1$ and $1 \leq \beta < m$ we have

$$Y_{r,l}^\beta(m) = \sum_{1 \leq f \leq e \leq \beta} N_{e-1,f-1} \cdot \text{SSYT}_{r-f,l-f}(m-e),$$

where

$$N_{e,f} = \frac{1}{e} \binom{e}{f+1} \binom{e}{f}$$

are the Narayana numbers.

Proof. Summing over in the row x of the last column that contains β and the number y that is in row x of the first column, we get the recursion

$$\begin{aligned} Y_{r,l}^\beta(m) &= \sum_{x=1}^l \sum_{y=1}^\beta X_x^{y,\beta} Z_{r-x+1,l-x+1}^{y,\beta}(m), \\ &= \sum_{x=1}^l \sum_{y=1}^\beta X_x^{y,\beta} Z_{r-x+1,l-x+1}^{1,\beta-y+1}(m-y+1), \end{aligned}$$

where we have used Lemma 6.2 in the second line. This leads to

$$Y_{r,l}^\beta(m) = \sum_{x=1}^l \sum_{y=1}^\beta \left[\binom{\beta}{x-1} \binom{y-1}{x-1} - \binom{\beta-1}{x-2} \binom{y}{x} \right] \\ \times \sum_{1 \leq i \leq j \leq \beta-y+1} (-1)^{i+1} \binom{j-2}{i-2} \binom{\beta-y-j+i}{i-1} \text{SS}_{x+i-1}(y+j-1),$$

where we have reused the notation $\text{SS}_i(j)$ from Lemma 6.2 and fixed r, l, m .

Note that all except the y -sum are natural in the sense that we can replace the limits by \mathbb{Z} . The terms outside of the specified limits will then be zero by definition of the binomial coefficient.

As a first step, replace i by $f = x + i - 1$ and j by $e = y + j - 1$ to get

$$Y_{r,l}^\beta(m) = \sum_x \sum_{y=1}^\beta \left[\binom{\beta}{x-1} \binom{y-1}{x-1} - \binom{\beta-1}{x-2} \binom{y}{x} \right] \\ \times \sum_f \sum_e (-1)^{f+x} \binom{e-y-1}{f-x-1} \binom{\beta-e+f-x}{f-x} \text{SS}_f(e).$$

We are now going to move the x, y sums inside the e, f sums

$$Y_{r,l}^\beta(m) = \sum_f \sum_e \sum_x \sum_{y=1}^\beta (-1)^{f+x} \left[\binom{\beta}{x-1} \binom{y-1}{x-1} - \binom{\beta-1}{x-2} \binom{y}{x} \right] \\ \times \binom{e-y-1}{f-x-1} \binom{\beta-e+f-x}{f-x} \text{SS}_f(e).$$

Now the last binomial coefficient is zero unless $\beta \geq e + x - f$, and the one before that is zero unless $e - y \geq f - x$. Therefore, we can replace the upper limit of the y -sum by $e + x - f$. Now we are in a position to do the y -sums. We use the identity

$$\sum_{m=0}^n \binom{m}{j} \binom{n-m}{k-j} = \binom{n+1}{k+1}, \text{ if } 0 \leq j \leq k \leq n,$$

to obtain

$$\sum_{y=1}^{e+x-f} \binom{e-y-1}{f-x-1} \binom{y-1}{x-1} = \binom{e-1}{f-1}, \\ \sum_{y=1}^{e+x-f} \binom{e-y-1}{f-x-1} \binom{y}{x} = \binom{e}{f}.$$

Note that it does not matter if the sum runs to $e + x - f$ or e . Thereafter we are left with

$$Y_{r,l}^\beta(m) = \sum_f \sum_e \text{SS}_f(e) \\ \times \sum_x (-1)^{f+x} \binom{\beta-e+f-x}{f-x} \left[\binom{\beta}{x-1} \binom{e-1}{f-1} - \binom{\beta-1}{x-2} \binom{e}{f} \right].$$

Now the x -sum can be done by first using $\binom{n}{k} = (-1)^k \binom{k-n-1}{k}$ and then variants of the Chu-Vandermonde identity,

$$\sum_{x=1}^f (-1)^x \binom{\beta - e + f - x}{f - x} \binom{\beta}{x - 1} = (-1)^f \binom{e - 1}{f - 1},$$

$$\sum_{x=2}^f (-1)^{x+1} \binom{\beta - e + f - x}{f - x} \binom{\beta - 1}{x - 2} = (-1)^{f+1} \binom{e - 2}{f - 2}.$$

Noting that

$$\binom{e - 1}{f - 1}^2 - \binom{e}{f} \binom{e - 2}{f - 2} = N_{e-1, f-1},$$

the Narayana number, completes the proof. □

Corollary 6.5. *For $l \leq r$ and $l + m - r \leq \beta \leq m$, $Y_{r,l}^\beta(m)$ is independent of β .*

Proof. We use the formula for $Y_{r,l}^\beta(m)$ in Theorem 6.4. For clarity, we rewrite the formula and explicitly include the binomial coefficients,

$$Y_{r,l}^\beta(m) = \sum_{1 \leq f \leq e \leq \beta} \frac{r - l + 1}{(e - 1)(r - f + 1)} \binom{e - 1}{f} \binom{e - 1}{f - 1} \binom{m - e}{r - f} \binom{m - e + 1}{l - f}.$$

First of all, $f \leq l$ by the last binomial coefficient. More importantly, $e \leq m - r + f$ by the penultimate one. Using the constraint on β , we get that

$$m + f - r \leq m + l - r \leq \beta.$$

Therefore we can replace the upper limit of e by $m + l - r$ in the summation. □

Remark 6.6. There is a nice combinatorial way of proving Corollary 6.5 directly. Given $l + m - r < \beta \leq m$ we can define a bijection from $Y_{r,l}^\beta(m)$ to $Y_{r,l}^{\beta-1}(m)$ as follows. Assume $\lambda \in Y_{r,l}^\beta(m)$ with $\lambda_{2,b} = \beta$ for some row $b \leq l$. There are two cases. If $\lambda_{2,b-1} = \beta - 1$, then λ is mapped to itself. If $\lambda_{2,b-1} < \beta - 1$, then we map λ to the partition $\lambda' \in Y_{r,l}^{\beta-1}(m)$ which differs from λ only by $\lambda'_{2,b} = \beta - 1$. This can always be done since $\lambda_{1,b} \leq \lambda_{1,l} \leq l + m - r < \beta$ by assumption. The inverse of the bijection is defined in the same way.

7. THREE-POINT CORRELATIONS

In line with the work of Amir, Angel and Valkó on the infinite species TASEP, we now present results for three-point correlations. Just as in [4, Section 7.4], we will only prove results for three consecutive particles. Without loss of generality, we calculate the probability of particles of three different species being at sites 1, 2 and 3. What differentiates these cases is the relative order of the species. For convenience we will denote the lowest of these species by i , the next by j , and the highest by k . It is then clear that there are $3! = 6$ different possibilities, the permutations of the letters i, j, k . Let π be any such permutation. We fix n and denote $E_\pi(n) \equiv E_\pi$ to be the probability in the stationary distribution of the TASEP to have the π_i th particle at site i for $i = 1, 2, 3$. Clearly, $E_{a,b,c}$ is zero if any two of a, b or c are equal, and hence can be thought of as a probability distribution on $\{1, \dots, n\}^3$.

The proofs here will, by analogy with two-point correlations in Section 4, sum over the number of semistandard Young tableau (SSYT) with three columns of

lengths $a \geq b \geq c$. See the paragraph below Remark 4.3 for the general definition of SSYT. We will repeatedly use the following formula that is easy to deduce from the hook-content formula (see [18, Corollary 7.21.4]):

$$(12) \quad \text{SSYT}_{a,b,c}(m) = \begin{cases} \frac{(a-b+1)(a-c+2)(b-c+1)}{(a+1)(a+2)(b+1)} \\ \times \binom{m}{a} \binom{m+1}{b} \binom{m+2}{c}, & \text{if } m \geq a \geq b \geq c \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

We will also follow the standard convention that $\text{SSYT}_{0,0,0}(m) = 1$ for all $m \geq 0$.

The proofs in this section will follow the same lines as the proof of Theorem 4.2. Here we will project to the four-species system $\Omega_{m,r,s,t}$, where $m = (r, s, t, n - r - s - t)$. That is, to a system with r 1s, s 2s, t 3s and $n - r - s - t$ particles of class 4. The case with the most uniform answer is the decreasing case, corresponding to the permutation 321.

Theorem 7.1. *For $i < j < k$,*

$$E_{k,j,i} = \frac{6(j-i)(k-i)(k-j)}{n^3(n-1)^2(n-2)}.$$

Proof. Let $T_{r,s,t} = \mathbb{P}(w_1 = 4, w_2 = 3, w_3 = 2)$ be the probability at stationarity in the Markov chain $\Omega_{m,r,s,t}$ of having the first positions occupied by 4, 3, 2. We will compute $T_{r,s,t}$ in two different ways. First by the projection principle we have for all $0 \leq r, s, t < n$ with $r + s + t \leq n$,

$$(13) \quad T_{r,s,t} = \sum_{k=r+s+t+1}^n \sum_{j=r+s+1}^{r+s+t} \sum_{i=r+1}^{r+s} E_{k,j,i}.$$

We can evaluate the sum explicitly by substituting the value of the summand. Thus it suffices to prove that

$$(14) \quad T_{r,s,t} = \sum_{k,j,i} \frac{6(j-i)(k-i)(k-j)}{n^3(n-1)^2(n-2)} = \frac{st(s+t)(n-r)(n-r-s)(n-r-s-t)}{n^3(n-1)^2(n-2)}.$$

Note that we can, by inclusion-exclusion, obtain the $E_{k,j,i}$'s from the $T_{r,s,t}$'s.

Second we can compute $T_{r,s,t}$ by counting the number of multiline queues that gives a word starting with 4, 3, 2. The only possibility is

$$(15) \quad \begin{array}{cccccccc} \circ & \circ & \circ & . & \dots & . & & \\ \circ & \circ & \bullet & . & \dots & . & & \\ \circ & \bullet & \bullet & . & \dots & . & & \\ \hline 4 & 3 & 2 & . & \dots & . & & \end{array}$$

No particles may be queueing in the beginning of this multiline queue, and thus we get the following set of inequalities. Let $z_{x,y}$ be the position of the x th vacant position in row y :

$$\begin{array}{ccccccccccc} 4 \leq z_{1,4} < & \dots & < z_{1,s+3} < & \dots & < z_{1,s+t+2} < & \dots & < z_{1,n-r} & \leq n \\ & & \vee & \dots & \vee & \dots & \vee & & & & \\ & 4 & \leq z_{2,3} < & \dots & < z_{2,t+2} < & \dots & < z_{2,n-r-s} & \leq n \\ & & & & \vee & \dots & \vee & & & & \\ & & & & 4 & \leq z_{3,2} < & \dots & < z_{3,n-r-s-t} & \leq n. \end{array}$$

As in the proof of Theorem 4.2 the number of possible values of the $z_{x,y}$ is counted by the number of semistandard Young tableaux, here

$$\text{SSYT}_{n-r-3, n-r-s-2, n-r-s-t-1}(n-3).$$

The total number of multiline queues is $\binom{n}{r}\binom{n}{r+s}\binom{n}{r+s+t}$, since we may choose the r occupied positions in the first row, the $r + s$ occupied positions in the second row and the $r + s + t$ occupied positions in the third row in all possible ways. Thus the probability $T_{r,s,t}$ is the number of SSYT divided by this product of binomials. Using (12) we get

$$\begin{aligned} T_{r,s,t} &= \frac{\binom{n-3}{r}\binom{n-2}{r+s}\binom{n-1}{r+s+t}}{\binom{n}{r}\binom{n}{r+s}\binom{n}{r+s+t}} \frac{st(s+t)}{(n-r-2)(n-r-1)(n-r-s-1)} \\ &= \frac{(n-r)(n-r-s)(n-r-s-t)st(s+t)}{n(n-1)(n-2)n(n-1)n}, \end{aligned}$$

which is equal to (14).

To complete the proof, one needs to check that the special cases $r = 0, s = 0$ and $t = 0$ also satisfy the identity. In principle, these need to be done separately because these will involve projections to multiline queues with less than 4 queues. When s or t are zero, $T_{r,s,t}$ is identically zero as needed. The case of $r = 0$ needs to be checked separately, and this can be done just as in the special case of Lemma 5.7. \square

The next case we consider is the permutation 213. This also gives us the answer for the permutation 132 by particle-hole symmetry Proposition 2.1(ii), $E_{j,i,k} = E_{n+1-k, n+1-i, n+1-j}$.

Theorem 7.2. For $i < j < k$,

$$E_{j,i,k} = \begin{cases} \frac{2(j-i)}{n^3(n-1)}, & k > j+1, \\ \frac{2(j-i)}{n(n-1)} \left(\frac{1}{n^2} + \frac{j(n-j)}{n^2(n-1)} \right) \\ \quad + \frac{2j(j-1)(n-j)}{n^3(n-1)^2(n-2)}, & k = j+1. \end{cases}$$

Proof. This time we define $T_{r,s,t} = \mathbb{P}(w_1 = 3, w_2 = 2, w_3 = 4)$. By the projection principle we can first use

$$(16) \quad T_{r,s,t} = \sum_{k=r+s+t+1}^n \sum_{j=r+s+1}^{r+s+t} \sum_{i=r+1}^{r+s} E_{j,i,k}.$$

Summing over the expression we want to prove we obtain, using a computer algebra package,

$$(17) \quad T_{r,s,t} = \frac{s(n-r-s-t)(n^2st+n^2t^2+2nrt+nrs+ns^2-ns-nt-nr-nt^2+2rs+2r^2)}{n^3(n-1)^2(n-2)}.$$

Second, we need to count the number of multiline queues that makes the TASEP word start with 324. There are four possible different configurations:

$$\begin{array}{cc}
 \begin{array}{ccccccc}
 \circ & \circ & \circ & . & \dots \\
 \circ & \bullet & \circ & . & \dots \\
 \bullet & \bullet & \circ & . & \dots \\
 \hline
 3 & 2 & 4 & . & \dots
 \end{array} & , & & & \begin{array}{ccccccc}
 \circ & \circ & \bullet & . & \dots \\
 \circ & \bullet & \circ & . & \dots \\
 \bullet & \bullet & \circ & . & \dots \\
 \hline
 3 & 2 & 4 & . & \dots
 \end{array} \\
 \\
 \begin{array}{ccccccc}
 \circ & \circ & \circ & . & \dots \\
 \circ & \bullet & \bullet & . & \dots \\
 \bullet & \bullet & \circ & . & \dots \\
 \hline
 3 & 2 & 4 & . & \dots
 \end{array} & , & & & \begin{array}{ccccccc}
 \circ & \circ & \bullet & . & \dots \\
 \circ & \bullet & \bullet & . & \dots \\
 \bullet & \bullet & \circ & . & \dots \\
 \hline
 3 & 2 & 4 & . & \dots
 \end{array}
 \end{array}$$

These are, by arguments similar to above, counted by

$$(18) \quad \sum_{x=0}^1 \sum_{y=0}^1 \text{SSYT}_{n-r-2-x, n-r-s-1-y, n-r-s-t-1}(n-3).$$

Thus, using (12), we get

$$\begin{aligned}
 & \binom{n}{r} \binom{n}{r+s} \binom{n}{r+s+t} T_{r,s,t} \\
 &= \sum_{x=0}^1 \sum_{y=0}^1 \binom{n-3}{r+x-1} \binom{n-2}{r+s+y-1} \binom{n-1}{r+s+t} \\
 & \quad \times \frac{(s+y-x)(t-y+1)(s+t-x+1)}{(n-r-x-1)(n-r-x)(n-r-s-y)}.
 \end{aligned}$$

Performing the sums and simplifying, $T_{r,s,t}$ becomes equal to the expression in (17), completing the proof. □

The last case for which we can prove the formula for correlations is 312, and by particle-hole symmetry Proposition 2.1(ii) 231, that is, $E_{j,k,i} = E_{n+1-i, n+1-k, n+1-j}$.

Theorem 7.3. For $i < j < k$,

$$E_{j,k,i} = \begin{cases} \frac{3(j-i)(2n-j-i-1)}{n^3(n-1)(n-2)} - \frac{4(j-i)(n-k)}{n^3(n-1)^2}, & k > j+1, \\ \frac{(j-i)(n-1-j)}{n^2(n-1)^2} \left(\frac{1}{n-2} + \frac{3(n-i-1)}{n} \right. \\ \left. - \frac{(n-1-j)(3n-3i+j-1)}{n(n-2)} \right) + \frac{6(j-i)(n-i)}{n^3(n-1)(n-2)}, & k = j+1. \end{cases}$$

Proof. This time we define $T_{r,s,t} = \mathbb{P}(w_1 = 3, w_2 = 4, w_3 = 2)$. Summing $E_{j,k,i}$'s just as we did in (13), we again get, using a standard computer algebra package,

$$(19) \quad \begin{aligned}
 T_{r,s,t} = & s(n-r-s-t) \left(2n^2st + 2n^2t^2 + nrs - ns - nt - nt^2s - ns^2t \right. \\
 & - 2nstr + 2nrt + ns^2 - 2nt^2r - nt^2 - nr + r^2 \\
 & \left. + rs - 2trs - r^2s - rs^2 - 2tr^2 \right) / \left(n^3(n-1)^2(n-2) \right).
 \end{aligned}$$

Now, we need to count the number of multiline queues that make the TASEP word start with 342. There are four possible configurations:

$$\begin{array}{ccc}
 \begin{array}{cccc}
 \circ & \circ & \circ & . \dots \\
 \circ & \bullet & \circ & . \dots \\
 \bullet & \circ & \bullet & . \dots \\
 \hline
 3 & 4 & 2 & . \dots
 \end{array} & , &
 \begin{array}{cccc}
 \circ & \circ & \circ & . \dots \\
 \circ & \circ & \bullet & . \dots \\
 \bullet & \circ & \bullet & . \dots \\
 \hline
 3 & 4 & 2 & . \dots
 \end{array} \\
 \\
 \begin{array}{cccc}
 \circ & \circ & \bullet & . \dots \\
 \circ & \bullet & \circ & . \dots \\
 \bullet & \circ & \bullet & . \dots \\
 \hline
 3 & 4 & 2 & . \dots
 \end{array} & , &
 \begin{array}{cccc}
 \circ & \circ & \circ & . \dots \\
 \circ & \bullet & \bullet & . \dots \\
 \bullet & \circ & \bullet & . \dots \\
 \hline
 3 & 4 & 2 & . \dots
 \end{array}
 \end{array}$$

Here, as before, no particle can be in the queue in any of the rows in the first column. These are counted by

$$\begin{aligned}
 & 2 \text{SSYT}_{n-r-3, n-r-s-2, n-r-s-t-1}(n-3) \\
 & + \text{SSYT}_{n-r-2, n-r-s-2, n-r-s-t-1}(n-3) \\
 & + \text{SSYT}_{n-r-3, n-r-s-1, n-r-s-t-1}(n-3).
 \end{aligned}$$

Using (12) and dividing by $\binom{n}{r} \binom{n}{r+s} \binom{n}{r+s+t}$ we get

$$T_{r,s,t} = \frac{(n-r)(n-r-s)(n-r-s-t)}{n^3(n-1)^2(n-2)} \left(2st(s+t) + \frac{(s+1)t(s+t+1)r}{n-r} + \frac{(s-1)(t+1)(s+t)}{n-r-s} \right),$$

and a simple calculation shows that this is equal to (19). □

For the increasing case we have not been able to compute the correlations using this method. This is in one sense the most interesting formula because the probability is completely independent of i, j, k when they are far apart. This generalizes the behaviour of $E_{i,j}$ in Theorem 4.2.

To prove this conjecture with the same approach as for the other patterns, we would need a formula for the number of SSYT and for the number of near-SSYT, that is, tableaux that are not SSYT but become SSYT if the last column is moved one step up (and the top value deleted).

Conjecture 7.4. For $i < j < k$,

$$E_{i,j,k} = \begin{cases} \frac{1}{n^3}, & i < j - 1 < k - 2, \\ \frac{n - 1 + i(n - i)}{n^3(n - 1)}, & i = j - 1 < k - 2, \\ \frac{n - 1 + j(n - j)}{n^3(n - 1)}, & i < j - 1 = k - 2, \\ \frac{(n - 1 + i(n - i))(n - 1 + (i + 1)(n - i - 1))}{n^3(n - 1)^2} \\ \quad + \frac{2i(i + 1)(n - i)(n - i - 1)}{n^3(n - 1)^2(n - 2)}, & i = j - 1 = k - 2. \end{cases}$$

Note added in proof. Since the submission of this manuscript Conjecture 7.4 has been proven by Erik Aas [1]. The first and second cases follow since he has proved Conjecture 8.3 below. The third case follows by symmetry from the second, and finally the fourth case since for fixed i and j we know what the sum over k must be by Theorem 4.2. □

We can obtain, as in Corollary 4.4, a finite strengthening for the probability in the 9 cases corresponding to Theorems 7.1, 7.2 and 7.3 and conjecture one for the four cases of Conjecture 7.4. The proof is just a summation over all possibilities for each case. These probabilities are not written down in [4, Theorem 7.7] but can be easily obtained by integrating the densities therein.

Corollary 7.5. *Assuming Conjecture 7.4, the probability of all cases for three adjacent positions in the TASEP on permutations is given by Table 2. It is assumed that $i < j < k$ and the more general cases in the beginning do not cover the more specific cases later.*

Using Theorems 7.1, 7.2 and 7.3 and assuming Conjecture 7.4, one may also deduce as a corollary similar to Corollary 4.5, the joint distributions of three consecutive particles in the infinite multispecies TASEP process [4, Theorem 7.7]. This should come as no surprise; the proof in [4] also amounts to projecting the TASEP to four-particle systems and then studying the multiline queues.

Corollary 7.6 (Amir, Angel and Valkó [4]). *Assuming Conjecture 7.4, in the limit as $n \rightarrow \infty$, the probability distribution on $\{1, \dots, n\}^3$ given by E_{w_1, w_2, w_3} converges, after suitable rescaling, to the probability distribution on $[-1, 1]^3$ whose density is exactly that given in Table 2 of [4].*

Proof. The idea of the proof is essentially identical to that of Corollary 4.5 and we will not repeat all the details. In the first part of the table, one obtains the density by multiplying the probability mass functions by $(n/2)^3$ and taking the limit. In the second and third parts, one does the same, except that the prefactor becomes $(n/2)^2$ and $(n/2)$ respectively. □

TABLE 2. Probability of three adjacent positions by their relative size; see Corollary 7.5.

(w_1, w_2, w_3)	Probability	Follows from
(k, j, i)	$\frac{(n+1)(n+2)}{30n(n-1)} = \frac{1}{30} + \frac{2n+1}{15n(n-1)}$	Theorem 7.1
(i, k, j) & (j, i, k)	$\frac{(n-2)(n-3)}{12n^2} = \frac{1}{12} - \frac{5n-6}{12n^2}$	Theorem 7.2
$(i, k, i+1)$ & $(j, i, j+1)$	$\frac{n^3+8n^2-23n+10}{20n^2(n-1)} = \frac{1}{20} + \frac{9n^2-23n+10}{20n^2(n-1)}$	Theorem 7.2
(k, i, j) & (j, k, i)	$\frac{(n+1)(n-3)(7n-10)}{60n^2(n-1)} = \frac{7}{60} - \frac{17n^2+n-30}{60n^2(n-1)}$	Theorem 7.3
$(k, i, i+1)$ & $(j, j+1, i)$	$\frac{(n+1)(n^2+7n-10)}{20n^2(n-1)} = \frac{1}{20} + \frac{9n^2-3n-10}{20n^2(n-1)}$	Theorem 7.3
(i, j, k)	$\frac{(n-2)(n-3)(n-4)}{6n^3} = \frac{1}{6} + \frac{9n^2-26n+24}{6n^3}$	Conjecture 7.4
$(i, i+1, k)$ & $(i, j, j+1)$	$\frac{(n-2)(n-3)(n+8)}{12n^3} = \frac{1}{12} + \frac{3n^2-34n+48}{12n^3}$	Conjecture 7.4
$(i, i+1, i+2)$	$\frac{n^4+13n^3+32n^2-160n+120}{30n^3(n-1)} = \frac{1}{30} + \frac{7n^3+16n^2-80n+60}{15n^3(n-1)}$	Conjecture 7.4

8. DISCUSSIONS AND OPEN PROBLEMS

As expected, the correlation of several particles seems intractable in general. We can however state one fact and several conjectures. We denote, just as in Section 7, the general nearest neighbour correlation E_{i_1, \dots, i_r} to be the joint probability of seeing i_a in position a , for $a = 1, \dots, r$. For the decreasing case we conjectured the following Vandermonde formula.

Conjecture 8.1. For $r \leq n$ and $i_1 > \dots > i_r$,

$$E_{i_1, i_2, \dots, i_r} = r! \frac{\prod_{1 \leq a < b \leq r} (i_a - i_b)}{r-1 \prod_{i=0}^{i_1} (n-i)^{r-i}}$$

This conjecture has now been proven using determinantal techniques in [2], but it is natural to ask for a simpler proof of this result.

The writing of this paper was delayed partly because we really wanted to prove the $1/n^3$ formula in Conjecture 7.4. It is interesting that the simplest three-point correlation formula has proved impervious to our proof technique. The $1/n^3$ formula also strongly suggests that there is some more conceptual independence to be discovered and proved for the cyclic multispecies TASEP. In the infinite species TASEP [4], the independence of certain events is obvious. However it is not guaranteed in the finite case, but still mysteriously seems to be present. We offer a sequence of conjectures generalizing this observation.

Conjecture 8.2 (Proved in [1]). For $i_1 < i_2 - 1 < i_3 - 2 < \dots < i_r - (r - 1)$,

$$E_{i_1, \dots, i_r} = \frac{1}{n^r}$$

Note that the first formula in Theorem 7.2 says $E_{j,i,k} = E_{j,i} \cdot \frac{1}{n}$ if $j + 1 < k$. This naturally suggests the following generalization of Conjecture 8.2.

Conjecture 8.3 (Proved in [1]). *For any i_1, i_2, \dots, i_r and $k > 1 + \max_a i_a$, then*

$$E_{i_1, \dots, i_r, k} = \frac{1}{n} \cdot E_{i_1, \dots, i_r}.$$

Note that this is a finite cyclic version of [4, Lemma 7.3]. Inspired by our study in Section 5 of the correlation of particles further apart, in particular Corollary 5.2, there is reason to believe the following natural generalization of Conjecture 8.3.

Conjecture 8.4. *For any i_1, i_2, \dots, i_r and b, k such that $k > b - r + \max_a i_a$, then*

$$\mathbb{P}(w_a = i_a \text{ for } 1 \leq a \leq r, \text{ and } w_b = k) = \frac{1}{n} \cdot E_{i_1, \dots, i_r}.$$

We end with the most general independence for nearest neighbour correlations in two blocks.

Conjecture 8.5. *For any $i_1, \dots, i_r, j_{r+1}, \dots, j_{r+s}$ such that for all $r+1 \leq b \leq r+s$ and $\min_b j_b > 1 + \max_a i_a$, we have the independence*

$$\mathbb{P}(w_a = i_a, 1 \leq a \leq r, \text{ and } w_b = j_b, r+1 \leq b \leq r+s) = E_{i_1, \dots, i_r} E_{j_{r+1}, \dots, j_{r+s}}.$$

The reader should compare this to [3, Corollary 2.4], where such independence is proved if $r + s = n$ and $\min_b j_b > \max_a i_a$. All the above conjectures fit with data for values of $n \leq 8$.

The techniques used to prove the two-point case in Theorem 4.2 fail for Conjecture 7.4, and it is clear that new ideas are needed for the conjectures above. Maybe the key to all these conjectures is to find a conceptual proof of $E_{i,j} = 1/n^2$ if $j > i + 1$ in Theorem 4.2.

Problem 8.6. Find a more conceptual proof of the independence $E_{i,j} = 1/n^2$ if $j > i + 1$.

On the combinatorial side it would be interesting to find a simpler proof of Theorem 6.4.

Problem 8.7. Find a proof of Theorem 6.4 that explains the occurrence of the Narayana numbers.

ACKNOWLEDGEMENTS

The authors thank MSRI for their hospitality, where this research was initiated. We thank Thomas Lam, James Martin, Greta Panova, Anne Schilling and Lauren Williams for fruitful discussions. The first author would like to acknowledge support in part by a UGC Centre for Advanced Study grant and in part by the DST grant DST/INT/SWD/VR/P-01/2014. The second author has had partial support from the Swedish Science Council 621-2014-4780. The authors also thank the anonymous referees for useful comments on the draft.

REFERENCES

- [1] Erik Aas, *A Markov Process on Cyclic Words*, PhD-thesis at KTH, Sweden (Dec. 2014), [urn:nbn:se:kth:diva-156862](https://nbn-resolving.org/urn:nbn:se:kth:diva-156862)
- [2] Erik Aas and Svante Linusson, *Continuous Multi-line Queues and the TASEP*, [arXiv:1501.04417](https://arxiv.org/abs/1501.04417).

- [3] Erik Aas and Jonas Sjöstrand, *A product formula for the TASEP on a ring*, Random Struct. Alg. **48** (2015), no. 2, 247–259, arXiv.org/abs/1312.2493.
- [4] Gideon Amir, Omer Angel, and Benedek Valkó, *The TASEP speed process*, Ann. Probab. **39** (2011), no. 4, 1205–1242, DOI 10.1214/10-AOP561. MR2857238
- [5] Arvind Ayyer and Svante Linusson, *An inhomogeneous multispecies TASEP on a ring*, Adv. in Appl. Math. **57** (2014), 21–43, DOI 10.1016/j.aam.2014.02.001. MR3206520
- [6] Anders Björner and Francesco Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics, vol. 231, Springer, New York, 2005. MR2133266 (2006d:05001)
- [7] Pablo A. Ferrari, Patricia Gonçalves, and James B. Martin, *Collision probabilities in the rarefaction fan of asymmetric exclusion processes* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), no. 4, 1048–1064, DOI 10.1214/08-AIHP303. MR2572163 (2010m:60337)
- [8] Pablo A. Ferrari and J. B. Martin, *Multi-class processes, dual points and M/M/1 queues*, Markov Process. Related Fields **12** (2006), no. 2, 175–201. MR2249628 (2008m:60200)
- [9] Pablo A. Ferrari and James B. Martin, *Stationary distributions of multi-type totally asymmetric exclusion processes*, Ann. Probab. **35** (2007), no. 3, 807–832, DOI 10.1214/009117906000000944. MR2319708 (2008h:60403)
- [10] James E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics, vol. 29, Cambridge University Press, Cambridge, 1990. MR1066460 (92h:20002)
- [11] Kurt Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), no. 2, 437–476, DOI 10.1007/s002200050027. MR1737991 (2001h:60177)
- [12] Thomas Lam, *The shape of a random affine Weyl group element and random core partitions*, Ann. Probab. **43** (2015), no. 4, 1643–1662, DOI 10.1214/14-AOP915. MR3353811
- [13] Thomas Lam, Luc Lapointe, Jennifer Morse, and Mark Shimozono, *Affine insertion and Pieri rules for the affine Grassmannian*, Mem. Amer. Math. Soc. **208** (2010), no. 977, xii+82, DOI 10.1090/S0065-9266-10-00576-4. MR2741963 (2012f:05314)
- [14] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Markov chains and mixing times*, With a chapter by James G. Propp and David B. Wilson. American Mathematical Society, Providence, RI, 2009. MR2466937 (2010c:60209)
- [15] Svante Linusson and James Martin, *Stationary probabilities for an inhomogeneous multi-type TASEP*, in preparation.
- [16] Thomas Mountford and Hervé Guioil, *The motion of a second class particle for the TASEP starting from a decreasing shock profile*, Ann. Appl. Probab. **15** (2005), no. 2, 1227–1259, DOI 10.1214/105051605000000151. MR2134103 (2006d:60152)
- [17] H. Rost, *Nonequilibrium behaviour of a many particle process: density profile and local equilibria*, Z. Wahrsch. Verw. Gebiete **58** (1981), no. 1, 41–53, DOI 10.1007/BF00536194. MR635270 (83a:60176)
- [18] Richard P. Stanley, *Enumerative combinatorics. Vol. 2*. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999. MR1676282 (2000k:05026)

DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF SCIENCE, BANGALORE - 560012, INDIA
E-mail address: arvind@math.iisc.ernet.in

DEPARTMENT OF MATHEMATICS, KTH-ROYAL INSTITUTE OF TECHNOLOGY, SE-100 44, STOCKHOLM, SWEDEN
E-mail address: linusson@math.kth.se