Asymptotics of a Locally Dependent Statistic on Finite Reflection Groups

Frank Röttger

Max Planck Institute for Mathematics in the Sciences
Leipzig, Germany

`rottger@mis.mpg.de`

Abstract

This paper discusses the asymptotic behaviour of the number of descents in a random signed permutation and its inverse, which was listed as an interesting direction by Chatterjee and Diaconis (2017). For that purpose, we generalize their result for the asymptotic normality of the number of descents in a random permutation and its inverse to other finite reflection groups. This is achieved by applying their proof scheme to signed permutations, i.e. elements of Coxeter groups of type B_n, which are also known as the hyperoctahedral groups. Furthermore, a similar central limit theorem for elements of Coxeter groups of type D_n is derived via Slutsky's Theorem and a bound on the Wasserstein distance of certain normalized statistics with local dependency structures and bounded local components is proven for both types of Coxeter groups. In addition, we show a two-dimensional central limit theorem via the Cramér-Wold device.

Mathematics Subject Classifications: 20F55, 60C05, 60F05

1 Introduction

A recent result of Chatterjee and Diaconis in [5] was a new proof of the asymptotic normality of the number of descents in a random permutation and its inverse, normalized by its expected value and its variance. This was shown via the *method of interaction graphs*, which is an approach first introduced by Chatterjee [4] in 2008 that was derived from Stein's method. Chatterjee's result gives a bound on the Wasserstein distance to the standard normal distribution. As convergence with respect to the Wasserstein distance implies convergence in distribution, such bounds are powerful tools to investigate asymptotic normality. The Wasserstein distance, also known as Kantorovich-Rubinstein metric, is a common metric on probability distributions, although it is maybe less known

as for example the Kolmogorov distance. An intuitive definition of the Wasserstein distance comes from optimal transport: The Wasserstein distance between two probability measures corresponds to the minimal cost to transport the stochastic mass of one measure into the other. The method of interaction graphs and the bound on the Wasserstein distance as introduced by Chatterjee [4] are shortly summarized in Section 2.

The asymptotic normality of the number of descents in a random permutation and its inverse was already shown by Vatutin via generating functions in [8] in 1996, but was not generalized to other statistics depending on both a random permutation and its inverse. Chatterjee and Diaconis [5] showed such a generalization through a bound of the Wasserstein distance to the standard normal distribution for a wider class of normalized statistics that depend on a random permutation and its inverse. In the last section of [5], they issued the asymptotic normality of the number of descents in an element of a finite reflection group and its inverse, for example for random signed permutations, as an interesting direction and indicated, that their approach should also suffice in this case. This paper confirms their intuition by applying their proof scheme on signed permutations, that is elements of the Coxeter group of type B_n. For that purpose we construct random signed permutations and their inverses from the same random variables. With this construction, we are able to apply the method of interaction graphs, exactly like Chatterjee and Diaconis. Together with the bound on the Wasserstein distance between the normalized statistic and a standard normal distribution mentioned before (see Theorem 3), we can show the asymptotic normality by plugging in the formulas for the variance of the sum of the statistics into the bounds. Kahle and Stump listed the expected values and variances of the sum of the statistics for all finite irreducible Coxeter groups in [6, Corollary 5.2].

Using this result for signed permutations, we can extend the result to elements of the Coxeter group of type D_n, which are signed permutations with an even number of negative signs. This is done via an application of Slutsky's Theorem (see Theorem 5).

To generalize these results to certain sums of statistics of both a random signed permutation and its inverse, which have a bounded local degree and local components which are bounded by 1, we again follow Chatterjee and Diaconis and modify the interaction graphs in the right way so that we can apply Theorem 3. From this, we show that this also works for elements of Coxeter groups of type D_n (see Section 5). The last section discusses the asymptotic behaviour of the two-dimensional statistic formed by the number of descents in an element of a Coxeter group of both type B_n and D_n and its inverse via the Cramér-Wold device.

2 Interaction Graphs

We give a short overview over the method of interaction graphs as it is presented in [5]. Let $(\mathcal{X}, \mathcal{A})$ be a measurable space and $f : \mathcal{X}^n \to \mathbb{R}$ a measurable map. Consider a map $G(x)$, which connects every $x \in \mathcal{X}^n$ with a simple graph on $[n] := \{1, 2, \ldots, n\}$. This *graphical rule* is *symmetric*, if for a permutation π the graph $G(x_{\pi(1)}, \ldots, x_{\pi(n)})$ has exactly the permuted edge set $\{(\pi(i), \pi(j)) |\ (i, j)$ is an edge of $G(x_1, \ldots, x_n)\}$. For $m \geqslant n$, let $G'(x)$

for $x \in \mathcal{X}^m$ be a symmetric graphical rule on $\mathcal{X}^m$. $G'(x)$ is an *extension* of $G(x)$, if $G(x) = G(x_1, \ldots, x_n)$ is a subgraph of $G'(x) = G'(x_1, \ldots, x_m)$ for all $x \in \mathcal{X}^m$. To define an interaction rule, let for $x, x' \in \mathcal{X}^n$

$$x^i := (x_1, \ldots, x_{i-1}, x'_i, x_{i+1}, \ldots, x_n).$$

Furthermore, let x^{ij} be the vector x with replacements in the i-th and j-th position. Then, i and j are *non-interacting*, if

$$f(x) - f(x^j) = f(x^i) - f(x^{ij}).$$

A graphical rule G is an *interaction rule* for a function f, if for any $x, x' \in \mathcal{X}^n$ and any i, j, the edge (i, j) not being an edge of either $G(x), G(x^i), G(x^j)$ or $G(x^{ij})$ implies that i and j are non-interacting.

The Wasserstein distance is a distance function on the space of probability measures [1, Chapter 7].

Definition 1 (Wasserstein distance, also known as Kantorovich–Rubinstein metric)**.** Let (M, d) be a metric space where every probability measure is a Radon measure and let $P_p(M)$ be the collection of probability measures on M with finite p-th moments. The L^p-Wasserstein distance between $X \sim \mu \in P_p(M)$ and $Y \sim \nu \in P_p(M)$ is defined as

$$\delta_p(\mu, \nu) = (\inf \mathbb{E}\,[d(X, Y)^p])^{\frac{1}{p}},$$

where the infimum is taken over all joint distributions of $(X, Y)^T$ on $M \times M$ with marginals μ and ν.

Definition 2. Let $Y = (Y_1, \ldots, Y_n)$ be a vector of real-valued random variables distributed according to a continuous distribution. The rank statistic is defined as $R(Y_i) = \sum_{j=1}^n \mathbb{1}_{\{Y_i \geqslant Y_j\}}$, where $\mathbb{1}_{\{.\}}$ denotes the indicator function. The value of $R(Y_i)$ gives the position of Y_i when Y is sorted in ascending order.

We later apply the following theorem from [4], which can also be found in [5], on signed permutations. The Theorem gives a bound on the Wasserstein distance between a normalized statistic that admits a graphical interaction rule and the standard normal distribution.

Theorem 3 (Chatterjee)**.** *Let $f : \mathcal{X}^n \to \mathbb{R}$ be a measurable map that admits a symmetric interaction rule $G(x)$. Let $X_1, X_2, \ldots$ be independent and identically distributed $\mathcal{X}$-valued random variables and let $X := (X_1, \ldots, X_n)$. Let $W := f(X)$ and $\sigma^2 := \mathbb{V}(W)$. Let $X' = (X'_1, \ldots, X'_n)$ be an independent copy of X. For each j, define*

$$\Delta_j f(X) = W - f(X_1, \ldots, X_{j-1}, X'_j, X_{j+1}, \ldots, X_n)$$

and let $M := \max_j |\Delta_j f(X)|$. Let $G'(x)$ be an extension of $G(x)$ on $\mathcal{X}^{n+4}$ and define

$$\delta := 1 + \text{ degree of the vertex } 1 \text{ in } G'(X_1, \ldots, X_{n+4}).$$

Then, the Wasserstein distance δ_W between $\frac{W - \mathbb{E}(W)}{\sigma}$ and $N(0,1)$ satisfies

$$\delta_W \leqslant \frac{C\sqrt{n}}{\sigma^2}\mathbb{E}(M^8)^{\frac{1}{4}}\mathbb{E}(\delta^4)^{\frac{1}{4}} + \frac{1}{2\sigma^3}\sum_{j=1}^{n}\mathbb{E}|\Delta_j f(X)|^3$$

for some constant C independent of n.

Chatterjee and Diaconis used the theorem above to show a central limit theorem for statistics of the form $F_1(\pi) + F_2(\pi^{-1})$, where both F_1 and F_2 have bounded local degree and their local components' absolute values are bounded by 1. Hereby π denoted a permutation, hence an element of a Coxeter group of type $\mathtt{A}_n$. We apply the same proof scheme to statistics on signed permutation by modifying their model.

3   Signed Permutations

Chatterjee and Diaconis modeled elements of the symmetric group $S_n = \mathtt{A}_{n-1}$ and their inverses by ranking functions on series of uniformly distributed random variables on the unit square. We slightly modify this model by additionally introducing a random sign. The Coxeter group of type $\mathtt{B}_n$ is the symmetry group of the n-hypercube. It is isomorphic to the signed permutation group of rank n, which is the subgroup of all permutations on $\{\pm 1, \ldots, \pm n\}$ with the antisymmetric constraint $-\tilde{\pi}(i) = \tilde{\pi}(-i)$. In a one-line notation we write $\tilde{\pi} = (\tilde{\pi}(1), \ldots, \tilde{\pi}(n))$ where $\tilde{\pi}(i) \in \{\pm 1, \ldots, \pm n\}$ and $\{|\tilde{\pi}(1)|, \ldots, |\tilde{\pi}(n)|\} = [n]$ (see [6, Section 2.1]). Compare also the book of Björner–Brenti [3, p. 245], where an element $\tilde{\pi} \in \mathtt{B}_n$ is written in *window notation*, if $\tilde{\pi} = [a_1, \ldots, a_n]$ with $\tilde{\pi}(i) = a_i$ for $i = 1, \ldots, n$.

Following [3, Proposition 8.1.2], it holds that the descents in some signed permutation $\tilde{\pi} \in \mathtt{B}_n$ in the one-line notation are

$$\mathrm{Des}(\tilde{\pi}) = \{0 \leqslant i < n : \tilde{\pi}(i) > \tilde{\pi}(i+1)\},$$

where $\tilde{\pi}(0) = 0$. We write for $\tilde{\pi} \in \mathtt{B}_n$

$$\mathrm{des}(\tilde{\pi}) = |\mathrm{Des}(\tilde{\pi})| = \mathbb{1}_{\{0 > \tilde{\pi}(1)\}} + \sum_{i=1}^{n-1}\mathbb{1}_{\{\tilde{\pi}(i) > \tilde{\pi}(i+1)\}}. \tag{1}$$

In the following theorem, we study the asymptotic behaviour of the statistic

$$t(\tilde{\pi}) = \mathrm{des}(\tilde{\pi}) + \mathrm{des}(\tilde{\pi}^{-1}).$$

If $\tilde{\pi}$ is picked uniformly from $\mathtt{B_n}$, the statistic $t(\tilde{\pi})$ gives rise to a random variable $T_{\mathtt{B}_n}$. We show a central limit theorem for the sequence $(T_{\mathtt{B}_n})_n$, normalized by its expected value and its variance, so

$$\frac{T_{\mathtt{B}_n} - \mathbb{E}(T_{\mathtt{B}_n})}{\sqrt{\mathbb{V}(T_{\mathtt{B}_n})}} \xrightarrow{D} N(0,1), \tag{2}$$

by adapting the proof of Theorem 1.1 in [5] for the modified model.

Theorem 4. *Given a sequence of Coxeter groups of type* $\mathtt{B}_n$ *of growing rank. Then,* $T_{\mathtt{B}_n}$ *satisfies the central limit theorem, if* n *tends to infinity.*

Proof. Let $\mathcal{X} := [0,1]^2 \times \{-1,1\}$ and $X_1, X_2, \ldots$ be independent and identically distributed of the form (U_i, V_i, B_i) with $(U_i, V_i) \sim \mathrm{Unif}([0,1]^2)$ and $B_i \sim \mathrm{Ber}(\frac{1}{2})$ on $\{-1,1\}$ and independent of (U_i, V_i). Let $X := (X_1, \ldots, X_n)$ and let the x-rank of X_i be the rank statistic (cf. Definition 2) of U_i among $(U_1, \ldots, U_n)$ and the y-rank of X_i the rank statistic of V_i among $(V_1, \ldots, V_n)$, so that $X_{(1)}, \ldots, X_{(n)}$ denote the X_i ordered with respect to their x-ranks and $X^{(1)}, \ldots, X^{(n)}$ with respect to their y-ranks. This means that $\pi(i) = $ y-rank of $X_{(i)}$ is a random permutation and $\sigma(i) = $ x-rank of $X^{(i)}$ is its inverse. Now, to see that

$$\tilde{\pi}(i) := B_{(|i|)}\mathrm{sign}(i)\pi(|i|), \qquad \tilde{\sigma}(i) := B^{(|i|)}\mathrm{sign}(i)\sigma(|i|)$$

define random signed permutations, just check that $\tilde{\pi}(-i) = -\tilde{\pi}(i)$ and $\tilde{\sigma}(-i) = -\tilde{\sigma}(i)$ and that $\tilde{\pi}(i)$ and $\tilde{\sigma}(i)$ are injective. Furthermore it follows that $\tilde{\sigma} = \tilde{\pi}^{-1}$, as $B_{(\sigma(|i|))} = B^{(|i|)}$ and

$$\tilde{\pi}(\tilde{\sigma}(i)) = B_{(|\tilde{\sigma}(i)|)}\mathrm{sign}(\tilde{\sigma}(i))\pi(|\tilde{\sigma}(i)|) = B_{(\sigma(|i|))}\mathrm{sign}(B^{(|i|)}\mathrm{sign}(i))\pi(\sigma(|i|)) = i.$$

Therefore the number of descents in the signed permutation and its inverse is given by:

$$T_{\mathtt{B}_n} := f(X) = \sum_{i=0}^{n-1} \mathbb{1}_{\{\tilde{\pi}(i) > \tilde{\pi}(i+1)\}} + \sum_{i=0}^{n-1} \mathbb{1}_{\{\tilde{\sigma}(i) > \tilde{\sigma}(i+1)\}} \tag{3}$$

$$= \mathbb{1}_{\{0 > B_{(1)}\pi(1)\}} + \sum_{i=1}^{n-1} \mathbb{1}_{\{B_{(i)}\pi(i) > B_{(i+1)}\pi(i+1)\}} + \mathbb{1}_{\{0 > B^{(1)}\sigma(1)\}} + \sum_{i=1}^{n-1} \mathbb{1}_{\{B^{(i)}\sigma(i) > B^{(i+1)}\sigma(i+1)\}}$$

To apply Theorem 3, it is required to define a symmetric interaction rule:

For $x \in \mathcal{X}^n$, define a simple graph $G(x)$ on $[n]$ as follows: For any $1 \leqslant i \neq j \leqslant n$, let $\{i, j\}$ be an edge if and only if the x-rank of x_i and the x-rank of x_j or the y-rank of x_i and the y-rank of x_j differ by at most 1. To check that this graphical rule is symmetric, see that the edge set of a relabeled Graph $G(x_{\pi(1)}, \ldots, x_{\pi(n)})$, where π is an arbitrary permutation, has the edge set $\{(\pi(i), \pi(j)) \mid (i, j) \text{ is an edge of } G(x_1, \ldots, x_n)\}$. This is true, since the x-ranks or the y-ranks of $x_{\pi(i)}$ are equal to the respective ranks of x_i. Hence this graph is invariant under relabeling of the indices and it is therefore a symmetric graphical rule.

We now show that $G(x)$ is an interaction rule. For this, we define x^i as the vector $(x_1, \ldots x_{i-1}, x_i', x_{i+1}, \ldots, x_n)$ where $x, x' \in \mathcal{X}^n$. This means that x^i is the vector x in which the i-th entry is replaced by the i-th entry of x'. Furthermore, x^{ij} is the vector with replacements in the i-th and the j-th entry. Now, suppose that (i, j) is not an edge in $G(x), G(x^i), G(x^j)$ or $G(x^{ij})$. Then, the equation

$$f(x) - f(x^j) = f(x^i) - f(x^{ij})$$

holds, as j is not a neighbour of i in either of the four graphs. To better visualize this, check that

$$f(x) = f(x^i) + f(x^j) - f(x^{ij}). \tag{4}$$

Any indicator function in $f(x)$, that is not dependent of either x_i or x_j, appears in $f(x^i), f(x^j)$ and $f(x^{ij})$, as it is left unchanged by the replacements in x^i, x^j or x^{ij}. Those indicator functions, that depend on x_i but not on x_j, are unchanged in $f(x^j)$. As i and j are no neighbours in all four graphs, these indicator functions, that depend on x^i but not on x^j, appear in both $f(x^i)$ and $f(x^{ij})$. Therefore, the indicator functions that either depend on x_i or on x_j turn up exactly once on both sides of the equation. Hence Equation (4) holds, since there cannot be an indicator functions that depend on both x^i and x^j, as i and j are no neighbours in all four graphs. This means, that $G(x)$ is a symmetric interaction rule for f.

Now, we construct an extension $G'(x)$ of $G(x)$ on $\mathcal{X}^{n+4}$. For any $1 \leqslant i \neq j \leqslant n+4$, let $\{i, j\}$ be an edge in $G'(x)$ if and only if the x-rank of x_i and the x-rank of x_j or the y-rank of x_i and the y-rank of x_j, differ by at most 5. As this graph is invariant under relabeling of the indices, it is a symmetric graphical rule. Obviously, every edge in $G(x)$ is also an edge in $G'(x)$, as the distance between two connected nodes in $G(x)$ can be 5 at most through the insertion of four additional nodes. Therefore $G'(x)$ is an extension of $G(x)$.

It remains to check the conditions in Theorem 3: As $T_{\mathtt{B}_n}$ can differ in at most 4 summands from $f(X_1, \ldots, X_{j-1}, X'_j, X_{j+1}, \ldots, X_n)$, it follows that $|\Delta_j f(X)| \leqslant 4$. Furthermore, the degree of any node in $G'(x)$ is bounded by 20, as either the difference in the x-ranks or in the y-ranks has to be smaller or equal to 5. This means, that $|\delta| \leqslant 21$. Then, by Theorem 3,

$$\delta_{T_{\mathtt{B}_n}} \leqslant \frac{C\sqrt{n}}{\sigma^2} + \frac{Cn}{\sigma^3}$$

for some constant C. As [6] shows, $\sigma^2 = \mathbb{V}(T_{\mathtt{B}_n}) = \frac{n+3}{6}$. Therefore, $T_{\mathtt{B}_n}$ follows the central limit theorem. $\square$

4 Coxeter Group of Type $\mathtt{D}_n$

This section reproduces the previous section's result for sequences of Coxeter groups of type $\mathtt{D}_n$. The Coxeter group of type $\mathtt{D}_n$ is the symmetry group of the n-demicube. It is isomorphic to the subgroup of the signed permutation group of rank n that consist of all signed permutation with an even number of negative signs. This means, that

$$\mathtt{D}_n = \{\pi \in \mathtt{B}_n : \prod_{i=1}^n \pi(i) > 0\}.$$

For some $\pi \in \mathtt{D}_n$, it holds that

$$\mathrm{Des}(\pi) = \{0 \leqslant i < n : \pi(i) > \pi(i+1)\},$$

where $\pi(0) = -\pi(2)$ [3, Proposition 8.2.2]. We write for $\pi \in \mathtt{D}_n$

$$\text{des}(\pi) = |\text{Des}(\pi)| = \mathbb{1}_{\{-\pi(2)>\pi(1)\}} + \sum_{i=1}^{n-1} \mathbb{1}_{\{\pi(i)>\pi(i+1)\}}. \tag{5}$$

We can reuse the model from the proof of Theorem 4 to generate $T_{\mathtt{D}_n}$, with a slight modification: One sign-generating random variable is set to be the product of all the others. Therefore, the number of negative signs is always even. Of course it is not possible to directly apply the method of interaction graphs, as the local dependency structure is destroyed by one random variable being dependent of all the others. This problem is solved via an application of Slutsky's Theorem.

Theorem 5. *Let W_n be a sequence of growing rank of Coxeter groups of type $\mathtt{D}_n$. Then, $T_{\mathtt{D}_n}$ satisfies the central limit theorem, if n tends to infinity.*

Proof. Let $\mathcal{X} := [0,1]^2 \times \{-1,1\}$ and $X_1, X_2, \ldots, X_{n-1}$ be independent and identically distributed of the form (U_i, V_i, B_i) with $(U_i, V_i) \sim \text{Unif}([0,1]^2)$ and $B_i \sim \text{Ber}(\frac{1}{2})$ on $\{-1,1\}$. Furthermore, set $X_n = (U_n, V_n, \prod_{i=1}^{n-1} B_i)$ with $(U_n, V_n) \sim \text{Unif}([0,1]^2)$ and $B_n = \prod_{i=1}^{n-1} B_i$. The product of independent $\text{Ber}(\frac{1}{2})$-distributed random variables on $\{-1,1\}$ is again $\text{Ber}(\frac{1}{2})$-distributed on $\{-1,1\}$. Let $X := (X_1, \ldots, X_n)$ and let the x-rank and the y-rank of X be defined as in the proof of Theorem 4. $X_{(1)}, \ldots, X_{(n)}$ denote the X_i ordered in respect to their x-ranks and $X^{(1)}, \ldots, X^{(n)}$ in respect to their y-ranks. Then, as in (5), if $\tilde{\pi} \in \mathtt{D}_n$ and $\tilde{\pi}^{-1} = \tilde{\sigma}$, with $\tilde{\pi}(0) = -\tilde{\pi}(2)$ we obtain

$$T_{\mathtt{D}_n} = \sum_{i=0}^{n-1} \mathbb{1}_{\{\tilde{\pi}(i)>\tilde{\pi}(i+1)\}} + \sum_{i=0}^{n-1} \mathbb{1}_{\{\tilde{\sigma}(i)>\tilde{\sigma}(i+1)\}}$$

$$= \mathbb{1}_{\{-B_{(2)}V_{(2)}>B_{(1)}V_{(1)}\}} + \sum_{i=1}^{n-1} \mathbb{1}_{\{B_{(i)}V_{(i)}>B_{(i+1)}V_{(i+1)}\}}$$

$$+ \mathbb{1}_{\{-B^{(2)}U^{(2)}>B^{(1)}U^{(1)}\}} + \sum_{i=1}^{n-1} \mathbb{1}_{\{B^{(i)}U^{(i)}>B^{(i+1)}U^{(i+1)}\}}.$$

Now, $T_{\mathtt{D}_n}$ is equal in distribution to $T_{\mathtt{B}_{n-1}} + Y_n$, where $Y_n = T_{\mathtt{D}_n} - T_{\mathtt{B}_{n-1}}$ is a random variable with $|Y_n| \leqslant c$ for some positive constant c and all n. This holds as we directly obtain $T_{\mathtt{B}_{n-1}}$ from $X_1, \ldots, X_{n-1}$. Comparing $T_{\mathtt{B}_{n-1}}$ and $T_{\mathtt{D}_n}$, we see that all indicator functions are identical except where $B_{(i)}, B^{(i)}, B_{(i+1)}$ or $B^{(i+1)}$ equal B_n. This implies $|Y_n| \leqslant c$.

Then, as $\mathbb{E}(T_{\mathtt{D}_n}) = n$ and $\mathbb{E}(T_{\mathtt{B}_{n-1}}) = n-1$ (see for example in [6]),

$$\frac{T_{\mathtt{D}_n} - \mathbb{E}(T_{\mathtt{D}_n})}{\sqrt{\mathbb{V}(T_{\mathtt{D}_n})}} = \frac{T_{\mathtt{B}_{n-1}} + Y_n - n}{\sqrt{\mathbb{V}(T_{\mathtt{D}_n})}} = \frac{\sqrt{\mathbb{V}(T_{\mathtt{B}_{n-1}})}}{\sqrt{\mathbb{V}(T_{\mathtt{D}_n})}} \frac{T_{\mathtt{B}_{n-1}} - (n-1)}{\sqrt{\mathbb{V}(T_{\mathtt{B}_{n-1}})}} + \frac{Y_n - 1}{\sqrt{\mathbb{V}(T_{\mathtt{D}_n})}}. \tag{6}$$

We know from Theorem 4 that $\frac{T_{\mathtt{B}_{n-1}} - (n-1)}{\sqrt{\mathbb{V}(T_{\mathtt{B}_{n-1}})}}$ converges in distribution to a standard normal distribution. Y_n is bounded, as it is a finite sum of indicator functions. Therefore,

$\lim\limits_{n\to\infty} \frac{Y_n - 1}{\sqrt{\mathbb{V}(T_{D_n})}} = 0$ almost surely and $\lim\limits_{n\to\infty} \frac{\sqrt{\mathbb{V}(T_{B_{n-1}})}}{\sqrt{\mathbb{V}(T_{D_n})}} = 1$ (compare [6, Corollary 5.2]).
Therefore, T_{D_n} satisfies the central limit theorem (see Slutsky's theorem, for example in [7, Theorem 2.3.3]). $\qquad\square$

5 Generalization to a Class of Statistics with Local Degree k

As in [5], it is possible to generalize the proof of Theorem 4 via the method of interaction graphs as introduced by Chatterjee in [4, Section 2.3] to a wider class of statistics of local degree k. These statistics are of the form

$$F_1(\pi) + F_2(\pi^{-1}),$$

where the local components' absolute value is bounded by 1. If π is a signed permutation, a bound for the Wasserstein distance between the normalized statistic and the standard normal distribution follows. Therefore the central limit theorem for these statistics holds, if the variance of the statistics is of order $n^{\frac{1}{2}+\varepsilon}$ for an $\varepsilon > 0$. The Theorem is implied from a generalization of the proof of Theorem 4 by constructing the symmetric interaction rule in the right way.

Theorem 6. *Let W_n be a sequence of growing rank of Coxeter groups of type B and let F_1, F_2 be statistics of local degree k, with the absolute value of their local components bounded by 1. The statistic $F_1(\pi) + F_2(\pi^{-1})$ gives rise to a random variable F. The Wasserstein distance between F, normalized by its mean and variance, and the standard normal distribution satisfies*

$$\delta_F \leqslant C(k) \left(\frac{\sqrt{n}}{s^2} + \frac{n}{s^3} \right)$$

for $s^2 := \mathbb{V}(F_1(\pi) + F_2(\pi^{-1}))$ and some constant $C(k)$.

Proof. If the statistics F_1 and F_2 are of local degree k and their local components' absolute value is bounded by 1, let $\{i, j\}$ be an edge in $G(x)$ if and only if the x-ranks or the y-ranks differ by at most $k-1$. For the extension $G'(x)$, we say that $\{i, j\}$ is an edge if and only if the ranks differ by at most $k+3$. Then, Theorem 3 applies, and the Wasserstein distance is bounded:

$$\delta_F \leqslant C(k) \left(\frac{\sqrt{n}}{s^2} + \frac{n}{s^3} \right).$$

Here, $C(k)$ is a large enough constant. $\qquad\square$

To see that the bound in Theorem 6 also holds when π is an element of a Coxeter group of type D_n, we use the same technique as in the proof of Theorem 5. Hence, we decompose the statistic into a part that is the same statistic depending on a signed permutation on $\{\pm 1, \ldots, \pm(n-1)\}$ and a finitely bounded random variable.

Theorem 7. *Let W_n be a sequence of growing rank of Coxeter groups of type* D *and let F_1, F_2 be statistics of local degree k, with the absolute value of their local components bounded by 1. The statistic $F_1(\pi) + F_2(\pi^{-1})$ gives rise to a random variable F. Then, if we assume that $\mathbb{V}(F) \to \infty$, the Wasserstein distance between F, normalized by its mean and variance, and the standard normal distribution satisfies*

$$\delta_F \leqslant C(k)\left(\frac{\sqrt{n-1}}{s^2} + \frac{n-1}{s^3}\right) + o(1)$$

for $s^2 := \mathbb{V}(F_1(\pi) + F_2(\pi^{-1}))$ and some constant $C(k)$.

Proof. Let $F = F_1(\pi_1) + F_2(\pi_1^{-1}) = f(X)$ where π_1 is a uniformly chosen element of the Coxeter group of type D_n. Let $X = (X_1, \ldots, X_n)$ be generated as in the proof of Theorem 5, so $X_i = (U_i, V_i, B_i)$ with $(U_i, V_i) \sim \text{Unif}([0, 1]^2)$. B_i is an independent random sign for $1 \leqslant i \leqslant n - 1$ and $B_n = \prod_{i=1}^{n-1} B_i$. Then, F' is the statistic where we remove all local components that depend on B_n. Subsequently we add local components, so that the resulting statistic is $F' = F_1(\pi_2) + F_2(\pi_2^{-1})$, where π_2 is a random signed permutation on $\{\pm 1, \ldots \pm (n-1)\}$ generated by $(X_1, \ldots, X_{n-1})$. Then, as the local degree is k, $F - F' = O(1)$ and therefore $\mathbb{E}(F - F') = O(1)$ and $\mathbb{V}(F - F') = O(1)$, which implies that $\mathbb{V}(F') = \mathbb{V}(F) + O(1)$. Now, see that Eq. (6) from the proof of Theorem 5 generalizes to

$$\frac{F - \mathbb{E}(F)}{\sqrt{\mathbb{V}(F)}} = \frac{\sqrt{\mathbb{V}(F')}}{\sqrt{\mathbb{V}(F)}} \frac{F' - \mathbb{E}(F')}{\sqrt{\mathbb{V}(F')}} + \frac{F - F' - \mathbb{E}(F - F')}{\sqrt{\mathbb{V}(F)}},$$

which immediately shows that the Wasserstein distance between F and F' tends to zero, because $\lim_{n \to \infty} \frac{\mathbb{V}(F')}{\mathbb{V}(F)} = 1$ and $\lim_{n \to \infty} \frac{F - F' - \mathbb{E}(F - F')}{\sqrt{\mathbb{V}(F)}} = 0$. Therefore it holds that $\delta_F \leqslant \delta_{F'} + o(1)$ and the theorem follows. $\square$

6 The Statistic $(\text{des}(\pi), \text{des}(\pi^{-1}))$

This section derives a two-dimensional central limit theorem for the vector statistic defined as $(\text{des}(\pi), \text{des}(\pi^{-1}))$ for π being either an element of a Coxeter group of type B_n or D_n. This is achieved with the Cramér–Wold device and a slight modification of the proofs of Theorems 4 and 5. The Cramér–Wold device shows the equivalence of the convergence in distribution between a random vector and every linear combination of its elements. It is also known as the Theorem of Cramér–Wold (see for example in [2, Theorem 29.4]).

Theorem 8 (Cramér–Wold). *Let $\bar{X}_n = (X_{n1}, \ldots, X_{nk})$ and $\bar{X} = (X_1, \ldots, X_k)$ be random vectors of dimension k. Then, $\bar{X}_n \xrightarrow{D} \bar{X}$, if and only if*

$$\sum_{i=1}^{k} t_i X_{ni} \xrightarrow{D} \sum_{i=1}^{k} t_i X_i$$

for each $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$ and for $n \to \infty$.

We use the short-hand notation (D_n, D_n') for the random variable that rises from $(\mathrm{des}(\pi), \mathrm{des}(\pi^{-1}))$. With Theorem 8, we can show the convergence of (D_n, D_n') by studying linear combinations of the form $t_1 D_n + t_2 D_n'$. It is sufficient to only check linear combinations with $t \in S^1$, since the investigated statistic is normalized by the square root of the variance $\mathbb{V}(t_1 D_n + t_2 D_n')$. This leads to the following theorem:

Theorem 9. *Let W_n be a sequence of Coxeter groups of growing rank of either type B_n or D_n. Then, the statistic (D_n, D_n') satisfies a two-dimensional central limit theorem of the form*

$$\Sigma_n^{-\frac{1}{2}} \begin{pmatrix} D_n - \mathbb{E}(D_n) \\ D_n' - \mathbb{E}(D_n') \end{pmatrix} \xrightarrow{D} N_2(0, I)$$

for $n \to \infty$, where I denotes the two-dimensional identity matrix and Σ_n is the covariance matrix of (D_n, D_n').

Proof. Via the Theorem of Cramér–Wold, we can study the convergence of (D_n, D_n') by studying $t_1 D_n + t_2 D_n'$ for $t^T = (t_1, t_2) \in S^1$. We derive a convergence

$$t^T \frac{1}{\sqrt{\mathbb{V}(D_n)}} \begin{pmatrix} D_n - \mathbb{E}(D_n) \\ D_n' - \mathbb{E}(D_n') \end{pmatrix} \xrightarrow{D} N(0, 1) \tag{7}$$

to show the Theorem via an application of Slutsky's Theorem. (7) is equivalent to

$$\frac{1}{\sqrt{\mathbb{V}(D_n)}} (t_1 D_n + t_2 D_n' - (t_1 + t_2)\mathbb{E}(D_n)) \xrightarrow{D} N(0, 1), \tag{8}$$

as $\mathbb{E}(D_n) = \mathbb{E}(D_n')$. Now, since $t \in S^1$, the proofs of the Theorems 4 and 5 apply, which means that

$$\frac{t_1 D_n + t_2 D_n' - (t_1 + t_2)\mathbb{E}(D_n)}{\sqrt{\mathbb{V}(t_1 D_n + t_2 D_n')}} \xrightarrow{D} N(0, 1).$$

This convergence is also a consequence of Theorem 6 or Theorem 7, as the local components of $t_1 D_n + t_2 D_n'$ are still bound by 1 and the local dependency structure is not changed by multiplying the sum of indicator functions that model D_n and D_n' with constants. Furthermore, the variance $\mathbb{V}(t_1 D_n + t_2 D_n')$ is of order n and therefore, the Wasserstein distance to the standard normal distribution is bound by a vanishing function in n. Now, by Slutsky's Theorem, (8) and therefore (7) is satisfied as

$$\frac{\mathbb{V}(t_1 D_n + t_2 D_n')}{\mathbb{V}(D_n)} \xrightarrow{a.s} 1.$$

This results from the fact that $\mathbb{V}(D_n) = \mathbb{V}(D_n'))$ and $\mathrm{Cov}(D_n, D_n') = O(1)$ (see [6]) and that $t_1^2 + t_2^2 = 1$. Because of the convergence in (7), the theorem follows via another application of Slutsky's Theorem, as

$$\frac{1}{\mathbb{V}(D_n)} \Sigma_n = \frac{1}{\mathbb{V}(D_n)} \begin{pmatrix} \mathbb{V}(D_n) & \mathrm{Cov}(D_n, D_n') \\ \mathrm{Cov}(D_n, D_n') & \mathbb{V}(D_n') \end{pmatrix}$$
$$\xrightarrow{a.s.} I,$$

since $\mathrm{Cov}(D_n, D_n') = O(1)$ and $\mathbb{V}(D_n) = \mathbb{V}(D_n')$. $\qquad \square$

Remark 10. Theorem 9 can be generalized to certain statistics $(F_1(\pi), F_2(\pi^{-1}))$, if F_1 and F_2 meet the constraints of Theorem 6 or Theorem 7, $\mathbb{V}(F_1(\pi)) = \mathbb{V}(F_2(\pi^{-1}))$ holds and $\mathbb{V}(F_1(\pi))$ is big enough so that the constraint to the Wasserstein distance in Theorem 6 or Theorem 7 converges to zero for n going to infinity.

Acknowledgements

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[2] P. Billingsley. *Probability and Measure.* Wiley Series in Probability and Statistics. Wiley, 1995.

[3] A. Björner and F. Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics.* Springer, New York, 2005.

[4] S. Chatterjee. A new method of normal approximation. *Ann. Probab.*, 36(4):1584–1610, 2008.

[5] S. Chatterjee and P. Diaconis. A central limit theorem for a new statistic on permutations. *Indian J. Pure Appl. Math.*, 48(4):561–573, 2017.

[6] T. Kahle and C. Stump. Counting inversions and descents of random elements in finite Coxeter groups. *Math. Comp.*, 89(321):437–464, 2020.

[7] E. L. Lehmann. *Elements of Large-Sample Theory.* Springer Texts in Statistics. Springer, New York, 1998.

[8] V. A. Vatutin. The numbers of ascending segments in a random permutation and in one inverse to it are asymptotically independent. *Diskret. Mat.*, 8(1):41–51, 1996.