Electron. Commun. Probab. **25** (2020), no. 28, 1–6.
https://doi.org/10.1214/20-ECP309
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On the central limit theorem for the two-sided descent statistics in Coxeter groups

Valentin Féray*

Abstract

In 2018, Kahle and Stump raised the following problem: identify sequences of finite Coxeter groups W_n for which the two-sided descent statistics on a uniform random element of W_n is asymptotically normal. Recently, Brück and Röttger provided an almost-complete answer, assuming some regularity condition on the sequence W_n. In this note, we provide a shorter proof of their result, which does not require any regularity condition. The main new proof ingredient is the use of the second Wasserstein distance on probability distributions, based on the work of Mallows (Ann. Math. Statist., 1972).

Keywords: asymptotic normality; Wasserstein distance; Coxeter groups.
AMS MSC 2010: 60C05; 60F05; 05E15.
Submitted to ECP on December 9, 2019, final version accepted on March 13, 2020.
Supersedes arXiv:1911.10939.

We recall that a sequence of random variables $(X_n)_{n\geq 0}$ is said to be *asymptotically normal* if $\frac{X_n - \mathbb{E}[X_n]}{\sqrt{\mathrm{Var}(X_n)}}$ converges in distribution to a standard random variable $Z \sim \mathcal{N}(0,1)$. Asymptotic normality of permutation statistics is a vast topic in discrete probability, dating back at least to Goncharov [Gon44] and Hoeffding [Hoe51]; we refer also to [Vat96, Ful04, CD17, Özd19] for more recent works on the descent and two-sided descent statistics. Recently, there has been some interest into generalizing such asymptotic normality results to statistics of Coxeter group elements[1]. In particular, Kahle and Stump [KS20] have given sufficient and necessary conditions on a sequence W_n of finite Coxeter groups so that the number of inversions (resp. of descents) of a uniform random element in W_n is asymptotically normal. They then asked for a similar characterization for the two sided descent statistics t defined as follows: for an element w of a Coxeter group W, we set $t(w) = \mathrm{des}(w) + \mathrm{des}(w^{-1})$, where $\mathrm{des}(w)$ is the number of descents of w. Unlike for inversions and descents, the two sided-descent statistics on a uniform random element does not decompose as a sum of independent Bernoulli variable, making the problem more difficult. For further background on the topic, we refer to [KS20] and [BR19].

The main result of this note is a complete answer to the Kahle–Stump question.

Theorem 1. Let $(W_n)_{n\geq 1}$ be a sequence of finite Coxeter groups. For each n, we let $\boldsymbol{w}_n$ be a uniform random element in W_n. Then the following assertions are equivalent:

*University of Zurich, Switzerland. E-mail: valentin.feray@math.uzh.ch

[1] For the reader's convenience, we provide an appendix with the necessary definitions regarding Coxeter groups, in particular the notion of *descent*.

a) The sequence $t(\boldsymbol{w}_n)$ is asymptotically normal;

b) $\mathrm{Var}\left[t(\boldsymbol{w}_n)\right]$ tends to $+\infty$.

This had been previously proved by Brück and Röttger in [BR19] under a regularity assumption on the sequence $(W_n)_{n\geq 1}$ (the sequence should be *well-behaved* in the terminology of [BR19]). In addition to not requiring any regularity assumption, the proof that we provide here is shorter. In particular, we do not need any fourth moment estimates.

As in [BR19], we will take as granted that asymptotic normality holds when $(W_n)_{n\geq 1}$ is one of the infinite families A_n, B_n and D_n; this was proved previously in [Vat96, CD17, Röt18, Özd19]. In addition to the fact that $t(w)$ is bounded by $2\,\mathrm{rk}(W)$, this is the only specific information we will need on the two-sided descent statistics. All other arguments are of probabilistic nature. In particular, we shall use characteristic function analysis, and Lindeberg type arguments to prove the asymptotic normality (as in [BR19]). We also introduce a new proof ingredient: the second Wasserstein metric for probabilistic distributions.

We first recall the definition of this Wasserstein metric, and some useful properties of it, and then proceed to the proof of the main theorem.

A (real-valued) random variable X is square integrable if $\mathbb{E}[X^2] < +\infty$. A probability distribution (on $\mathbb{R}$) is square integrable if a random variable with that probability distribution is.

Definition 2 (see Lemma 2 in [Mal72])**.** Let μ and ν be square integrable probability distributions on $\mathbb{R}$. Then we define

$$d_2(\mu,\nu) = \inf_{X\sim\mu, Y\sim\nu} \|X-Y\|_2,$$

where the infimum is taken over all pairs (X,Y) of random variables defined on the same probability space and with distributions μ and ν, respectively.

As usual in probability theory, we sometimes identify a random variable and its distribution: namely for random variables Z and T (not necessarily on the same probability space), we write $d_2(Z,T) = d_2(\mathbb{P}_Z, \mathbb{P}_T)$, where $\mathbb{P}_Z$ and $\mathbb{P}_T$ are the distributions of Z and T.

The introduction of the Wasserstein metric (using L^1 norm instead of L^2 norm, and for general metric space) is usually attributed to Wasserstein (sometimes also spelled Vasershtein), though it seems that it appeared in several earlier works [EOM11]. The L^2 case and its relation with asymptotic normality were studied by Mallows [Mal72]. In particular, he established the following lemmas (Lemmas 1 and 3 in [Mal72]):

Lemma 3. Let X_n and X be square integrable random variables. Then $d_2(X_n, X)$ tends to 0 if and only if $X_n \to X$ in distribution and $\mathbb{E}[X_n^2] \to \mathbb{E}[X^2]$.

Lemma 4. Let $k > 0$ be an integer and Z be standard normal random variable. If $X_1, \cdots, X_k$ are independent random variables and $(a_j)_{j\leq k}$ are real coefficients with $\sum_{j\leq k} a_j^2 = 1$, then

$$d_2\left(\sum_{j\leq k} a_j X_j, Z\right) \leq \sum_{j\leq k} a_j^2\, d_2(X_j, Z).$$

We can now prove the main result of this note.

Proof of Theorem 1. The implication a) ⇒ b) is immediate: since $t(\boldsymbol{w}_n)$ is integer valued, it cannot tend to a continuous distribution without a renormalization factor tending

to $+\infty$; see [KS20, Proposition 6.15] for details. We focus on b) $\Rightarrow$ a) and assume that $\mathrm{Var}\big[t(\boldsymbol{w}_n)\big]$ tends to $+\infty$.

For each $n \geq 1$, we can decompose the group W_n as a direct product of irreducible factors $W_n = \prod_{j \leq r_n} W_{n,j}$. For each $j \leq r_n$, we denote by $\boldsymbol{w}_{n,j}$ uniform random elements in $W_{n,j}$ and by $\boldsymbol{t}_{n,j} = t(\boldsymbol{w}_{n,j})$ the associated two-sided descent statistics. Setting $\boldsymbol{t}_n = t(\boldsymbol{w}_n)$, we have the following decomposition:

$$\boldsymbol{t}_n \;=\; \sum_{j=1}^{r_n} \boldsymbol{t}_{n,j}, \tag{1}$$

where the $\boldsymbol{t}_{n,j}$ in the right-hand side are independent; see [BR19, Lemma 2.2]. We denote $s_{n,j}^2 = \mathrm{Var}\big[t(\boldsymbol{w}_{n,j})\big]$ and $s_n^2 = \sum_{j \leq r_n} s_{n,j}^2 = \mathrm{Var}\big[t(\boldsymbol{w}_n)\big]$. Introducing the renormalized random variables

$$\widetilde{\boldsymbol{t}_n} := \frac{\boldsymbol{t}_n - \mathbb{E}[\boldsymbol{t}_n]}{s_n}, \quad \widetilde{\boldsymbol{t}_{n,j}} := \frac{\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}]}{s_{n,j}},$$

the decomposition (1) writes as

$$\widetilde{\boldsymbol{t}_n} = \sum_{j=1}^{r_n} \tfrac{s_{n,j}}{s_n} \widetilde{\boldsymbol{t}_{n,j}}.$$

Here and in the following, all tilde variables are centered with variance 1.

We recall that irreducible finite Coxeter groups are of the following types: A_p ($p \geq 1$), B_p ($p \geq 2$), D_p ($p \geq 4$), $I_2(m)$ ($m \geq 3$) or one of the exceptional types (H_3, H_4, E_6, E_7, E_8) [Cox35]. We write $\boldsymbol{a}_p$, $\boldsymbol{b}_p$ and $\boldsymbol{d}_p$ for uniform random elements in A_p, B_p and D_p respectively. As mentioned above, from previous results [Vat96, CD17, Röt18, Özd19], we know that the three sequences

$$\widetilde{\boldsymbol{a}_p} := \frac{\boldsymbol{a}_p - \mathbb{E}[\boldsymbol{a}_p]}{\sqrt{\mathrm{Var}(\boldsymbol{a}_p)}}, \quad \widetilde{\boldsymbol{b}_p} := \frac{\boldsymbol{b}_p - \mathbb{E}[\boldsymbol{b}_p]}{\sqrt{\mathrm{Var}(\boldsymbol{b}_p)}}, \quad \widetilde{\boldsymbol{d}_p} := \frac{\boldsymbol{d}_p - \mathbb{E}[\boldsymbol{d}_p]}{\sqrt{\mathrm{Var}(\boldsymbol{d}_p)}}$$

converge in distribution to a standard normal random variable Z. In addition, their second moment is equal to 1 for all p, so we also have convergence of second moments. From Lemma 3, the distributions of $\widetilde{\boldsymbol{a}_p}$, $\widetilde{\boldsymbol{b}_p}$ and $\widetilde{\boldsymbol{d}_p}$ converge to that of Z for the d_2 metric.

Fix $\varepsilon > 0$ (everything below, including the definitions of large and small components, depends on ε). We can find $p_0 = p_0(\varepsilon)$ such that for $p \geq p_0$, we have

$$d_2\big(\widetilde{\boldsymbol{a}_p}, Z\big) \leq \varepsilon,\; d_2\big(\widetilde{\boldsymbol{b}_p}, Z\big) \leq \varepsilon,\; d_2\big(\widetilde{\boldsymbol{d}_p}, Z\big) \leq \varepsilon. \tag{2}$$

We now split the irreducible components $(W_{n,j})_{j \leq r_n}$ into two groups: those of type A_p, B_p or D_p for some $p \geq p_0$, which we call *large* and those of other types to which we will refer to as *small*. Up to reordering, we can assume that there is an index $q_n = q_n(\varepsilon)$ such that large components are exactly those with $j \leq q_n$.

We further write $s_{n,+}^2 = \sum_{j=1}^{q_n} s_{n,j}^2$ and $s_{n,-}^2 = \sum_{j=q_n+1}^{r_n} s_{n,j}^2$. We also introduce

$$\widetilde{\boldsymbol{t}_{n,+}} = \sum_{j=1}^{q_n} \tfrac{s_{n,j}}{s_{n,+}} \widetilde{\boldsymbol{t}_{n,j}}, \quad \widetilde{\boldsymbol{t}_{n,-}} = \sum_{j=q_n+1}^{r_n} \tfrac{s_{n,j}}{s_{n,-}} \widetilde{\boldsymbol{t}_{n,j}},$$

so that the renormalized two-sided descent statistics decomposes into a large component part and a small component part

$$\widetilde{\boldsymbol{t}_n} = \tfrac{s_{n,+}}{s_n} \widetilde{\boldsymbol{t}_{n,+}} + \tfrac{s_{n,-}}{s_n} \widetilde{\boldsymbol{t}_{n,-}}.$$

(Summands in the right-hand-side of these equations are independent.)

Estimates for the large component part Using the definition of large components and Eq. (2), we have that $d_2(\widetilde{\boldsymbol{t}_{n,j}}, Z) \le \varepsilon$ for $j \le q_n$. From Lemma 4, this implies

$$d_2\big(\widetilde{\boldsymbol{t}_{n,+}}, Z\big) = d_2\left(\sum_{j=1}^{q_n} \frac{s_{n,j}}{s_{n,+}} \widetilde{\boldsymbol{t}_{n,j}}, Z\right) = \sum_{j=1}^{q_n} \frac{s_{n,j}^2}{s_{n,+}^2} d_2\big(\widetilde{\boldsymbol{t}_{n,j}}, Z\big) \le \varepsilon \left(\sum_{j=1}^{q_n} \frac{s_{n,j}^2}{s_{n,+}^2}\right) = \varepsilon.$$

In the following equation, we assume that $\widetilde{\boldsymbol{t}_{n,+}}$ and Z are defined on the same probability space in such a way that $\big\|\widetilde{\boldsymbol{t}_{n,+}} - Z\big\|_2 \le \varepsilon$. Using that $u \mapsto \exp(iu)$ is a 1-Lipschitz function on $\mathbb{R}$, we have, for ζ in $\mathbb{R}$:

$$\begin{aligned}\Big|\mathbb{E}\big[\exp(i\zeta \tfrac{s_{n,+}}{s_n}\widetilde{\boldsymbol{t}_{n,+}})\big] - \exp(-\tfrac{\zeta^2 s_{n,+}^2}{2s_n^2})\Big| &\le \mathbb{E}\Big[\big|\exp(i\zeta \tfrac{s_{n,+}}{s_n}\widetilde{\boldsymbol{t}_{n,+}}) - \exp(i\zeta \tfrac{s_{n,+}}{s_n} Z)\big|\Big] \\ &\le \tfrac{s_{n,+}}{s_n}|\zeta|\,\mathbb{E}\Big[\big|\widetilde{\boldsymbol{t}_{n,+}} - Z\big|\Big] \le |\zeta|\,\big\|\widetilde{\boldsymbol{t}_{n,+}} - Z\big\|_2 \le |\zeta|\,\varepsilon, \quad (3)\end{aligned}$$

where the second to last inequality uses $\frac{s_{n,+}}{s_n} \le 1$ and Cauchy-Schwartz inequality.

Estimates for the small component part Here, we will use classical characteristic function estimates, as used in Lindeberg central limit theorem (see, *e.g.*, [Bil86, Theorem 27.2]). By definition, small components are of some exceptional type, of type $I_2(m)$ or of type A_p, B_p or D_p for $p < p_0$. Their rank is therefore at most $\max(8, p_0)$ ($I_2(m)$ has rank 2, the largest exceptional group E_8 has rank 8 and A_p, B_p or D_p have rank p). But the two sided-descent statistics on any Coxeter group W cannot exceed $2\,\mathrm{rk}(W)$. We conclude that there is a uniform bound $K = K(\varepsilon) = 2\max(8, p_0)$ on all the $\boldsymbol{t}_{n,j}$ corresponding to small components ($j > q_n$). In particular, for $j > q_n$, we have $s_{n,j} \le K$.

Fix ζ in $\mathbb{R}$. Using, the definition of $\widetilde{\boldsymbol{t}_{n,-}}$, we have

$$\mathbb{E}\Big[\exp\big(i\zeta \tfrac{s_{n,-}}{s_n}\widetilde{\boldsymbol{t}_{n,-}}\big)\Big] = \prod_{j=q_n+1}^{r_n} \mathbb{E}\Big[\exp\big(i\tfrac{\zeta}{s_n}(\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}])\big)\Big] \quad (4)$$

We assumed $\lim s_n = +\infty$ and argued above that $s_{n,j}$ is uniformly bounded for $j > q_n$. Thus, for n sufficiently large and $j > q_n$, we have $\frac{\zeta^2}{s_n^2} s_{n,j}^2 \le 1$. This implies (see [Bil86, eqs. (27.11) and (27.15)] that, for $j > q_n$, we have

$$\left|\mathbb{E}\Big[\exp\big(i\tfrac{\zeta}{s_n}(\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}])\big)\Big] - \exp\big(-\tfrac{\zeta^2 s_{n,j}^2}{2s_n^2}\big)\right| \le \tfrac{|\zeta|^3}{s_n^3}\mathbb{E}\Big[|\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}]|^3\Big] + \tfrac{|\zeta|^4}{s_n^4} s_{n,j}^4 \quad (5)$$

Since $\boldsymbol{t}_{n,j}$ is bounded by K, we have

$$\mathbb{E}\Big[|\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}]|^3\Big] \le K\,\mathbb{E}\Big[|\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}]|^2\Big] = K s_{n,j}^2.$$

Using also $s_{n,j}^4 \le K^2 s_{n,j}^2$ and taking n large enough so that $|\zeta| \le s_n$, we can simplify the upper bound in (5) to

$$\left|\mathbb{E}\Big[\exp\big(i\tfrac{\zeta}{s_n}(\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}])\big)\Big] - \exp\big(-\tfrac{\zeta^2 s_{n,j}^2}{2s_n^2}\big)\right| \le 2K^2|\zeta|^3 \tfrac{s_{n,j}^2}{s_n^3}. \quad (6)$$

We now use the following basic inequality: if $(a_i)_{i\le t}$ and $(b_i)_{i\le t}$ are collections of numbers of absolute values at most one, then $\big|\prod_{i\le t} a_i - \prod_{i\le t} b_i\big| \le \sum_{i\le t} |a_i - b_i|$ (see, *e.g.*, [Bil86, eq. (27.3)]). Therefore, (6) implies

$$\begin{aligned}&\left|\prod_{j=q_n+1}^{r_n} \mathbb{E}\Big[\exp\big(i\tfrac{\zeta}{s_n}(\boldsymbol{t}_{n,j} - \mathbb{E}[\boldsymbol{t}_{n,j}])\big)\Big] - \prod_{j=q_n+1}^{r_n} \exp\big(-\tfrac{\zeta^2 s_{n,j}^2}{2s_n^2}\big)\right| \\ &\qquad \le \sum_{j=q_n+1}^{r_n} 2K^2|\zeta|^3 \tfrac{s_{n,j}^2}{s_n^3} = 2K^2|\zeta|^3 \frac{s_{n,-}^2}{s_n^3} \le \frac{2K^2|\zeta|^3}{s_n}.\end{aligned}$$

The first term is the left-hand side is exactly $\mathbb{E}\Big[\exp\big(i\zeta\frac{s_{n,-}}{s_n}\widetilde{\boldsymbol{t}_{n,-}}\big)\Big]$; see (4). Since s_n tends to $+\infty$, the upper bound in the last display tends to 0. Therefore for n large enough, we have

$$\left|\mathbb{E}\Big[\exp\big(i\zeta\tfrac{s_{n,-}}{s_n}\widetilde{\boldsymbol{t}_{n,-}}\big)\Big]-\exp\big(-\tfrac{\zeta^2 s_{n,-}^2}{2s_n^2}\big)\right|\leq\varepsilon. \tag{7}$$

Conclusion of the proof. We recall that $s_n^2=s_{n,+}^2+s_{n,-}^2$ and $\widetilde{\boldsymbol{t}_n}=\frac{s_{n,+}}{s_n}\widetilde{\boldsymbol{t}_{n,+}}+\frac{s_{n,-}}{s_n}\widetilde{\boldsymbol{t}_{n,-}}$. Using again that $|a_1a_2-b_1b_2|\leq|a_1-b_1|+|a_2-b_2|$ for numbers of absolute values at most 1, Eqs. (3) and (7) imply that, for n large enough,

$$\left|\mathbb{E}\Big[\exp\big(i\zeta\widetilde{\boldsymbol{t}_n}\big)\Big]-\exp\big(-\tfrac{\zeta^2}{2}\big)\right|\leq(|\zeta|+1)\,\varepsilon.$$

Since this holds for any ε and any ζ in $\mathbb{R}$ (with a threshold value for n depending on ε and ζ), we have proved that the characteristic function of $\widetilde{\boldsymbol{t}_n}$ converges pointwise towards $\exp\big(-\frac{\zeta^2}{2}\big)$, which is the characteristic function of a Gaussian random variable. By Lévy's continuity theorem, this concludes our proof. □

Technical comment: a naive characteristic function estimates for the large component part would lead to an upper bound in (3) depending on the number q_n of large components. Since we have no control on this number, we would have not been able to conclude. Using the second Wasserstein distance avoids this problem.

A Coxeter groups

Coxeter groups have been introduced by Coxeter in the '30s [Cox34, Cox35]. They are now standard objects in combinatorial geometry; we give here a short introduction to the topic to make this note self-contained. Classical references are [Hum92, Bou02, BB05].

A Coxeter matrix $M=(m_{ij})_{i,j\in S}$ indexed by some set S is a symmetric matrix with entries in $\{1,2,3,\cdots\}\cup\{+\infty\}$ such that $m_{ij}=1$ if and only if $i=j$. A group W is a *Coxeter group* if one can find a set S of generators and a Coxeter matrix M indexed by S such that W admits the presentation

$$W\simeq\Big\langle s\in S\,\Big|\,(st)^{m_{st}}=1,\ s,t\in S\Big\rangle.$$

The pair (W,S) is then called a *Coxeter system*. When we consider a Coxeter group W, we often also consider a fixed set S such that (W,S) is a Coxeter system. The rank of a Coxeter group (or rather of a Coxeter system) is the size of S. Apart from this combinatorial definition, *finite* Coxeter groups can also be characterized geometrically: they are finite subgroups of general linear groups generated by reflections.

The direct product of two Coxeter groups is a Coxeter group. A Coxeter group (or rather a Coxeter system) is irreducible if it cannot be written as a direct product of two smaller Coxeter groups. Trivially, any finite Coxeter group is a direct product of irreducible factors. Finite irreducible Coxeter groups have been classified by Coxeter in 1935:

- there are three infinite families of increasing rank, commonly denoted A_n, B_n and D_n. the group A_n is the symmetric group on $n+1$ elements, B_n is the group of permutations of n elements with 2 colors, and D_n is a index 2 subgroup of B_n.
- there is one infinite family $I_2(m)$ of groups all of rank 2, called *dihedral groups*. These are the groups of symmetry of regular polygons.

- Finally, there are 6 *exceptional groups*, commonly denotes E_6, E_7, E_8, F_4, H_3 and H_4 (the index is always the rank or the group).

This classification and previous results for the infinite families are crucial in this note.

We end this appendix by defining the notion of *descent* in a Coxeter group studied in this note. This generalizes the notion of descents in permutations, corresponding to Coxeter groups of type A_n. For an element w in a Coxeter group W, we write $\ell(w)$ for the minimal number of factors needed to write w as a product of elements of S. Then, by definition, a generator s in S is a descent of w if $\ell(ws) < \ell(w)$.

References

[BB05] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, vol. **231**, Springer, 2005. MR-2133266

[BR19] B. Brück and F. Röttger. A central limit theorem for the two-sided descent statistic on Coxeter groups. Preprint arXiv:1908.07955, 2019.

[Bil86] P. Billingsley. *Probability and Measure*, 2nd edn. Wiley Series in Probability and Statistics, Wiley, 1986. MR-2893652

[Bou02] N. Bourbaki. *Lie Groups and Lie Algebras: Chapters 4–6*, Elements of Mathematics, Springer, 2002. MR-1890629

[CD17] 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. MR-3741694

[Cox34] H. S. M. Coxeter. Discrete groups generated by reflections, *Ann. Math.*, **35** (3): 588–621, 1934. MR-1503182

[Cox35] H. S. M. Coxeter. The complete enumeration of finite groups of the form $r_i^2 = (r_i r_j)^{k_{ij}} = 1$. *J. London Math. Soc.*, **10** (1): 21–25, 1935. MR-1575146

[EOM11] Wasserstein metric. *Encyclopedia of Mathematics*. Available at: this url.

[Ful04] J. Fulman. Stein's method and non-reversible Markov chains. *Lecture Notes-Monograph Series*, **46**: 66–74, 2004. MR-2118603

[Gon44] V. L. Goncharov. Some facts from combinatorics. Izv. Akad. Nauk SSSR, Ser. Mat. **8**: 3–48, 1944. [See also: On the field of combinatory analysis. Transl. Amer. Math. Soc. 19 (1962) 1-46]. MR-0131369

[Hoe51] W. Hoeffding. A combinatorial central limit theorem. *Ann. Math. Statist.* **22** (4): 558–566, 1951. MR-0044058

[Hum92] J. E. Humphreys. *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics, vol. **29**, Cambridge University Press, 1992. MR-1066460

[KS20] T. Kahle and C. Stump. Counting inversions and descents of random elements in finite Coxeter groups. *Math. Comp.*, **89**: 437–464, 2020. MR-4011551

[Mal72] C. L. Mallows. A note on asymptotic joint normality. *Ann. Math. Statist.*, **43** (2): 508–515, 1972. MR-0298812

[Özd19] A. Y. Özdemir. Martingales and descent statistics. Preprint arXiv:1901.01719, 2019.

[Röt18] F. Röttger. Asymptotics of a locally dependent statistic on finite reflection groups. Preprint arXiv:1812.00372, 2018.

[Vat96] 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. MR-1388382

[Was69] L.N. Wasserstein. Markov processes over denumerable products of spaces describing large systems of automata. *Probl. Inform. Transmission*, **5**: 47–52, 1969. MR-0314115

Acknowledgments. I would like to thank Christian Stump for introducing me to the subject of central limit theorems in Coxeter groups in the "Einstein Workshop on Algebraic Combinatorics" in Berlin, November 2018. I learned about Mallow's paper [Mal72] through the *Encyclopedia of Mathematics* page on Wasserstein metric [EOM11].