EXTREMES OF JOINT INVERSIONS AND DESCENTS ON FINITE COXETER GROUPS

PHILIP DÖRR AND JOHANNES HEINY

Abstract. The numbers of inversions and descents of random permutations are known to be asymptotically normal. On general finite Coxeter groups, the central limit theorem (CLT) is still valid under mild conditions. The extreme values of these two statistics are attracted by the Gumbel distribution. The joint distribution of inversions and descents is a likewise interesting object, but only the CLT on symmetric groups has been established thus far. In this paper, we comprehensively extend the knowledge of the joint distribution of inversions and descents. We prove both the CLT and the extreme value attraction for the joint distribution of inversions and descents by using Hájek projections and a suitable Gaussian approximation. On the signed permutation groups, we additionally show that these results are still valid when the choice of the random signs is biased. Furthermore, we investigate the applicability of these techniques to products of classical Weyl groups.

1. Introduction

The numbers of inversions and descents are two of the most important characteristics of permutations or, more generally, of Coxeter group elements. On the symmetric group S_n of permutations $\pi\colon \{1,\dots,n\} \to \{1,\dots,n\}$, an inversion is any tuple (i,j) with $i<j$, but $\pi(i)>\pi(j)$. A descent is any inversion of two adjacent numbers, that is, any i with $\pi(i)>\pi(i+1)$. Symmetric groups belong to the class of finite Coxeter groups, on which the concept of inversions and descents can be readily generalized, see [3, Section 1.4]. We can treat the underlying Coxeter group as a probability space and draw its elements at random. The study of stochastic properties of quantities such as the number of inversions and descents belongs to the field of statistical algebra and is the aim of this paper. We use the notations X_{inv} and X_{des} to indicate the nature of these numbers as random variables. In all following scenarios, we suppose that elements of a symmetric group or finite Coxeter group are drawn *uniformly* at random, unless stated otherwise. For some basics on finite Coxeter groups, we refer to [3].

This paper will focus on the classes S_n (symmetric groups), B_n (signed permutation groups) and D_n (even-signed permutation groups), for which we use the umbrella term *classical Weyl groups* throughout.

The asymptotic distribution of X_{inv} and X_{des} has been investigated by several authors (see, e.g., [2, 6, 16, 17, 18, 23]), who showed that inversions and descents each satisfy a central limit theorem (CLT) and thus are asymptotically normal. We say that a family of real-valued

2010 *Mathematics Subject Classification.* Primary: 60G70, 05A16; Secondary: 20F55, 62R01.

Key words and phrases. Permutation statistics, joint distribution, extreme value theory, central limit theorem, maximum, Coxeter group.

random variables $X_1, X_2, \dots$ (or their respective distributions) satisfies the CLT if

$$\frac{X_n - \mathbb{E}(X_n)}{\mathrm{Var}(X_n)^{1/2}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1)\,, \qquad n \to \infty\,,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. If the X_n are sums of independent random variables, one can verify the classical Lindeberg or Lyapunov conditions to establish a CLT.

For the number of inversions or descents on finite Coxeter groups, there are several techniques and approaches toward the CLT. Chatterjee & Diaconis provide an overview of proofs of the CLT for X_{des} on the family of symmetric groups in [6, Section 3]. One approach is based on a representation of X_{des} via m-dependent variables, which will be essential for our own methods in Section 2. Alternative proofs of the asymptotic normality of X_{des} are based on the zeros of its generating function [17, 23] or on certain regularity properties of the generating function, see [2, Ex. 3.5 and 5.3]. Stein's method of exchangeable pairs has been applied in [10, 16], where the latter reference and [2, Ex. 5.5] also cover inversions. We note that Stein's method has been used in various settings, see, e.g., [11] for permutations on multisets, [22] for generalized inversions, and [1] for non-uniformly random permutations.

Kahle & Stump [18] developed a full characterization of CLTs for inversions and descents on sequences of finite Coxeter groups, by giving necessary and sufficient conditions on the asymptotics of $\mathrm{Var}(X_{\mathrm{inv}})$ and $\mathrm{Var}(X_{\mathrm{des}})$. The work of Dörr & Kahle [13] was the first to provide extreme value theory for these classical permutation statistics. They showed that the numbers of inversions and descents are in the maximum-domain of attraction of the standard Gumbel distribution, assuming a triangular array where the number of samples per row obeys an exponential upper bound.

While the asymptotic normality of inversions and descents as *univariate statistics* is well studied, the knowledge of their *joint distribution* is comparatively sparse. Fang & Röllin [15] gave a CLT for arbitrarily large collections of permutation statistics based on antisymmetric matrices, including $(X_{\mathrm{inv}}, X_{\mathrm{des}})$ as a special case. Their work can also be seen as a multivariate extension of [16]. The aforementioned papers based on Stein's method typically achieve an $O(n^{-1/2})$ rate of convergence. Our novel approach will not achieve this rate, but it will generalize the multivariate CLT to other classical Weyl groups and even further, it will cover the asymptotic extreme value behavior of $(X_{\mathrm{inv}}, X_{\mathrm{des}})$.

This paper is structured as follows. Section 2 introduces the Hájek projection of inversions and descents on symmetric groups and justifies its use to approximate X_{inv}. Section 3 introduces the Gaussian approximation result on m-dependent random vectors by Chang *et al.* [5], which we use to prove the asymptotic normality of $(X_{\mathrm{inv}}, X_{\mathrm{des}})$. Section 4 presents the corresponding extreme value limit theorem (EVLT) as the main result of this paper. In Section 5, these results are extended to the larger groups B_n and D_n, which are equipped with a new family of probability measures, namely the so-called p-biased signed permutations. Section 6 studies the CLT and EVLT for direct products of classical Weyl groups, and Section 7 concludes the paper with some open questions. The remaining technical proofs are gathered in the appendix in Section 8.

Throughout this paper, we denote the standard uniform distribution on the interval $[0, 1]$ by $U(0, 1)$, and the discrete uniform distribution on the set $\{0, 1, \dots, n\}$ by $U(\{0, 1, \dots, n\})$. We sometimes also use these notations for accordingly distributed random variables when the meaning is clear from the context. The symbol $\stackrel{\mathcal{D}}{=}$ means equality in distribution and $\sum_{i<j}$ will be used as an abbreviation of the double-indexed sum $\sum_{1\leq i<j\leq n}$. For any random variable X, we denote its standard deviation by $\sigma(X) = \sqrt{\mathrm{Var}(X)}$. Moreover, we use typical Landau notation for positive sequences a_n, b_n as follows:

- $a_n = O(b_n)$ means that a_n grows at most as fast as b_n, i.e., $\limsup_{n\to\infty} a_n/b_n < \infty$.
- $a_n = o(b_n)$ means that a_n grows slower than b_n, i.e., $\lim_{n\to\infty} a_n/b_n = 0$. This is also written as $a_n \ll b_n$ or $b_n \gg a_n$.
- $a_n = \Theta(b_n)$ means that a_n and b_n have the same order of magnitude, i.e., both $a_n = O(b_n)$ and $b_n = O(a_n)$ hold.
- $a_n = b_n + o_{\mathbb{P}}(1)$ means that a_n, b_n are sequences of random variables with $a_n - b_n \xrightarrow{\mathbb{P}} 0$.

2. Hájek projections on symmetric groups

A permutation drawn uniformly at random from the symmetric group S_n of permutations on n letters is induced by the ranks of independent random variables $Z_1, \ldots, Z_n \sim U(0,1)$. Thus, the numbers of inversions and descents of this permutation can be represented by

$$X_{\mathrm{inv}} = \sum_{1\leq i<j\leq n} \mathbf{1}\{Z_i > Z_j\}, \tag{1}$$

$$X_{\mathrm{des}} = \sum_{i=1}^{n-1} \mathbf{1}\{Z_i > Z_{i+1}\}. \tag{2}$$

The primary challenge in dealing with the joint permutation statistic $(X_{\mathrm{inv}}, X_{\mathrm{des}})$ is the dependence structure between X_{inv} and X_{des}. The random variables $\mathbf{1}\{Z_i > Z_j\}, 1 \leq i < j \leq n$ in (1) are also dependent. It is worth noting that X_{inv} has the following representation as a sum of $n-1$ independent terms:

$$X_{\mathrm{inv}} \stackrel{\mathcal{D}}{=} \sum_{i=1}^{n-1} U(\{0, 1, \ldots, i\})\,. \tag{3}$$

This follows, e.g., from [13, Corollary 2.5a)] within the framework of all finite Coxeter groups. The representation of X_{des} in (2) has m-dependent variables (precisely, $m = 1$). There is also a decomposition of X_{des} into independent summands, based on the splitting of its generating function (known as the *Eulerian polynomial*) into linear factors of its real-valued roots [13, Corollary 2.5b)].

From the representations in [13, Corollary 2.5a)] and [13, Corollary 2.5b)] it is easy to derive the generating functions of X_{inv} and X_{des}. That X_{inv} and X_{des} are not independent follows from the fact that the joint generating function of $(X_{\mathrm{des}}, X_{\mathrm{inv}})$ does not factor into those of X_{inv} and X_{des}. It is an interesting question whether this polynomial factors at all. Some tests were made on S_3 or S_4, but there was no regularity detected so far.

To handle the dependence between X_{inv} and X_{des}, we approximate X_{inv} through its Hájek projection $\hat{X}_{\mathrm{inv}}$ and write $(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})$ as a sum of m-dependent two-dimensional random vectors. Then, we apply a Gaussian approximation theorem by Chang *et al.* [5] for triangular arrays of m-dependent random vectors. This will give a new proof of the CLT, and more importantly, it will be essential to derive the extreme value limit theorem (EVLT) for $(X_{\mathrm{inv}}, X_{\mathrm{des}})$ on classical Weyl groups.

Definition 2.1. For independent random variables $Z_1, \dots, Z_n$ and any random variable X_n, the *Hájek projection* of X_n with respect to $Z_1, \dots, Z_n$ is given by

$$\hat{X}_n := \sum_{k=1}^{n} \mathbb{E}(X_n \mid Z_k) - (n-1)\mathbb{E}(X_n).$$

Note that $\mathbb{E}(\hat{X}_n) = \mathbb{E}(X_n)$. Since each $\mathbb{E}(X_n \mid Z_k)$ is a measurable function only in Z_k, the Hájek projection is a sum of independent random variables, regardless of the original dependence structure between X_n and Z_k. In Sections 2–4, we always assume that $Z_1, \dots, Z_n \sim U(0,1)$ for our purposes. To decide whether the Hájek projection is a sufficiently accurate approximation, the following criterion is useful.

Theorem 2.2 (cf. [24], Theorem 11.2)**.** *Consider a sequence $(X_n)_{n\geq 1}$ of random variables and their associated Hájek projections $(\hat{X}_n)_{n\geq 1}$. If $\mathrm{Var}(\hat{X}_n) \sim \mathrm{Var}(X_n)$ as $n \to \infty$, then*

$$\frac{X_n - \mathbb{E}(X_n)}{\mathrm{Var}(X_n)^{1/2}} = \frac{\hat{X}_n - \mathbb{E}(\hat{X}_n)}{\mathrm{Var}(\hat{X}_n)^{1/2}} + o_{\mathbb{P}}(1).$$

In particular, if $\mathrm{Var}(\hat{X}_n) \sim \mathrm{Var}(X_n)$ and $(\hat{X}_n)_{n\geq 1}$ satisfies a CLT, then Theorem 2.2 guarantees that $(X_n)_{n\geq 1}$ also satisfies a CLT.

In what follows, for a random variable or vector X with finite variance, we write Y for its standardization, that is, $Y = \big(X - \mathbb{E}(X)\big)/\sqrt{\mathrm{Var}(X)}$. In particular, Y_{inv} is the standardization of X_{inv} and $\hat{Y}_{\mathrm{inv}}$ is that of $\hat{X}_{\mathrm{inv}}$. We use the symbols $X_{\mathrm{inv}}, Y_{\mathrm{inv}}, \hat{X}_{\mathrm{inv}}, \hat{Y}_{\mathrm{inv}}, X_{\mathrm{des}}, Y_{\mathrm{des}}$ with suppression of n, the underlying symmetric group or its rank, unless needed for clarification.

The next result provides the Hájek projection of X_{inv} defined in (1) and verifies the variance equivalence condition stated in Theorem 2.2 for X_{inv}.

Lemma 2.3. *The Hájek projection $\hat{X}_{\mathrm{inv}}$ of X_{inv} is given by*

$$\hat{X}_{\mathrm{inv}} = \frac{n(n-1)}{4} + \sum_{k=1}^{n} (n-2k+1) Z_k$$

and it holds that $\mathrm{Var}(X_{\mathrm{inv}}) \sim \mathrm{Var}(\hat{X}_{\mathrm{inv}})$ as $n \to \infty$.

Proof. We first study the conditional expectations $\mathbb{E}(X_{\mathrm{inv}} \mid Z_k)$ for $k = 1, \dots, n$ and get

$$\mathbb{E}(X_{\mathrm{inv}} \mid Z_k) = \sum_{1\leq i<j\leq n} \mathbb{P}(Z_i > Z_j \mid Z_k) = \sum_{1\leq i<j\leq n} \begin{cases} 1/2, & \text{if } k \notin \{i,j\}, \\ Z_k, & \text{if } k = i, \\ 1 - Z_k, & \text{if } k = j. \end{cases}$$

We fix $k \in \{1, \dots, n\}$ and analyze the frequency of the three cases on the right-hand side. As $\{1, \dots, n\} \setminus \{k\}$ has cardinality $n-1$, there are $\binom{n-1}{2}$ subsets $\{i,j\} \subseteq \{1, \dots, n\} \setminus \{k\}$. When picking $i = k$, there are $n-k$ indices j with $j > k$, for which we have $\mathbb{P}(Z_k > Z_j \mid Z_k) = \mathbb{P}(Z_j < Z_k \mid Z_k) = Z_k$ since $Z_k \sim U(0,1)$. Likewise, when picking $j = k$, there are

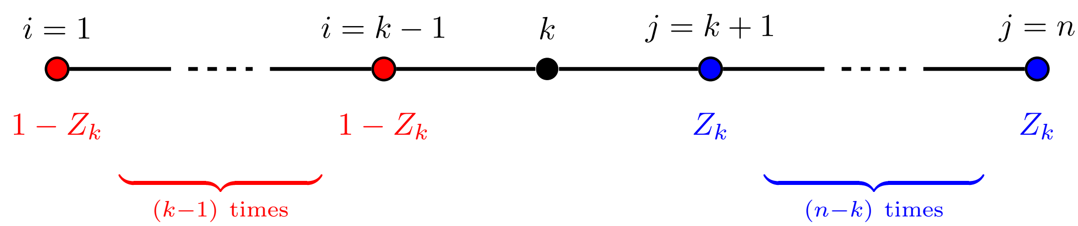

FIGURE 1. Display of the non-constant contributions to $\mathbb{E}(X_{\text{inv}} \mid Z_k)$

$k-1$ indices i with $i<k$, which gives $\mathbb{P}(Z_i > Z_k \mid Z_k) = 1 - Z_k$. These contributions are illustrated in Figure 1. Therefore, we obtain

$$\begin{aligned}\mathbb{E}(X_{\text{inv}} \mid Z_k) &= \frac{1}{2}\binom{n-1}{2} + (n-k)Z_k + (k-1)(1-Z_k)\\ &= \frac{1}{2}\binom{n-1}{2} + (n-2k+1)Z_k + (k-1),\end{aligned}$$

from which we deduce that

$$\begin{aligned}\hat{X}_{\text{inv}} &= \sum_{k=1}^{n} \mathbb{E}(X_{\text{inv}} \mid Z_k) - (n-1)\mathbb{E}(X_{\text{inv}})\\ &= \frac{n}{2}\binom{n-1}{2} + \sum_{k=1}^{n}(n-2k+1)Z_k + \sum_{k=1}^{n}(k-1) - \frac{n-1}{2}\binom{n}{2}\\ &= \frac{n(n-1)}{4} + \sum_{k=1}^{n}(n-2k+1)Z_k. \qquad (4)\end{aligned}$$

Since the Z_k's are i.i.d. , the variance of the Hájek projection is

$$\operatorname{Var}(\hat{X}_{\text{inv}}) = \sum_{k=1}^{n} \operatorname{Var}((n-2k+1)Z_k).$$

Due to $\operatorname{Var}(Z_k) = 1/12$, we get

$$\begin{aligned}\operatorname{Var}(\hat{X}_{\text{inv}}) &= \frac{1}{12}\sum_{k=1}^{n}(2k-n-1)^2 = \frac{1}{12}\sum_{k=1}^{n}\left(4k^2 + (n+1)^2 - 4k(n+1)\right)\\ &= \frac{1}{12}\left(4\sum_{k=1}^{n}k^2 + n(n+1)^2 - 4(n+1)\frac{n(n+1)}{2}\right)\\ &= \frac{1}{12}\left(4\frac{n(n+1)(2n+1)}{6} - n(n+1)^2\right) = \frac{1}{36}n^3 - \frac{n}{36}\,, \qquad n \to \infty.\end{aligned}$$

By [18, Corollary 3.2], we have $\operatorname{Var}(X_{\text{inv}}) = \dfrac{1}{36}n^3 + \dfrac{9n^2+7n}{72}$ and therefore $\operatorname{Var}(X_{\text{inv}}) \sim \operatorname{Var}(\hat{X}_{\text{inv}})$ as $n \to \infty$. □

A combination of Theorem 2.2 and Lemma 2.3 yields

$$Y_{\text{inv}} = \hat{Y}_{\text{inv}} + o_{\mathbb{P}}(1)\,, \qquad n \to \infty\,.$$

Remark 2.4. Interestingly, this approach fails for X_{des}, since $\operatorname{Var}(X_{\text{des}}) \sim \operatorname{Var}(\hat{X}_{\text{des}})$ does not hold. Repeating the considerations in the proof of Lemma 2.3 for X_{des}, we first obtain

$$\mathbb{E}(X_{\text{des}} \mid Z_k) = \sum_{i=1}^{n-1} \mathbb{P}(Z_i > Z_{i+1} \mid Z_k) = \sum_{i=1}^{n-1} \begin{cases} 1/2, & k \notin \{i, i+1\}, \\ Z_k, & k = i, \\ 1 - Z_k, & k = i+1. \end{cases}$$

Now, except for the border cases $k = 1$ and $k = n$, the summands for $k = i$ and $k = i + 1$ are each used exactly once, so the Z_k in their sum $Z_k + (1 - Z_k)$ cancel out. In total, we obtain

$$\hat{X}_{\text{des}} = Z_1 - Z_n + c_n\,,$$

where c_n is some constant that depends only on n. Therefore, $\operatorname{Var}(\hat{X}_{\text{des}}) = 2/12$ does not have the linear order of $\operatorname{Var}(X_{\text{des}}) = (n + 1)/12$ (see [18, Corollary 4.2]).

For these reasons, our results will be based on the following consequence of Theorem 2.2 and Lemma 2.3.

Corollary 2.5. *Let* $(X_{\text{inv}}, X_{\text{des}})^\top$ *be given from the symmetric group* S_n*. For the standardized random vector* $(Y_{\text{inv}}, Y_{\text{des}})^\top$ *and the standardized Hájek projection* $\hat{Y}_{\text{inv}}$*, we have*

$$\begin{pmatrix} Y_{\text{inv}} \\ Y_{\text{des}} \end{pmatrix} = \begin{pmatrix} \hat{Y}_{\text{inv}} + o_{\mathbb{P}}(1) \\ Y_{\text{des}} \end{pmatrix}.$$

A decomposition of $(\hat{X}_{\text{inv}}, X_{\text{des}})^\top$ *into* 1*-dependent summands is given by*

$$\begin{pmatrix} \hat{X}_{\text{inv}} \\ X_{\text{des}} \end{pmatrix} = \sum_{k=1}^{n-1} \begin{pmatrix} (n - 2k + 1) Z_k \\ \mathbf{1}\{Z_k > Z_{k+1}\} \end{pmatrix} + \begin{pmatrix} -(n-1) Z_n + \frac{n(n-1)}{4} \\ 0 \end{pmatrix}. \tag{5}$$

Likewise, a 1*-dependent decomposition for* $(\hat{Y}_{\text{inv}}, Y_{\text{des}})^\top$ *can be found by standardization.*

It is worth noting that the correlation $\operatorname{Corr}(X_{\text{inv}}, X_{\text{des}})$ is not zero. However, we now show that $\operatorname{Corr}(X_{\text{inv}}, X_{\text{des}}) \to 0$ as $n \to \infty$. Moreover, by Corollary 2.5 the same holds true for $(\hat{X}_{\text{inv}}, X_{\text{des}})^\top$ as well (in fact, $\operatorname{Corr}(\hat{X}_{\text{inv}}, X_{\text{des}})$ is even easier to compute). To proceed, we need the covariance of $\hat{X}_{\text{inv}}$ and X_{des}. Our next result additionally provides $\operatorname{Cov}(X_{\text{inv}}, X_{\text{des}})$, which – to the best of our knowledge – is not available in the literature.

Lemma 2.6 (see Subsections 8.1, 8.2 for the proof)**.** *On the symmetric group* S_n*, we have*

a) $\operatorname{Cov}(X_{\text{inv}}, X_{\text{des}}) = (n - 1)/4$.
b) $\operatorname{Cov}(\hat{X}_{\text{inv}}, X_{\text{des}}) = (n - 1)/6$.

Corollary 2.7. *Since* $\operatorname{Var}(X_{\text{inv}})\operatorname{Var}(X_{\text{des}}) = \Theta(n^4)$ *according to* [18, Corollaries 3.2 and 4.2], *and the same holds true if* $\operatorname{Var}(X_{\text{inv}})$ *is replaced by* $\operatorname{Var}(\hat{X}_{\text{inv}})$*, we conclude from Lemma 2.6 that*

$$\operatorname{Corr}(X_{\text{inv}}, X_{\text{des}}) = \frac{\operatorname{Cov}(X_{\text{inv}}, X_{\text{des}})}{\sqrt{\operatorname{Var}(X_{\text{inv}})\operatorname{Var}(X_{\text{des}})}} = \Theta(1/n)\,,$$

$$\operatorname{Corr}(\hat{X}_{\text{inv}}, X_{\text{des}}) = \Theta(1/n)\,, \qquad n \to \infty.$$

3. The bivariate Central Limit Theorem

In this section, we establish the joint normality of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ by using the 1-dependent decomposition of $(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ and applying a recent (and quite optimized) CLT for m-dependent triangular arrays from Chang *et al.* [5]. Their work provides Gaussian approximations for high-dimensional data under various dependency frameworks, including m-dependence. It gives error rates over the system of all hyperrectangles, including the Kolmogorov distance as a special case. Moreover, the high-dimensional framework implicitly covers the finite-dimensional one by repeating the components of a vector.

Refining the notation of [5], we consider triangular arrays $(X_t^{(n)})_{t=1,\dots,n}$ whose entries $X_1^{(n)}, \dots, X_n^{(n)}$ are mean zero random vectors in $\mathbb{R}^{\mathfrak{p}}$, where $\mathfrak{p} = \mathfrak{p}(n)$ can grow with respect to n. For the sum

$$X^{(n)} := \sum_{t=1}^{n} X_t^{(n)} \quad \text{with covariance matrix } \Sigma^{(n)} := \mathrm{Var}(X^{(n)}), \tag{6}$$

the work of Chang *et al.* [5] gives bounds and rates of convergence for

$$r_n(\mathcal{A}) := \sup_{A \in \mathcal{A}} |\mathbb{P}(X^{(n)} \in A) - \mathbb{P}(\mathcal{N}_n \in A)|,$$

where $\mathcal{A}$ encompasses a system of Borel sets and $\mathcal{N}_n \sim \mathrm{N}(0, \Sigma^{(n)})$ is a normal distribution with the same covariance structure as $X^{(n)}$. Bounds for $r_n(\mathcal{A})$ in both constant and high dimensions have been strongly investigated for *independent* variables. A seminal work for the system $\mathcal{A} := \{[\mathbf{x}, \infty) \mid \mathbf{x} \in \mathbb{R}^{\mathfrak{p}}\}$ to compare the maxima of $X^{(n)}$ and $\mathcal{N}_n$ in high dimensions is given by Chernozhukov *et al.* [7]. In recent years, there have been great efforts to improve the error bounds and the growth of dimension within the independent framework, see [8, 9, 12, 14, 19].

An interesting feature of [5] is that the $X_t^{(n)}$ are allowed to be dependent which offers a wide range of applications beyond the independent framework, including the 1-dependent decomposition of $(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})^\top$. The following two conditions need to be imposed on the $X_t^{(n)} = (X_{t,1}^{(n)}, \dots, X_{t,\mathfrak{p}}^{(n)})^\top$.

Condition 1: There exists a sequence of constants $B_n \geq 1$ and a universal constant $\gamma_1 \geq 1$ such that

$$\max_{j=1,\dots,\mathfrak{p}} \mathbb{E}\left(\exp\left(\left|\sqrt{n} X_{t,j}^{(n)}\right|^{\gamma_1} B_n^{-\gamma_1}\right)\right) \leq 2\,, \qquad t = 1, \dots, n.$$

Condition 2: There exists a constant $K > 0$ such that for all $n \in \mathbb{N}$

$$\min_{j=1,\dots,\mathfrak{p}} \mathrm{Var}\left(\sum_{t=1}^{n} X_{t,j}^{(n)}\right) \geq K.$$

Remark 3.1. Condition 1 means sub-Gaussianity, i.e., by Markov's inequality,

$$\forall u > 0: \quad \mathbb{P}(|\sqrt{n} X_{t,j}^{(n)}| > u) \leq 2\exp(-u^{\gamma_1} B_n^{-\gamma_1}).$$

For sub-Gaussian variables, particularly for the bounded variables $X_{\mathrm{des}}, X_{\mathrm{inv}}$ and their Hájek projection $\hat{X}_{\mathrm{inv}}$, we can choose $\gamma_1 = 2$ and $B_n = O(1)$. Condition 2 implies non-degeneracy, which obviously holds true in our setting.

A very useful estimate of $r_n(\mathcal{A}^{\mathrm{re}})$ is given for the system of hyperrectangles

$$\mathcal{A}^{\mathrm{re}} := \Big\{\{\mathbf{w} \in \mathbb{R}^{\mathfrak{p}} \colon \mathbf{a} \le \mathbf{w} \le \mathbf{b}\} \mid \mathbf{a}, \mathbf{b} \in [-\infty, \infty]^{\mathfrak{p}}\Big\}.$$

Proposition 3.2 (see [5], Corollary 1)**.** *Let $(X_t^{(n)})_{t=1,\dots,n}$ be a triangular array of mean zero random vectors in high dimensions, i.e., $X_1^{(n)}, \dots, X_n^{(n)} \in \mathbb{R}^{\mathfrak{p}}$ with $\mathfrak{p} = \mathfrak{p}(n) \gg n^{\kappa}$ for a constant $\kappa > 0$. Assume that each row $X_1^{(n)}, \dots, X_n^{(n)}$ is m-dependent with a global constant $m \in \mathbb{N}$. Under Conditions 1 and 2, it holds that*

$$r_n(\mathcal{A}^{\mathrm{re}}) = O\left(\frac{B_n m^{2/3} \log(\mathfrak{p})^{7/6}}{n^{1/6}}\right), \qquad n \to \infty.$$

We note that if $\mathfrak{p}$ remains constant, we can artificially repeat the vector components (say n times) and therefore, the requirement $\mathfrak{p} \gg n^{\kappa}$ can be removed. We obtain the following corollary.

Corollary 3.3. *Let $(X_t^{(n)})_{t=1,\dots,n}$ be a triangular array of mean zero random vectors in fixed dimension $\mathfrak{p}$ and suppose that each row $X_1^{(n)}, \dots, X_n^{(n)}$ is m-dependent with a global constant $m \in \mathbb{N}$. Under Conditions 1 and 2, it holds that*

$$r_n(\mathcal{A}^{\mathrm{re}}) = O\left(\frac{B_n m^{2/3} \log(n)^{7/6}}{n^{1/6}}\right), \qquad n \to \infty.$$

Our next result establishes a CLT for the joint distribution of inversions and descents.

Theorem 3.4. *The joint distribution of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ on the family of symmetric groups satisfies the CLT. This means*

$$(Y_{\mathrm{inv}}, Y_{\mathrm{des}})^\top = \left(\frac{X_{\mathrm{inv}} - \mathbb{E}(X_{\mathrm{inv}})}{\sqrt{\mathrm{Var}(X_{\mathrm{inv}})}}, \frac{X_{\mathrm{des}} - \mathbb{E}(X_{\mathrm{des}})}{\sqrt{\mathrm{Var}(X_{\mathrm{des}})}}\right)^\top \xrightarrow{\mathcal{D}} \mathrm{N}_2(0, \mathrm{I}_2), \qquad n \to \infty.$$

Proof. Due to Corollary 2.5 and Slutsky's Theorem, it suffices to show that $(\hat{Y}_{\mathrm{inv}}, Y_{\mathrm{des}})^\top \xrightarrow{\mathcal{D}} \mathrm{N}_2(0, \mathrm{I}_2)$. By (5), we have that

$$\begin{pmatrix} \hat{X}_{\mathrm{inv}} - \mathbb{E}(\hat{X}_{\mathrm{inv}}) \\ X_{\mathrm{des}} - \mathbb{E}(X_{\mathrm{des}}) \end{pmatrix} = \sum_{k=1}^{n-1} \begin{pmatrix} (n-2k+1)(Z_k - 1/2) \\ \mathbf{1}\{Z_k > Z_{k+1}\} - 1/2 \end{pmatrix} + \begin{pmatrix} -(n-1)(Z_n - 1/2) \\ 0 \end{pmatrix} \tag{7}$$

is a sum of 1-dependent random vectors with mean zero. Setting

$$X_k^{(n)} := \begin{pmatrix} (n-2k+1)(Z_k - 1/2)/\sqrt{\mathrm{Var}(\hat{X}_{\mathrm{inv}})} \\ (\mathbf{1}\{Z_k > Z_{k+1}\} - 1/2)/\sqrt{\mathrm{Var}(X_{\mathrm{des}})} \end{pmatrix}, \qquad k = 1, \dots, n-1,$$

$$X_n^{(n)} := \begin{pmatrix} -(n-1)(Z_n - 1/2)/\sqrt{\mathrm{Var}(\hat{X}_{\mathrm{inv}})} \\ 0 \end{pmatrix},$$

we obtain the representation $(\hat{Y}_{\mathrm{inv}}, Y_{\mathrm{des}})^\top = \sum_{k=1}^n X_k^{(n)} =: X^{(n)}$. The covariance matrix of $X^{(n)}$ (see (6)) is given by $\Sigma^{(n)} = \begin{pmatrix} 1 & \rho_n \\ \rho_n & 1 \end{pmatrix}$, where $\rho_n := \mathrm{Corr}(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})$. An application of

Corollary 3.3 yields that for $\mathcal{N}_n \sim \mathrm{N}(0, \Sigma^{(n)})$,

$$\sup_{u \in \mathbb{R}^2} |\mathbb{P}(X^{(n)} \leq u) - \mathbb{P}(\mathcal{N}_n \leq u)| \leq r_n(\mathcal{A}^{\mathrm{re}}) = O\left(n^{-1/6} \log(n)^{7/6}\right) .$$

In combination with the fact that the correlation ρ_n vanishes in the limit (see Corollary 2.7), we can conclude that $(\hat{Y}_{\mathrm{inv}}, Y_{\mathrm{des}})^\top \xrightarrow{\mathcal{D}} \mathrm{N}_2(0, \mathrm{I}_2)$, completing the proof of the theorem. □

4. The Extreme Value Limit Theorem

In what follows, we use the Gaussian approximation of Proposition 3.2 to prove that the vector of componentwise maxima of i.i.d. copies of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ is attracted to the bivariate Gumbel distribution with independent margins. First, we briefly recapitulate the procedure used in [13] for the univariate case. Let $(k_n)_{n\in\mathbb{N}}$ be a divergent sequence of positive integers. It is well known (see, e.g., [20, Theorem 1.5.3]) that the maximum of k_n i.i.d. standard normal variables $(N_i)_{i=1,\ldots,k_n}$ is attracted toward the standard Gumbel distribution $\Lambda(x) = \exp(-\exp(-x))$ by virtue of

$$\alpha_n\big(\max\{N_1, \ldots, N_{k_n}\} - \alpha_n\big) \xrightarrow{\mathcal{D}} \Lambda, \qquad n \to \infty,$$

where

$$\alpha_n := \sqrt{2 \log k_n} - \frac{\log\log k_n + \log(4\pi)}{2\sqrt{2\log k_n}}.$$

The key step in the results of [13] was to employ large deviations theory in order to establish the tail equivalence

$$\mathbb{P}\big(X/\sqrt{\mathrm{Var}(X)} > x_n\big) \sim 1 - \Phi(x_n), \qquad n \to \infty\,,$$

where X denotes the centered version of either X_{inv} or X_{des} on S_n, and Φ is the standard normal distribution function and $x_n \to \infty$ is a suitable sequence of real numbers. In the multivariate case, we additionally need to control the dependence between X_{inv} and X_{des} which we facilitate by Proposition 3.2 after first replacing X_{inv} with its Hájek projection $\hat{X}_{\mathrm{inv}}$.

Let $(X_{\mathrm{inv}}^{(j)}, X_{\mathrm{des}}^{(j)})^\top, j = 1, \ldots, k_n$ be independent copies of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ on S_n. We are interested in the asymptotic joint distribution of the component-wise maxima of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$. Equivalently, we investigate the standardized maxima

$$M_{n,\mathrm{inv}} := \frac{\max_{j=1,\ldots,k_n} X_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}})}{\sigma(X_{\mathrm{inv}})} \quad \text{and} \quad M_{n,\mathrm{des}} := \frac{\max_{j=1,\ldots,k_n} X_{\mathrm{des}}^{(j)} - \mathbb{E}(X_{\mathrm{des}})}{\sigma(X_{\mathrm{des}})}\,. \tag{8}$$

We now postulate the main result of this paper. If the number of samples k_n is not too large, then the distribution of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ is in the maximum domain of attraction of a bivariate Gumbel distribution with independent margins.

Theorem 4.1. *Consider the setting from above and assume $(k_n \log k_n)/n \to 0$ as $n \to \infty$. Then, it holds that*

$$\lim_{n\to\infty} \mathbb{P}(\alpha_n(M_{n,\mathrm{inv}} - \alpha_n) \leq x, \alpha_n(M_{n,\mathrm{des}} - \alpha_n) \leq y) = \Lambda(x)\Lambda(y)\,, \qquad x, y \in \mathbb{R}\,. \tag{9}$$

In particular, the maxima of inversions and descents on S_n are asymptotically independent.

Proof. Let $(Z_i^j)_{i,j\geq 1}$ be a collection of independent $U(0,1)$ distributed random variables and recall that $\alpha_n \sim \sqrt{2\log k_n}$. Then we have the representation

$$\begin{pmatrix} X_{\mathrm{inv}}^{(j)} \\ X_{\mathrm{des}}^{(j)} \end{pmatrix}_{j=1,\dots,k_n} \stackrel{\mathcal{D}}{=} \begin{pmatrix} \sum_{1\leq i<k\leq n} \mathbf{1}\{Z_i^{(j)} > Z_k^{(j)}\} \\ \sum_{i=1}^{n-1} \mathbf{1}\{Z_i^{(j)} > Z_{i+1}^{(j)}\} \end{pmatrix}_{j=1,\dots,k_n}.$$

Therefore, by Slutsky's theorem, (9) is an immediate consequence of

$$\lim_{n\to\infty} \mathbb{P}(\alpha_n(\hat{M}_n - \alpha_n) \leq x, \alpha_n(M_{n,\mathrm{des}} - \alpha_n) \leq y) = \Lambda(x)\Lambda(y)\,, \qquad x,y \in \mathbb{R}\,, \tag{10}$$

and

$$\sqrt{\log k_n}\,|M_{n,\mathrm{inv}} - \hat{M}_n| \xrightarrow{\mathbb{P}} 0\,, \qquad n\to\infty\,, \tag{11}$$

where $\hat{M}_n := \sigma(\hat{X}_{\mathrm{inv}})^{-1}\Big(\max_{j=1,\dots,k_n} \hat{X}_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}})\Big)$. It remains to show (10) and (11). We begin with the proof of (11) and get

$$\begin{aligned} |M_{n,\mathrm{inv}} - \hat{M}_n| &\leq \max_{j=1,\dots,k_n} \left| \frac{X_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}})}{\sigma(X_{\mathrm{inv}})} - \frac{\hat{X}_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}})}{\sigma(\hat{X}_{\mathrm{inv}})} \right| \\ &= \max_{j=1,\dots,k_n} \left| \frac{X_{\mathrm{inv}}^{(j)} - \hat{X}_{\mathrm{inv}}^{(j)}}{\sigma(X_{\mathrm{inv}})} + (\hat{X}_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}}))\frac{\sigma(\hat{X}_{\mathrm{inv}}) - \sigma(X_{\mathrm{inv}})}{\sigma(X_{\mathrm{inv}})\sigma(\hat{X}_{\mathrm{inv}})} \right|. \end{aligned}$$

Thus, for any $\varepsilon > 0$, we obtain

$$\begin{aligned} \mathbb{P}\Big(\sqrt{\log k_n}\,|M_{n,\mathrm{inv}} - \hat{M}_n| > 2\varepsilon\Big) &\leq \mathbb{P}\Big(\sqrt{\log k_n} \max_{j=1,\dots,k_n} \left| \frac{X_{\mathrm{inv}}^{(j)} - \hat{X}_{\mathrm{inv}}^{(j)}}{\sigma(X_{\mathrm{inv}})} \right| > \varepsilon\Big) \\ &+ \mathbb{P}\Big(\sqrt{\log k_n} \max_{j=1,\dots,k_n} \left| (\hat{X}_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}}))\frac{\sigma(\hat{X}_{\mathrm{inv}}) - \sigma(X_{\mathrm{inv}})}{\sigma(X_{\mathrm{inv}})\sigma(\hat{X}_{\mathrm{inv}})} \right| > \varepsilon\Big) =: P_1 + P_2\,. \end{aligned}$$

Using the union bound and Markov's inequality, we have

$$\begin{aligned} P_1 &\leq k_n\,\mathbb{P}\Big(|X_{\mathrm{inv}} - \hat{X}_{\mathrm{inv}}| > \frac{\sigma(X_{\mathrm{inv}})\varepsilon}{\sqrt{\log k_n}}\Big) \leq k_n \frac{\log k_n}{\mathrm{Var}(X_{\mathrm{inv}})\varepsilon^2}\mathbb{E}|X_{\mathrm{inv}} - \hat{X}_{\mathrm{inv}}|^2 \\ &= \frac{k_n \log k_n}{\mathrm{Var}(X_{\mathrm{inv}})\varepsilon^2}\Big(\mathrm{Var}(X_{\mathrm{inv}}) + \mathrm{Var}(\hat{X}_{\mathrm{inv}}) - 2\mathrm{Cov}(X_{\mathrm{inv}}, \hat{X}_{\mathrm{inv}})\Big) \\ &= \frac{k_n \log k_n}{\varepsilon^2}\Big(1 - \frac{\mathrm{Var}(\hat{X}_{\mathrm{inv}})}{\mathrm{Var}(X_{\mathrm{inv}})}\Big)\,, \end{aligned} \tag{12}$$

where the last equality follows from the fact that $\mathrm{Cov}(X_{\mathrm{inv}}, \hat{X}_{\mathrm{inv}}) = \mathrm{Var}(\hat{X}_{\mathrm{inv}})$, see, e.g., [24, Theorem 11.1]. Plugging in the formulas for $\mathrm{Var}(X_{\mathrm{inv}})$ and $\mathrm{Var}(\hat{X}_{\mathrm{inv}})$ from the end of the proof of Lemma 2.3, we get $\mathrm{Var}(\hat{X}_{\mathrm{inv}})/\mathrm{Var}(X_{\mathrm{inv}}) = 1 + O(1/n)$ from which we conclude that

$$P_1 = k_n \log k_n\, O(1/n)\,, \qquad n\to\infty\,,$$

which tends to zero by the assumption on k_n. Repeating the above considerations for P_2 yields

$$P_2 \leq \frac{k_n \log k_n}{\varepsilon^2}\Big(\frac{\sigma(\hat{X}_{\mathrm{inv}}) - \sigma(X_{\mathrm{inv}})}{\sigma(X_{\mathrm{inv}})}\Big)^2 \mathbb{E}\Big(\frac{\hat{X}_{\mathrm{inv}}^{(j)} - \mathbb{E}(X_{\mathrm{inv}})}{\sigma(\hat{X}_{\mathrm{inv}})}\Big)^2$$

$$= k_n \log k_n \, O(1/n) = o(1)\,, \qquad n \to \infty\,, \tag{13}$$

which completes the proof of (11). Regarding (10), we recall that by Corollary 2.5,

$$\begin{pmatrix} \hat{X}^{(j)}_{\mathrm{inv}} - \mathbb{E}(\hat{X}^{(j)}_{\mathrm{inv}}) \\ X^{(j)}_{\mathrm{des}} - \mathbb{E}(X^{(j)}_{\mathrm{des}}) \end{pmatrix} = \sum_{k=1}^{n-1} \begin{pmatrix} (n-2k+1)(Z^{(j)}_k - 1/2) \\ \mathbf{1}\{Z^{(j)}_k > Z^{(j)}_{k+1}\} - 1/2 \end{pmatrix} + \begin{pmatrix} -(n-1)(Z^{(j)}_n - 1/2) \\ 0 \end{pmatrix}.$$

This is a sum of 1-dependent centered random vectors. Setting

$$Y^{(n,j)}_k := \begin{pmatrix} (n-2k+1)(Z^{(j)}_k - 1/2)/\sigma(\hat{X}_{\mathrm{inv}}) \\ (\mathbf{1}\{Z^{(j)}_k > Z^{(j)}_{k+1}\} - 1/2)/\sigma(X_{\mathrm{des}}) \end{pmatrix}, \qquad k = 1, \dots, n-1,$$

$$Y^{(n,j)}_n := \begin{pmatrix} -(n-1)(Z^{(j)}_n - 1/2)/\sigma(\hat{X}_{\mathrm{inv}}) \\ 0 \end{pmatrix},$$

we obtain the representation $(\hat{Y}^{(j)}_{\mathrm{inv}}, Y^{(j)}_{\mathrm{des}})^\top = \sum_{k=1}^n Y^{(n,j)}_k$. The covariance matrix of $(\hat{Y}^{(j)}_{\mathrm{inv}}, Y^{(j)}_{\mathrm{des}})^\top$ is given by $\Sigma^{(n)} = \begin{pmatrix} 1 & \rho_n \\ \rho_n & 1 \end{pmatrix}$, where $\rho_n := \mathrm{Corr}(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})$. For a centered normal random vector $\mathcal{N}_n = (N_1, \dots, N_{2k_n})^\top$ whose covariance matrix is block-diagonal with all k_n diagonal blocks equal to $\Sigma^{(n)}$, we write

$$P_n(x,y) := \mathbb{P}\left(\alpha_n \Big(\max_{j=1,\dots,k_n} N_{2j-1} - \alpha_n \Big) \le x, \alpha_n \Big(\max_{j=1,\dots,k_n} N_{2j} - \alpha_n \Big) \le y \right), \qquad x, y \in \mathbb{R}\,.$$

An application of Proposition 3.2 then yields, as $n \to \infty$,

$$\begin{aligned} &\left| \mathbb{P}(\alpha_n(\hat{M}_n - \alpha_n) \le x, \alpha_n(M_{n,\mathrm{des}} - \alpha_n) \le y) - P_n(x,y) \right| \\ &= O\left(n^{-1/6} \log(\max\{n, k_n\})^{7/6} \right) = o(1)\,. \end{aligned}$$

Finally, since $\rho_n \to 0$ (see Corollary 2.7), it is a standard result for maxima of bivariate Gaussian random vectors (with correlation strictly less than 1) that

$$P_n(x,y) \overset{n\to\infty}{\longrightarrow} \Lambda(x)\Lambda(y),$$

completing the proof of (10). □

Remark 4.2. The upper bound for the row-wise number of samples k_n is stricter than that for the individual statistics X_{inv} and X_{des} given in [13], due to the error that arises from replacing X_{inv} with $\hat{X}_{\mathrm{inv}}$. In particular, this excludes the choice of $k_n = n$ which gives a uniformly stretched triangular array. On the other hand, this new EVLT can be transferred to other individual and joint permutation statistics, since it is mainly based on a Gaussian approximation for m-dependent random vectors (or variables). Besides descents, some other examples of m-dependent permutation statistics on S_n are *peaks* (all indices i with $\pi(i-1) < \pi(i) > \pi(i+1)$) and *valleys* (all i with $\pi(i-1) > \pi(i) < \pi(i+1)$). Since the proof of Theorem 4.1 does not rely on any special property of descents other than m-dependence, any other m-dependent permutation statistic could be combined with inversions.

Furthermore, if we consider a permutation statistic consisting of one or two m-dependent components, then there is no need to use Hájek's projection and the corresponding part in the proof of Theorem 4.1 can be removed. In this case, only (10) needs to be shown and the upper bound of k_n only comes from

$$n^{-1/6}\log(k_n)^{7/6} = o(1) \Longleftrightarrow k_n = \exp\big(o(n^{1/7})\big).$$

Therefore, in this situation, Theorem 4.1 can be modified to give almost the same flexibility as [13, Theorem 4.2].

5. Signed and even-signed permutation groups

The previous two sections established the CLT and the EVLT for the joint distribution $(X_{\text{inv}}, X_{\text{des}})^\top$ on the family of symmetric groups. We now work toward these results for $(X_{\text{inv}}, X_{\text{des}})^\top$ on the *signed and even-signed permutation groups* B_n and D_n that generalize the symmetric groups S_n.

The group B_n of *signed permutations* on n letters arises from S_n by assigning any combination of positive or negative signs to the entries of a permutation. The *even-signed permutation group* D_n is the subgroup of B_n consisting of elements with an even number of negative signs. As in [18, Section 2.1], we use the in-line notation

$$\begin{aligned} B_n = \{\pi = (\pi(1), \ldots, \pi(n)) : \pi(i) \in \{\pm 1, \ldots, \pm n\}\ \forall i, \\ \{|\pi(1)|, \ldots, |\pi(n)|\} = \{1, \ldots, n\}\}, \end{aligned} \tag{14}$$

and we have $D_n = \{\pi \in B_n \colon \pi(1)\pi(2)\cdots\pi(n) > 0\}$.

According to [18], the combinatorial representation of inversions on B_n and D_n is given by

$$\text{Inv}(\pi) = \begin{cases} \text{Inv}^+(\pi) \cup \text{Inv}^-(\pi) \cup \text{Inv}^\circ(\pi), & \text{if } \pi \in B_n, \\ \text{Inv}^+(\pi) \cup \text{Inv}^-(\pi)\,, & \text{if } \pi \in D_n\,, \end{cases} \tag{15}$$

for the disjoint sets

$$\begin{aligned} \text{Inv}^+(\pi) &= \{1 \le i < j \le n \mid \pi(i) > \pi(j)\}, \\ \text{Inv}^-(\pi) &= \{1 \le i < j \le n \mid -\pi(i) > \pi(j)\}, \\ \text{Inv}^\circ\,(\pi) &= \{1 \le i \le n \mid \pi(i) < 0\}. \end{aligned}$$

Thus, the numbers of inversions on B_n and D_n are

$$\begin{aligned} X^B_{\text{inv}}(\pi) &= |\text{Inv}^+(\pi)| + |\text{Inv}^-(\pi)| + |\text{Inv}^\circ(\pi)|, \\ X^D_{\text{inv}}(\pi) &= |\text{Inv}^+(\pi)| + |\text{Inv}^-(\pi)|. \end{aligned}$$

The set Inv^+ is analogous to inversions on symmetric groups. Note that on B_n and D_n, one must pay attention to signs, that means, a pair (i, j) with $\pi(i), \pi(j) < 0$ and $|\pi(i)| < |\pi(j)|$ also adds to $\text{Inv}^+(\pi)$. The set Inv^- is that of negative sum pairs. The set $\text{Inv}^\circ(\pi)$ simply collects positions with negative entries. The latter two sets need to be counted so that $|\text{Inv}(\pi)|$ equals the word length on B_n with respect to the generating system $\{s_0, s_1, \ldots, s_{n-1}\}$, where $s_0 = (-1, 2, 3, \ldots, n)$ negates the first entry and $s_i = (1, \ldots, i-1, i+1, i, i+2, \ldots, n)$ is the transposition of i and $i+1$. The group D_n is generated by $\{\tilde{s}_0, s_1, \ldots, s_{n-1}\}$ with $\tilde{s}_0 = (-2, -1, 3, \ldots, n)$, thus it is sufficient to add only $|\text{Inv}^-(\pi)|$ to $|\text{Inv}^+(\pi)|$. For more details and formal proofs, see [3, Sections 8.1 and 8.2].

The knowledge of generating systems also allows to derive simple representations of X_{des} on B_n and D_n. Expand the in-line notation in (14) by setting $\pi(0) := 0$. Then, according to [3, Sections 8.1 and 8.2],

$$X^B_{\text{des}}(\pi) = \sum_{i=0}^{n-1} \mathbf{1}\{\pi(i) > \pi(i+1)\}, \tag{16a}$$

and on the even-signed permutation group D_n,

$$X^D_{\text{des}}(\pi) = \mathbf{1}\{-\pi(2) > \pi(1)\} + \sum_{i=1}^{n-1} \mathbf{1}\{\pi(i) > \pi(i+1)\}. \tag{16b}$$

To draw elements uniformly from B_n, one can first draw some uniform $\pi \in S_n$ and then multiply each $\pi(i)$ with a Rademacher variable independent of everything else. Instead of the Rademacher variables, we propose a more general approach by drawing signs with a fixed p-bias, i.e., each sign is -1 with probability $p \in [0,1]$ and $+1$ with probability $q := 1-p$. This yields a family of probability measures on B_n, where the case $p = 1/2$ corresponds to the uniform distribution on B_n, while if $p = 0$, all mass is on the symmetric group $S_n \subset B_n$.

A corresponding probability distribution on the even-signed permutation group D_n is obtained by first choosing the unsigned permutation $|\pi| \in S_n$ uniformly and then assigning $n-1$ signs for the entries $\pi(1), \dots, \pi(n-1)$ with p-bias, and finally specifying the sign of $\pi(n)$ such that $\pi(1) \cdots \pi(n-1)\pi(n) > 0$.

Definition 5.1. Let $p \in [0,1]$ and $q := 1-p$. Then, the *p-biased signed permutations measure* on the group B_n is the probability measure induced by the point masses

$$\mathbb{P}(\{\pi\}) = \frac{1}{n!} p^{\text{neg}(\pi)} q^{n-\text{neg}(\pi)}, \qquad \pi \in B_n\,,$$

where $\text{neg}(\pi)$ denotes the number of negative signs in π and we use the convention $0^0 := 1$. The p-biased signed permutations measure on D_n is derived as described above.

Therefore, the entries of π, with π distributed according to the p-biased signed permutations measure, can be represented by i.i.d. random variables $Z_1, \dots, Z_n$ with $\forall i = 1, \dots, n$: $Z_i = U_i R_i$, where R_i is a ± 1-valued random variable with

$$\mathbb{P}(R_i = -1) = p \quad \text{ and } \quad \mathbb{P}(R_i = 1) = q,$$

and $U_i \sim U(0,1)$ is independent of R_i. The probability distribution function of Z_1 is

$$F_p(z) := \mathbb{P}(Z_1 \le z) = \begin{cases} pz + p, & z \in [-1, 0], \\ qz + p, & z \in [0, 1], \end{cases}$$

and we simply write $Z_1 \sim \text{GR}(p)$ (generalized Rademacher with parameter p). Note the special cases $\text{GR}(0) = U(0,1)$, $\text{GR}(1/2) = U(-1,1)$ and $\text{GR}(1) = U(-1,0)$. Figure 2 illustrates F_p in the cases of $p = 0, p = 1/4$ and $p = 3/4$. Accordingly, the Lebesgue density of $\text{GR}(p)$ is $f_p(z) = p\mathbf{1}\{-1 < z < 0\} + q\mathbf{1}\{0 < z < 1\}$.

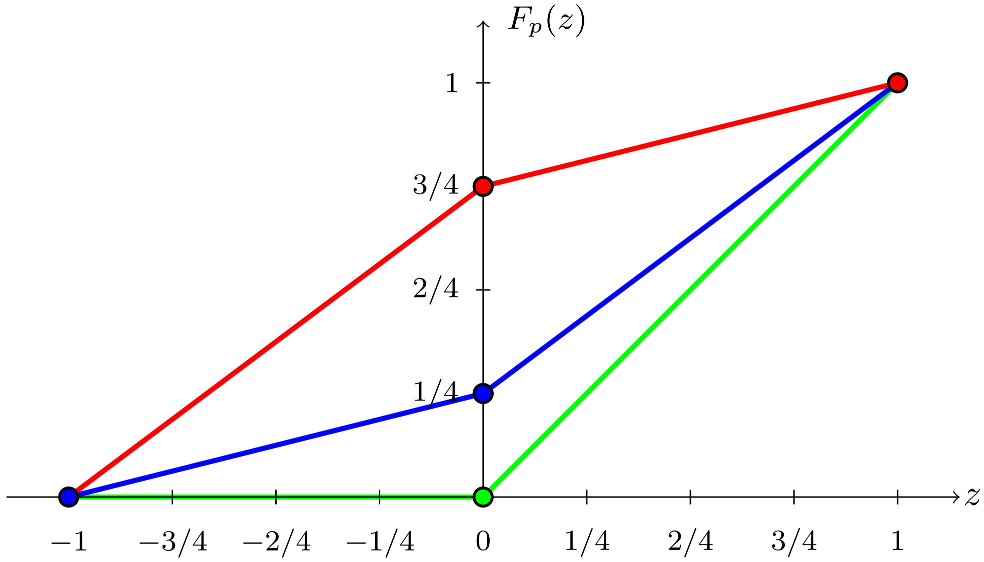

FIGURE 2. Probability distribution functions of $\mathrm{GR}(p)$ for $p = 0$ (green), $p = 1/4$ (blue) and $p = 3/4$ (red).

Remark 5.2. Let X_{inv}^B denote the random number of inversions on B_n and let X_{inv}^D denote that on D_n. According to (15), we can write

$$X_{\mathrm{inv}}^B = \sum_{i<j} \mathbf{1}\{Z_i > Z_j\} + \sum_{i<j} \mathbf{1}\{-Z_i > Z_j\} + \sum_{i=1}^{n} \mathbf{1}\{Z_i < 0\}, \tag{17a}$$

$$X_{\mathrm{inv}}^D = \sum_{i<j} \mathbf{1}\{Z_i > Z_j\} + \sum_{i<j} \mathbf{1}\{-Z_i > Z_j\}. \tag{17b}$$

Furthermore, (16a) and (16b) translate to

$$X_{\mathrm{des}}^B = \sum_{i=1}^{n-1} \mathbf{1}\{Z_i > Z_{i+1}\} + \mathbf{1}\{Z_1 < 0\}, \tag{18a}$$

$$X_{\mathrm{des}}^D = \sum_{i=1}^{n-1} \mathbf{1}\{Z_i > Z_{i+1}\} + \mathbf{1}\{-Z_2 > Z_1\}. \tag{18b}$$

In what follows, we use X_{inv} and X_{des} as an umbrella notation for the numbers of inversions and descents on each of the groups S_n, B_n and D_n. Again, both (18a) and (18b) give a 1-dependent representation of X_{des}, which yields a 1-dependent representation of $(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})^\top$.

In the uniform case ($p = 1/2$), the means and variances are known by [18, Corollaries 3.2 and 4.2]. We now calculate them within the general p-bias framework. To this end, we first observe that for any $i < j$,

$$\mathbb{P}(-Z_i > Z_j) = \mathbb{E}(\mathbb{P}(-Z_i > Z_j | Z_i)) = \mathbb{E}(F_p(-Z_i)) = \mathbb{E}(\mathbb{E}(F_p(-Z_i)|U_i))$$

$$= \mathbb{E}(pF_p(U_i) + qF_p(-U_i)) = \mathbb{E}(p(qU_i + p) + q(-pU_i + p)) = p\,.$$

Then, it follows straightforwardly from (17a) and (17b) that

$$\mathbb{E}\left(X_{\mathrm{inv}}^{B}\right) = \binom{n}{2}\left(p + \frac{1}{2}\right) + np, \qquad \mathbb{E}\left(X_{\mathrm{inv}}^{D}\right) = \binom{n}{2}\left(p + \frac{1}{2}\right).$$

The formula for $\mathrm{Var}(X_{\mathrm{inv}})$ is given in the next lemma.

Lemma 5.3 (see Subsection 8.3 for the proof)**.** *On the p-biased (even-)signed permutation groups, we have*

$$\begin{aligned}\mathrm{Var}\left(X_{\mathrm{inv}}^{B}\right) &= \left(-\frac{1}{3}p^2 + \frac{1}{3}p + \frac{1}{36}\right)n^3 - \left(3p^3 - 4p^2 + p - \frac{1}{24}\right)n^2 \\ &\quad + \left(3p^3 - \frac{14}{3}p^2 + \frac{5}{3}p - \frac{5}{72}\right)n, \\ \mathrm{Var}\left(X_{\mathrm{inv}}^{D}\right) &= \left(-\frac{1}{3}p^2 + \frac{1}{3}p + \frac{1}{36}\right)n^3 - \left(p^3 - 2p^2 + p - \frac{1}{24}\right)n^2 \\ &\quad + \left(p^3 - \frac{5}{3}p^2 + \frac{2}{3}p - \frac{5}{72}\right)n.\end{aligned}$$

In particular, if $p = 0$ or $p = 1/2$, we obtain the results in [18, Corollaries 3.2 and 4.2]*.*

For the variance of $\hat{X}_{\mathrm{inv}}$ on B_n and D_n, we get the same leading term as in Lemma 5.3, regardless of p. Hence, we obtain the Hájek approximation statement from Lemma 2.3 on the groups B_n and D_n with p-bias.

Lemma 5.4 (see Subsection 8.4 for the proof)**.** *On the p-biased (even-)signed permutation groups, we also have*

$$\mathrm{Var}(\hat{X}_{\mathrm{inv}}) = \left(-\frac{1}{3}p^2 + \frac{1}{3}p + \frac{1}{36}\right)n^3 + O(n^2),$$

so $\hat{Y}_{\mathrm{inv}} = Y_{\mathrm{inv}} + o_{\mathbb{P}}(1)$ applies according to Theorem 2.2.

The leading term, as a function of p, has no zeros in $[0, 1]$ and assumes its global maximum at $p = 1/2$, which is the unbiased case. This means that the order of $\mathrm{Var}(X_{\mathrm{inv}})$ and $\mathrm{Var}(\hat{X}_{\mathrm{inv}})$ is guaranteed to be cubic in n.

From Lemma 5.4, we obtain an extension of Corollary 2.5, which we present as a general statement on all three families of classical Weyl groups.

Corollary 5.5. *Let W be a classical Weyl group of rank n, that is, $W \in \{S_n, B_n, D_n\}$. Set $Z_0 := -\infty$ if $W = S_n$, $Z_0 := 0$ if $W = B_n$ and $Z_0 := -Z_2$ if $W = D_n$. Then,*

$$\begin{pmatrix}\hat{X}_{\mathrm{inv}} \\ X_{\mathrm{des}}\end{pmatrix} = \begin{pmatrix}\mathbb{E}(X_{\mathrm{inv}} \mid Z_1) \\ \mathbf{1}\{Z_0 > Z_1\}\end{pmatrix} + \ldots + \begin{pmatrix}\mathbb{E}(X_{\mathrm{inv}} \mid Z_{n-1}) \\ \mathbf{1}\{Z_{n-2} > Z_{n-1}\}\end{pmatrix} + \begin{pmatrix}\mathbb{E}(X_{\mathrm{inv}} \mid Z_n) - (n-1)\mathbb{E}(X_{\mathrm{inv}}) \\ \mathbf{1}\{Z_{n-1} > Z_n\}\end{pmatrix}$$

is a 1-dependent decomposition of $(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}})^{\top}$. On B_n and D_n, this applies with any sign bias.

Lemma 5.6 (see Subsections 8.5, 8.6 for the proof)**.** *On both of the groups B_n and D_n with p-bias, it holds that*

a) $\mathrm{Corr}(X_{\mathrm{inv}}, X_{\mathrm{des}}) \longrightarrow 0$ *as* $n \to \infty$*, and*

$$\mathrm{Cov}(X^B_{\mathrm{inv}}, X^B_{\mathrm{des}}) = (n-1)\left(\frac{p^2}{2} + p^2 q - \frac{p}{2} + \frac{1}{4}\right) + (p - p^2),$$

$$\mathrm{Cov}(X^D_{\mathrm{inv}}, X^D_{\mathrm{des}}) = (n-1)\left(\frac{p^2}{2} + p^2 q - \frac{p}{2} + \frac{1}{4}\right) + p^2.$$

b) $\mathrm{Cov}(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}}) = \Theta(1/n)$*, so again,* $\mathrm{Corr}(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}}) \longrightarrow 0$ *as* $n \to \infty$*.*

The CLT of $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ on B_n and D_n is now derived analogously to Theorem 3.4. Likewise, all arguments in the proof of the extreme value limit Theorem 4.1 apply on B_n and D_n.

Theorem 5.7. *For the joint statistic* $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ *on signed or even-signed permutation groups with p-bias, the following hold.*

a) $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top_{n\geq 1}$ *satisfies the CLT.*

b) The statement of Theorem 4.1 holds if W_n *is an arbitrary sequence of classical Weyl groups with* $\mathrm{rk}(W_n) = n$ *for all* $n \in \mathbb{N}$ *and with* k_n *chosen so that* $k_n \log(k_n) = o(n)$*.*

6. Products of classical Weyl groups

We now consider direct products of classical Weyl groups. Let $W = \prod_{i=1}^l W_i$ be such a product, where each W_i is one of S_n, B_n or D_n, and l is a fixed positive integer. By [18, Lemma 2.2], we know that

$$X^W_{\mathrm{inv}} = \sum_{i=1}^{l} X^{W_i}_{\mathrm{inv}}$$

is a sum of independent random variables, implying $\mathrm{Var}(X^W_{\mathrm{inv}}) = \sum_{i=1}^l \mathrm{Var}(X^{W_i}_{\mathrm{inv}})$. Let $X^{W_i}_{\mathrm{inv}}$ be constructed from variables $Z^{(i)}_1, \ldots, Z^{(i)}_{n_i}$, where n_i denotes the number of letters on which the group W_i acts, and each $Z^{(i)}_j$ is $\mathrm{GR}(p_i)$ for some $p_i \in [0,1]$, and the entire collection of all $Z^{(i)}_j$ is independent. Setting $n := n_1 + \ldots + n_l$, the overall Hájek projection $\hat{X}^W_{\mathrm{inv}}$ of X^W_{inv} is

$$\hat{X}^W_{\mathrm{inv}} = \sum_{i=1}^{l}\sum_{j=1}^{n_i} \mathbb{E}\left(X^W_{\mathrm{inv}} \mid Z^{(i)}_j\right) - (n-1)\mathbb{E}(X^W_{\mathrm{inv}}),$$

where $\mathbb{E}\left(X^W_{\mathrm{inv}} \mid Z^{(i)}_j\right) = \sum_{k=1}^l \mathbb{E}\left(X^{W_k}_{\mathrm{inv}} \mid Z^{(i)}_j\right)$. If $k \neq i$, then $X^{W_k}_{\mathrm{inv}}$ is independent of $Z^{(i)}_j$, which means that in this case $\mathbb{E}\left(X^{W_k}_{\mathrm{inv}} \mid Z^{(i)}_j\right) = \mathbb{E}(X^{W_k}_{\mathrm{inv}})$ is constant. We therefore obtain

$$\mathrm{Var}(\hat{X}^W_{\mathrm{inv}}) = \sum_{i=1}^{l}\sum_{j=1}^{n_i} \mathrm{Var}\left(\mathbb{E}\left(X^{W_i}_{\mathrm{inv}} \mid Z^{(i)}_j\right)\right) = \sum_{i=1}^{l} \mathrm{Var}(\hat{X}^{W_i}_{\mathrm{inv}}).$$

For any W_i, we have $\mathrm{Var}(X^{W_i}_{\mathrm{inv}}) \sim \mathrm{Var}(\hat{X}^{W_i}_{\mathrm{inv}})$. Furthermore, all variances are cubic as seen in Lemmas 2.3, 5.3, 5.4 and [18, Corollary 3.2], i.e., we have

$$\mathrm{Var}(X^{W_i}_{\mathrm{inv}}) = c_i n_i^3 + O(n_i^2) = \mathrm{Var}(\hat{X}^{W_i}_{\mathrm{inv}}), \qquad n_i \to \infty,$$

where $c_i := -\frac{1}{3}p_i^2 + \frac{1}{3}p_i + \frac{1}{36}$. It is seen from the calculations in the proofs of Lemmas 2.3 and 5.4 that $\mathrm{Var}(X_{\mathrm{inv}}) - \mathrm{Var}(\hat{X}_{\mathrm{inv}}) = \Theta(n^2)$. So we can write

$$\mathrm{Var}(X_{\mathrm{inv}}^W) = \sum_{i=1}^{l} \Big(c_i n_i^3 + \alpha_i n_i^2 + O(n_i)\Big), \qquad \mathrm{Var}(\hat{X}_{\mathrm{inv}}^W) = \sum_{i=1}^{l} \Big(c_i n_i^3 + \beta_i n_i^2 + O(n_i)\Big), \tag{19}$$

with $\alpha_i \neq \beta_i$ for all i.

Consider a sequence $(W_n)_{n\in\mathbb{N}}$ of products as introduced above, assuming that the number l of components remains bounded. Then, we see that $\mathrm{Var}(X_{\mathrm{inv}}^W) \sim \mathrm{Var}(\hat{X}_{\mathrm{inv}}^W)$ holds as $n \to \infty$, since the cubic terms are equal and cannot be dominated by the quadratic terms. Thus, to obtain the CLT, the Hájek projection approximation is sufficient and the above considerations show the following extension of Theorem 3.4.

Corollary 6.1. *On bounded products of classical Weyl groups, it holds that* $Y_{\mathrm{inv}} = \hat{Y}_{\mathrm{inv}} + o_{\mathbb{P}}(1)$, *and* $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ *satisfies the CLT.*

To repeat the proof of the extreme value limit Theorem 4.1, there is another issue to consider, namely the bounds (12) and (13), which require a suitable control of

$$1 - \frac{\mathrm{Var}(\hat{X}_{\mathrm{inv}})}{\mathrm{Var}(X_{\mathrm{inv}})}. \tag{20}$$

Since the number of components of W_n is bounded, we can assume w.l.o.g. that the components are sorted decreasingly by rank, meaning n_1 is the largest rank with $n_1 = \Theta(n)$. We can also assume that each group has exactly l components (as groups with fewer components can be filled with components of S_1, not giving any further inversions and descents).

Theorem 6.2. *For fixed* $l \in \mathbb{N}$, *let* $W_n = \prod_{i=1}^l W_{n,i}$ *be products of finite Coxeter groups with* $\mathrm{rk}(W_n) = n \ \forall n \in \mathbb{N}$. *Let* $k_n \log(k_n) = o(n)$ *and let* $(X_{\mathrm{inv}}^{(j)}, X_{\mathrm{des}}^{(j)})^\top$, $j = 1, \ldots, k_n$ *be independent copies of* $(X_{\mathrm{inv}}, X_{\mathrm{des}})^\top$ *on* W_n. *Let* $M_{n,\mathrm{inv}}, M_{n,\mathrm{des}}$ *be defined as in* (8). *Then, the statement of Theorem* 4.1 *applies for* $(W_n)_{n\in\mathbb{N}}$, *that is, we obtain* (9) *again.*

Proof. The proof of Theorem 4.1 carries over almost seamlessly, we only need to check the bound of (20). We can rephrase (19) as

$$\mathrm{Var}(\hat{X}_{\mathrm{inv}}^{W_n}) = \sum_{i=1}^{l} c_i n_i^3 + \alpha n^2 + O(n), \qquad \mathrm{Var}(X_{\mathrm{inv}}^{W_n}) = \sum_{i=1}^{l} c_i n_i^3 + \beta n^2 + O(n).$$

Then,

$$1 - \frac{\mathrm{Var}(\hat{X}_{\mathrm{inv}}^{W_n})}{\mathrm{Var}(X_{\mathrm{inv}}^{W_n})} = 1 - \frac{\sum_{i=1}^l c_i n_i^3 + \alpha n^2 + O(n)}{\sum_{i=1}^l c_i n_i^3 + \beta n^2 + O(n)}.$$

Depending on whether the residual $\beta n^2 + O(n)$ is positive or negative, we can bound this in both directions (assuming it is positive) via

$$(20) \geq 1 - \frac{\sum_{i=1}^l c_i n_i^3 + \alpha n^2 + O(n)}{\sum_{i=1}^l c_i n_i^3} = \frac{\alpha n^2 + O(n)}{\sum_{i=1}^l c_i n_i^3} = O\left(\frac{1}{n}\right),$$

$$(20) = \frac{(\beta - \alpha) n^2 + O(n)}{\sum_{i=1}^l c_i n_i^3 + \beta n^2 + O(n)} \leq \frac{(\beta - \alpha) n^2 + O(n)}{\sum_{i=1}^l c_i n_i^3} = O\left(\frac{1}{n}\right).$$

Therefore, we have the same bound for (20) as in the proof of Theorem 4.1. □

7. Conclusion and Outlook

In this work, we proved both a CLT and an EVLT for the joint statistic $(X_{\text{inv}}, X_{\text{des}})$ on all classical Weyl groups, as well as their bounded products. This addresses one of the open questions raised in [13] and gives a significant extension to [15], which only covered the CLT on symmetric groups. We benefited from the fact that the number of inversions X_{inv} can be suitably approximated by its Hájek projection, enabling us to apply Gaussian approximation theory for m-dependent vectors. In comparison with the univariate results in [13], the triangular array could not be stretched as generously, because the involvement of Hájek's projection required stronger assumptions.

On the symmetric groups S_n, both common inversions and descents are special instances of so-called *generalized inversions* or *d-inversions*. For any choice of $d \in \{1, \dots, n-1\}$, the statistic of d-inversions counts only inversions over pairs (i, j) with $1 \le j - i \le d$. It is interesting to choose $d = d_n$ as a function of n. The asymptotic normality of d-inversions under suitable conditions for d_n has been shown in [21]. Accordingly, it is worthwhile to investigate the constraints on d_n that guarantee the EVLT in the way of Theorem 4.1.

Some further open questions remain as well, especially about the *two-sided Eulerian statistic* $X_T(\pi) := X_{\text{des}}(\pi) + X_{\text{des}}(\pi^{-1})$. This statistic is known to satisfy a CLT according to [4] by the method of dependency graphs, but its extreme value behavior still remains unknown. Likewise, the extreme asymptotics of the joint distribution $\big(X_{\text{des}}(\pi), X_{\text{des}}(\pi^{-1})\big)$ give another open question.

Acknowledgements

PD's research was supported by the German Research Foundation through GRK 2297 "MathCoRe". This work was initiated during his research visit at the Department of Mathematics at Ruhr University Bochum. The authors would like to thank Martin Raič for pointing out the reference [15] in an email conversation. We are grateful to Christian Stump for valuable comments during the research process.

8. Appendix: remaining proofs

8.1. **Proof of Lemma 2.6a).** We compute $\text{Cov}(X_{\text{inv}}, X_{\text{des}})$ from (1) and (2), obtaining that it has lower order than $\sqrt{\text{Var}(X_{\text{inv}})\text{Var}(X_{\text{des}})}$. According to (1) and (2), we have

$$\begin{aligned}\text{Cov}(X_{\text{inv}}, X_{\text{des}}) &= \sum_{1 \le i < j \le n} \sum_{k=1}^{n-1} \text{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}) \\ &= \sum_{1 \le i < j \le n} \sum_{\substack{k \in \{i-1, i, j-1, j\} \\ 1 \le k \le n-1}} \text{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}),\end{aligned}$$

where we used that if $k \notin \{i-1, i, j-1, j\}$, then the events $\{Z_i > Z_j\}$ and $\{Z_k > Z_{k+1}\}$ are independent and therefore $\text{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}) = 0$. In what follows, we analyze the case $k \in \{i-1, i, j-1, j\}$, first assuming that all these numbers are distinct. Additionally,

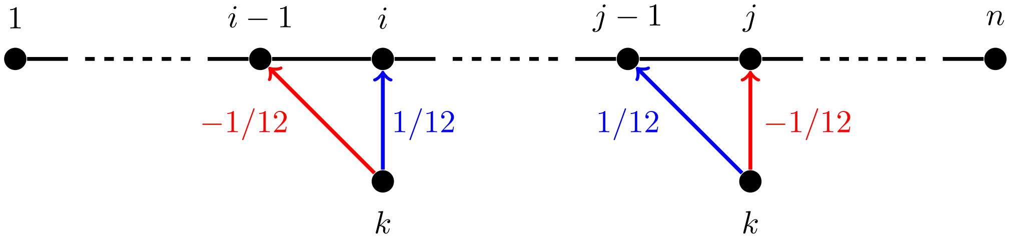

FIGURE 3. Canceling pairs of positive and negative covariances as k passes through $1, \dots, n-1$ in non-exceptional situations. The covariance is zero for all other values of k.

we temporarily ignore the border cases $i = 1$ (where $k = i - 1$ is outside the range) or $j = n$ (where $k = n$ is outside the range). This gives four possible constellations:

- type A: $k + 1 = i$ and $j > k$,
- type B: $k = i$ and $j > k + 1$,
- type C: $k + 1 = j$ and $i < k$,
- type D: $k = j$ and $i < k$.

For type A, we have

$$\begin{aligned}\mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_{i-1} > Z_i\}) &= \mathbb{P}(Z_i > Z_j, Z_{i-1} > Z_i) - \mathbb{P}(Z_i > Z_j)\mathbb{P}(Z_{i-1} > Z_i)\\ &= \mathbb{P}(Z_{i-1} > Z_i > Z_j) - \tfrac{1}{4}\\ &= \tfrac{1}{6} - \tfrac{1}{4} = -\tfrac{1}{12},\end{aligned}$$

since each of the six possible orderings of Z_{i-1}, Z_i, Z_j is equally likely as they are independent $U(0, 1)$ variables. For type B, we get

$$\begin{aligned}\mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_i > Z_{i+1}\}) &= \mathbb{P}(Z_i = \max\{Z_i, Z_{i+1}, Z_j\}) - \tfrac{1}{4}\\ &= \tfrac{1}{3} - \tfrac{1}{4} = \tfrac{1}{12}\end{aligned}$$

since each of Z_i, Z_{i+1}, Z_j are equally likely to be the maximum. Types C and D are handled the same way. For type C, we have $\mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_{j-1} > Z_j\}) = \frac{1}{12}$, and for type D, $\mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_j > Z_{j+1}\}) = -\frac{1}{12}$. So if $1 < i < i + 1 < j < n$, the inner sum

$$\sum_{k \in \{i-1, i, j-1, j\}} \mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\})$$

consists of two canceling pairs of $1/12$ and $-1/12$, and vanishes altogether. Figure 3 displays the passage of k and the positions of the positive and negative covariances, in which we see that the positive signs are located inside, while the negative ones are located outside. We also use this figure to explain what happens if $1 < i < i + 1 < j < n$ does not hold. We call such cases exceptional.

- If i and j are subsequent, i.e., $j = i + 1$, then the two locations in Figure 3 with positive contribution collide. If additionally $k = i$, we obtain $\mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}) = \mathrm{Cov}(\mathbf{1}\{Z_i > Z_{i+1}\}, \mathbf{1}\{Z_i > Z_{i+1}\}) = \mathrm{Var}(\mathbf{1}\{Z_i > Z_{i+1}\}) = 1/4$.

- If $i = 1$, then the leftmost negative term in Figure 3 disappears.
- If $j = n$, then the rightmost negative term in Figure 3 disappears.

As these situations are not mutually disjoint, we categorize the exceptional cases as follows:

- (E1): $j = i + 1$, but neither $i = 1$ nor $j = n$.
- (E2): $i = 1$ and $j = 3, \dots, n - 1$.
- (E3): $j = n$ and $i = 2, \dots, n - 2$.
- (E4): $i = 1$, $j = 2$.
- (E5): $i = n - 1$, $j = n$.
- (E6): $i = 1$, $j = n$.

As an example, we display situation (E1) in Figure 4.

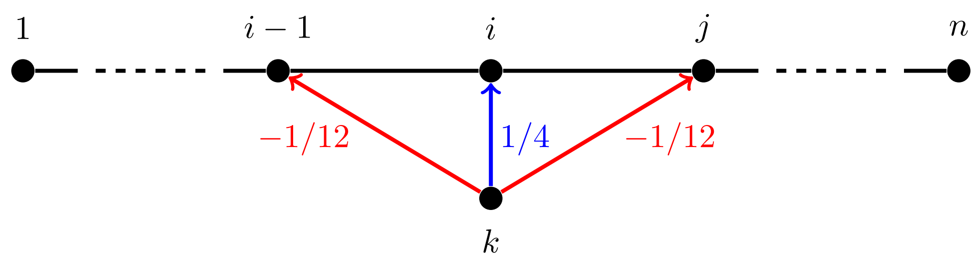

FIGURE 4. Display of covariances for the exceptional case (E1) when i and $j = i + 1$ are subsequent.

Looking at the contributions and frequencies of (E1), ..., (E6), we obtain the exact result

$$\operatorname{Cov}(X_{\text{inv}}, X_{\text{des}}) = \underbrace{(n-3)\left(\frac{1}{4} - \frac{1}{6}\right)}_{\text{(E1)}} + \underbrace{2(n-3)\frac{1}{12}}_{\text{(E2, E3)}} + \underbrace{2\left(\frac{1}{4} - \frac{1}{12}\right)}_{\text{(E4, E5)}} + \underbrace{\frac{1}{6}}_{\text{(E6)}} = \frac{n-1}{4}.$$

The claim follows. □

8.2. **Proof of Lemma 2.6b).** Recall that $Z_1, \dots, Z_n$ are i.i.d. $U(0, 1)$. By Lemma 2.3, we have

$$\hat{X}_{\text{inv}} = \sum_{j=1}^{n} (n - 2j + 1) Z_j + \frac{1}{2}\binom{n}{2},$$

which, together with the definition of X_{des}, yields

$$\begin{aligned} \operatorname{Cov}(\hat{X}_{\text{inv}}, X_{\text{des}}) &= \operatorname{Cov}\left(\sum_{j=1}^{n} (n-2j+1) Z_j, \sum_{k=1}^{n-1} \mathbf{1}\{Z_k > Z_{k+1}\}\right) \\ &= \Big(\sum_{j=2}^{n-1} + \sum_{j \in \{1,n\}}\Big) \sum_{k=1}^{n-1} (n-2j+1) \operatorname{Cov}(Z_j, \mathbf{1}\{Z_k > Z_{k+1}\}) \\ &=: T_1 + T_2 \,. \end{aligned}$$

In view of the independence of $Z_1, \ldots, Z_n$, we get

$$T_1 = \sum_{j=2}^{n-1} (n-2j+1)\big(\mathrm{Cov}(Z_j, \mathbf{1}\{Z_j < Z_{j-1}\}) + \mathrm{Cov}(Z_j, \mathbf{1}\{Z_j > Z_{j+1}\})\big) = 0\,,$$

where for the last equality, we used

$$\begin{aligned}
&\mathrm{Cov}(Z_j, \mathbf{1}\{Z_j < Z_{j-1}\}) + \mathrm{Cov}(Z_j, \mathbf{1}\{Z_j > Z_{j+1}\}) \\
&= \mathbb{E}(Z_j \mathbf{1}\{Z_j < Z_{j-1}\}) + \mathbb{E}(Z_j \mathbf{1}\{Z_j > Z_{j+1}\}) - \frac{1}{2} \\
&= \mathbb{E}(Z_j \mathbf{1}\{Z_j < Z_{j-1}\}) + \mathbb{E}(Z_j \mathbf{1}\{Z_j > Z_{j-1}\}) - \frac{1}{2} \\
&= \mathbb{E}(Z_j) - \frac{1}{2} = 0\,.
\end{aligned}$$

As $Z_1 \mathbf{1}\{Z_1 < Z_2\}$ is a function in two uniform variables with joint density $f\colon \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto \mathbf{1}\{(x,y) \in [0,1]^2\}$, we can apply Fubini's Theorem to obtain

$$\begin{aligned}
\mathbb{E}(Z_1 \mathbf{1}\{Z_1 < Z_2\}) &= \int_{[0,1]^2} x \mathbf{1}\{x < y\}\, \mathrm{d}(x,y) \\
&= \int_0^1 x \left(\int_0^1 \mathbf{1}\{x < y\} \mathrm{d}y \right) \mathrm{d}x \\
&= \int_0^1 x(1-x)\, \mathrm{d}x = \frac{1}{6}.
\end{aligned}$$

Therefore, we get for T_2:

$$\begin{aligned}
T_2 &= (n-1)\mathrm{Cov}(Z_1, \mathbf{1}\{Z_1 > Z_2\}) - (n-1)\mathrm{Cov}(Z_n, \mathbf{1}\{Z_{n-1} > Z_n\}) \\
&= (n-1)\big(\mathbb{E}(Z_1 \mathbf{1}\{Z_1 > Z_2\}) - \mathbb{E}(Z_1 \mathbf{1}\{Z_1 < Z_2\})\big) \\
&= (n-1)\big(1/2 - 2\mathbb{E}(Z_1 \mathbf{1}\{Z_1 < Z_2\})\big) = \frac{n-1}{6}\,,
\end{aligned}$$

which shows that $\mathrm{Cov}(\hat{X}_{\mathrm{inv}}, X_{\mathrm{des}}) = \dfrac{n-1}{6}$. □

8.3. **Proof of Lemma 5.3.** We follow the instructive calculation of $\mathrm{Var}(X_{\mathrm{inv}})$ in the uniform case provided in [18, Section 3]. So, the main task is to calculate $\mathbb{E}(X_{\mathrm{inv}}^2)$. For X_{inv}^B, we recall (17a) and use the abbreviations

$$X_{\mathrm{inv}}^B = \underbrace{\sum_{i<j} \mathbf{1}\{Z_i > Z_j\}}_{=:\, X^+} + \underbrace{\sum_{i<j} \mathbf{1}\{-Z_i > Z_j\}}_{=:\, X^-} + \underbrace{\sum_{i=1}^{n} \mathbf{1}\{Z_i < 0\}}_{=:\, X^\circ} = X^+ + X^- + X^\circ.$$

This means we need to compute the terms

$$\begin{aligned}
\mathbb{E}\big((X_{\mathrm{inv}}^B)^2\big) &= \mathbb{E}((X^+)^2) + \mathbb{E}((X^-)^2) + 2\mathbb{E}(X^+ X^-) \\
&\quad + \mathbb{E}((X^\circ)^2) + 2\mathbb{E}(X^+ X^\circ) + 2\mathbb{E}(X^- X^\circ).
\end{aligned}$$

The first term $\mathbb{E}((X^+)^2)$ is invariant under p, since it only involves events $\{Z_i > Z_j\}$ for which $\mathbb{P}(Z_i > Z_j) = 1/2$, even if the involved Z_i, Z_j are not uniformly distributed. Therefore, we can obtain $\mathbb{E}((X^+)^2)$ from [18, Section 3]:

$$\mathbb{E}((X^+)^2) = \frac{1}{2}\binom{n}{2} + \frac{1}{4}\binom{n}{2}\binom{n-2}{2} + \frac{5}{3}\binom{n}{3}.$$

Next, we turn to

$$\mathbb{E}((X^-)^2) = \sum_{i<j}\sum_{k<l}\mathbb{P}(-Z_i > Z_j, -Z_k > Z_l).$$

For the $\binom{n}{2}\binom{n-2}{2}$ choices of pairwise distinct i, j, k, l, we have that $\mathbb{P}(-Z_i > Z_j, -Z_k > Z_l) = p^2$ by independence, and for the $\binom{n}{2}$ cases where $(i, j) = (k, l)$, we simply get $\mathbb{P}(-Z_i > Z_j) = p$. The set of triples with two of the indices colliding need to be analyzed similarly as in the proof of Lemma 2.6a). Note that the cases $i = k$ and $j = l$ each contain two contributions. E.g., in the case of $i = l$, we calculate by case distinction according to the signs:

$$\begin{aligned}
\mathbb{P}(-Z_i > Z_j, -Z_k > Z_i) &= \mathbb{P}(-Z_i > Z_j, -Z_k > Z_i, Z_i > 0) \\
&\quad + \mathbb{P}(-Z_i > Z_j, -Z_k > Z_i, Z_i < 0) \\
&= \mathbb{P}(-Z_i > Z_j, -Z_k > Z_i, Z_i > 0, Z_j < 0, Z_k < 0) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j < 0, Z_k < 0) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j > 0, Z_k < 0, -Z_i > Z_j) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j < 0, Z_k > 0, -Z_k > Z_i) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j > 0, Z_k > 0, -Z_i > Z_j, -Z_k > Z_i) \\
&= \frac{1}{3}p^2 q + p^3 + \frac{1}{2}p^2 q + \frac{1}{2}p^2 q + \frac{1}{3}pq^2 \\
&= \frac{1}{3}p(2p+1).
\end{aligned}$$

It turns out that all six constellations of triplets give this contribution. So, overall,

$$\mathbb{E}((X^-)^2) = \binom{n}{2}p + \binom{n}{2}\binom{n-2}{2}p^2 + \binom{n}{3}2p(2p+1).$$

For $\mathbb{E}(X^+X^-)$, the disjoint quadruplets give a contribution of $p/2$ each. For the colliding pairs, we need to compute

$$\begin{aligned}
&\mathbb{P}(Z_i > Z_j, -Z_i > Z_j) \\
&= \underbrace{\mathbb{P}(Z_i > Z_j, -Z_i > Z_j, Z_i > 0, Z_j > 0)}_{=\,0} + \underbrace{\mathbb{P}(Z_i > Z_j, -Z_i > Z_j, Z_i < 0, Z_j > 0)}_{=\,0} \\
&\qquad + \mathbb{P}(Z_i > Z_j, -Z_i > Z_j, Z_i < 0, Z_j < 0) + \mathbb{P}(Z_i > Z_j, -Z_i > Z_j, Z_i > 0, Z_j < 0) \\
&= \mathbb{P}(Z_i < 0, Z_j < 0, Z_i > Z_j) + \mathbb{P}(Z_i > 0, Z_j < 0, -Z_i > Z_j) = \frac{p^2}{2} + p^2 q.
\end{aligned}$$

For the triplets, we repeat the procedure above. For the cases $j = k$ and $j = l$, we get

$$\mathbb{P}(Z_i > Z_j, -Z_j > Z_k) = -\frac{1}{6}p(2p-5).$$

For the cases $i = l$ and $i = k$, we derive

$$\begin{aligned}
\mathbb{P}(Z_i > Z_j, -Z_i > Z_l) &= \mathbb{P}(Z_i > Z_j, -Z_i > Z_l, Z_i > 0, Z_l < 0) \\
&\quad + \mathbb{P}(Z_i > Z_j, -Z_i > Z_l, Z_i < 0, Z_j < 0) \\
&= \mathbb{P}(Z_i > 0, Z_j < 0, Z_l < 0, -Z_i > Z_l) \\
&\quad + \mathbb{P}(Z_i > 0, Z_j > 0, Z_l < 0, Z_j < Z_i < -Z_l) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j < 0, Z_l < 0, Z_i > Z_j) \\
&\quad + \mathbb{P}(Z_i < 0, Z_j < 0, Z_l > 0, Z_j < Z_i < -Z_l) \\
&= \frac{1}{2}p^2 q + \frac{1}{6}pq^2 + \frac{1}{2}p^3 + \frac{1}{6}p^2 q \\
&= \frac{1}{6}p(2p+1),
\end{aligned}$$

and overall we obtain

$$\begin{aligned}
\mathbb{E}(X^+X^-) &= \binom{n}{2}\left(\frac{p^2}{2} + p^2 q\right) + \binom{n}{2}\binom{n-2}{2}\frac{p}{2} + 3\binom{n}{3}\underbrace{\left(\frac{1}{6}p(2p+1) - \frac{1}{6}p(2p-5)\right)}_{=\,p} \\
&= \binom{n}{2}\left(\frac{p^2}{2} + p^2 q\right) + \binom{n}{2}\binom{n-2}{2}\frac{p}{2} + 3\binom{n}{3}p.
\end{aligned}$$

The remaining three terms are easily calculated as

$$\begin{aligned}
\mathbb{E}(X^+X^\circ) &= \sum_{i<j}\sum_{k=1}^{n} \mathbb{P}(Z_i > Z_j, Z_k < 0) \\
&= 3\binom{n}{3}\frac{p}{2} + \sum_{i<j}\mathbb{P}(Z_i > Z_j, Z_i < 0) + \sum_{i<j}\mathbb{P}(Z_i > Z_j, Z_j < 0) \\
&= 3\binom{n}{3}\frac{p}{2} + \binom{n}{2}(p^2 + pq) = 3\binom{n}{3}\frac{p}{2} + \binom{n}{2}p, \\
\mathbb{E}(X^-X^\circ) &= \sum_{i<j}\sum_{k=1}^{n} \mathbb{P}(-Z_i > Z_j, Z_k < 0) \\
&= 3\binom{n}{3}p^2 + 2\binom{n}{2}(p^2 + p^2 q) = 3\binom{n}{3}p^2 + 2\binom{n}{2}(2p^2 - p^3), \\
\mathbb{E}((X^\circ)^2) &= \sum_{i,j=1}^{n} \mathbb{P}(Z_i < 0, Z_j < 0) = 2\binom{n}{2}p^2 + np.
\end{aligned}$$

Summing all of these terms and subtracting the square of the mean gives the claim for B_n. On D_n, we ignore the parts involving X° and get the desired result. □

8.4. **Proof of Lemma 5.4.** We first prove the claim on the even-signed permutation groups D_n, since the calculation will be the same for B_n. We proceed as in Lemma 2.3, starting from

$$\hat{X}_{\text{inv}} = \sum_{k=1}^{n} \mathbb{E}(X_{\text{inv}} \mid Z_k) - (n-1)\mathbb{E}(X_{\text{inv}}),$$

except here, X_{inv} is defined by (17b), and we have $Z_k \sim \text{GR}(p)$. According to (17b), we have

$$\begin{aligned} \mathbb{E}(X_{\text{inv}} \mid Z_k) &= \sum_{i<j} \mathbb{E}\big(\mathbf{1}\{Z_i > Z_j\} + \mathbf{1}\{-Z_i > Z_j\} \mid Z_k\big) \\ &= \sum_{i<j} \big(\mathbb{P}(Z_i > Z_j \mid Z_k) + \mathbb{P}(-Z_i > Z_j \mid Z_k)\big). \end{aligned} \tag{21}$$

Write $f(Z_i, Z_j) := \mathbf{1}\{Z_i > Z_j\} + \mathbf{1}\{-Z_i > Z_j\}$ for $i < j$ and set $U_k = |Z_k|$. A straightforward case distinction gives

- $Z_i, Z_j > 0$: $f(Z_i, Z_j) = \mathbf{1}\{U_i > U_j\}$,
- $Z_i > 0, Z_j < 0$: $f(Z_i, Z_j) = \mathbf{1}\{U_i > U_j\}$,
- $Z_i < 0, Z_j > 0$: $f(Z_i, Z_j) = \mathbf{1}\{U_i < U_j\} + 1$,
- $Z_i, Z_j < 0$: $f(Z_i, Z_j) = \mathbf{1}\{U_i < U_j\} + 1$.

For symmetry reasons, $f(Z_i, Z_j)$ depends only on the sign of Z_i but not on the sign of Z_j. To compute (21), we only need to consider the $n-k$ tuples (k, j) and the $k-1$ tuples (i, k), as the other tuples are independent of Z_k and produce constants which do not contribute to the variance.

Recall that for $k < j$, we have $\mathbb{P}(U_k > U_j \mid U_k) = U_k$ and $\mathbb{P}(U_k < U_j \mid U_k) = 1 - U_k$. Therefore, we can write

$$\mathbb{E}(X_{\text{inv}} \mid Z_k) = \sum_{i=1}^{k-1} \mathbb{E}\big(f(Z_i, Z_k) \mid Z_k\big) + \sum_{j=k+1}^{n} \mathbb{E}\big(f(Z_k, Z_j) \mid Z_k\big),$$

where

$$\begin{aligned} \mathbb{E}\big(f(Z_i, Z_k) \mid Z_k\big) &= \mathbb{P}(Z_i > 0)(1 - U_k) + \mathbb{P}(Z_i < 0)(1 + U_k) \\ &= q(1 - U_k) + p(1 + U_k) = 2pU_k - U_k + 1 \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}\big(f(Z_k, Z_j) \mid Z_k\big) &= \mathbf{1}\{Z_k > 0\}U_k + \mathbf{1}\{Z_k < 0\}(1 + 1 - U_k) \\ &= \mathbf{1}\{Z_k > 0\}U_k + \mathbf{1}\{Z_k < 0\}(2 - U_k). \end{aligned}$$

Overall, we obtain

$$\begin{aligned} \mathbb{E}(X_{\text{inv}} \mid Z_k) = (k-1)&(2pU_k - U_k + 1) \\ &+ (n-k)\big(\mathbf{1}\{Z_k > 0\}U_k + \mathbf{1}\{Z_k < 0\}(2 - U_k)\big) + \text{const}. \end{aligned}$$

To use the standard formula $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, where X is not affected by constant summands, we now compute

$$\begin{aligned} &\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)^2) \\ &= (k-1)^2 \mathbb{E}\left((2pU_k - U_k + 1)^2\right) \end{aligned} \tag{22a}$$

$$+ (n-k)^2\mathbb{E}\left(\left(\mathbf{1}\{Z_k > 0\}U_k + \mathbf{1}\{Z_k < 0\}(2-U_k)\right)^2\right) \tag{22b}$$

$$+ 2(k-1)(n-k)\mathbb{E}\left((2pU_k - U_k + 1)\big(\mathbf{1}\{Z_k > 0\}U_k + \mathbf{1}\{Z_k < 0\}(2-U_k)\big)\right). \tag{22c}$$

The random variables $\mathbf{1}\{Z_k > 0\}$ and U_k are independent by construction and therefore,

$$\begin{aligned}
(22a) &= (k-1)^2\mathbb{E}\left(4p^2U_k^2 - 4pU_k^2 + 4pU_k + U_k^2 - 2U_k + 1\right) \\
&= (k-1)^2\left(\frac{4}{3}p^2 - \frac{4}{3}p + 2p + \frac{1}{3}\right) = (k-1)^2\left(\frac{4}{3}p^2 + \frac{2}{3}p + \frac{1}{3}\right), \\
(22b) &= (n-k)^2\left(q\mathbb{E}(U_k)^2 + p\mathbb{E}((2-U_k)^2)\right) \\
&= (n-k)^2\left(\frac{q}{3} + \frac{7}{3}p\right) = (n-k)^2\left(2p + \frac{1}{3}\right), \\
(22c) &= 2(k-1)(n-k)\Big((1-p)\mathbb{E}(2pU_k^2 - U_k^2 + U_k) \\
&\qquad + p\mathbb{E}((2pU_k - U_k + 1)(2-U_k))\Big) \\
&= 2(k-1)(n-k)\left((1-p)\left(\frac{2}{3}p + \frac{1}{6}\right) + p\left(\frac{4}{3}p + \frac{5}{6}\right)\right) \\
&= 2(k-1)(n-k)\left(\frac{2}{3}p^2 + \frac{4p}{3} + \frac{1}{6}\right).
\end{aligned}$$

In total, we have

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)^2) &= (k-1)^2\left(\frac{4}{3}p^2 + \frac{2}{3}p + \frac{1}{3}\right) + (n-k)^2\left(2p + \frac{1}{3}\right) \\
&\quad + 2(k-1)(n-k)\left(\frac{2}{3}p^2 + \frac{4}{3}p + \frac{1}{6}\right).
\end{aligned}$$

We subtract the square of

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)) &= (k-1)\left(p + \frac{1}{2}\right) + (n-k)\left(\frac{1}{2} - \frac{p}{2} + \frac{3}{2}p\right) \\
&= \left(p + \frac{1}{2}\right)(n-1).
\end{aligned}$$

The variance of $\hat{X}_{\text{inv}}$ on D_n is

$$\begin{aligned}
\operatorname{Var}(\hat{X}_{\text{inv}}) &= \sum_{k=1}^{n}\left(\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)^2) - \mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k))^2\right) \\
&= \sum_{k=1}^{n}\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)^2) - n(n-1)^2\left(p + \frac{1}{2}\right)^2,
\end{aligned}$$

so, to conclude the proof, we compute

$$\sum_{k=1}^{n}\mathbb{E}(\mathbb{E}(X_{\text{inv}} \mid Z_k)^2) = \left(\frac{4}{3}p^2 + \frac{2}{3}p + \frac{1}{3}\right)\sum_{k=1}^{n}(k-1)^2 + \left(2p + \frac{1}{3}\right)\sum_{k=1}^{n}(n-k)^2$$

$$
\begin{aligned}
&+ \left(\frac{4}{3}p^2 + \frac{8p}{3} + \frac{1}{3}\right) \sum_{k=1}^{n} (k-1)(n-k) \\
&= \left(\frac{4}{3}p^2 + \frac{2}{3}p + \frac{1}{3}\right) \cdot \frac{1}{6} n(n-1)(2n-1) \\
&+ \left(2p + \frac{1}{3}\right) \cdot \frac{1}{6} n(n-1)(2n-1) \\
&+ \left(\frac{4}{3}p^2 + \frac{8p}{3} + \frac{1}{3}\right) \cdot \frac{1}{6} n(n-1)(n-2) \\
&= n^3 \left(\frac{2}{3}p^2 + \frac{4}{3}p + \frac{5}{18}\right) - n^2 \left(\frac{4}{3}p^2 - \frac{8}{3}p - \frac{1}{2}\right) \\
&+ n \left(\frac{2}{3}p^2 + \frac{4}{3}p + \frac{2}{9}\right).
\end{aligned}
$$

Subtracting $n(n-1)^2\left(p+\frac{1}{2}\right)^2$ gives exactly the desired leading term computed in Lemma 5.3. On the groups B_n, we achieve the same result since the extra parts in $\mathrm{Var}(\hat{X}_{\mathrm{inv}})$ yielded by $\sum_{i=1}^{n} \mathbf{1}\{Z_i < 0\}$ are not significant. Recall that

$$
X_{\mathrm{inv}}^B = X_{\mathrm{inv}}^D + \sum_{i=1}^{n} \mathbf{1}\{Z_i < 0\},
$$

and therefore,

$$
\begin{aligned}
\mathrm{Var}\left(\mathbb{E}(X_{\mathrm{inv}}^B \mid Z_k)\right) &= \mathrm{Var}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k) + \sum_{j=1}^{n} \mathbb{E}(\mathbf{1}\{Z_j < 0\} \mid Z_k)\right) \\
&= \mathrm{Var}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k) + \mathbb{E}(\mathbf{1}\{Z_k < 0\} \mid Z_k) + \mathrm{const}\right) \\
&= \mathrm{Var}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k) + \mathbf{1}\{Z_k < 0\} + \mathrm{const}\right) \\
&= \mathrm{Var}\Big((n-k)\mathbf{1}\{Z_k > 0\}U_k + (n-k)\mathbf{1}\{Z_k < 0\}(2-U_k) \\
&\qquad + (k-1)(2pU_k - U_k + 1) + \mathbf{1}\{Z_k < 0\} + \mathrm{const}\Big).
\end{aligned}
$$

Using the standard formula again, we have

$$
\begin{aligned}
\mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^B \mid Z_k)^2\right) &= \mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k)^2\right) + \mathbb{E}\left(\mathbf{1}\{Z_k < 0\}\right) \\
&\quad + 2\mathbb{E}\left((n-k)(2-U_k)\mathbf{1}\{Z_k < 0\}\right) \\
&\quad + 2\mathbb{E}((k-1)(2pU_k - U_k + 1)\mathbf{1}\{Z_k < 0\}) \\
&= \mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k)^2\right) + p + 3p(n-k) + (k-1)(2p+1)p, \\
\mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^B \mid Z_k)\right)^2 &= \mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k)\right)^2 + p^2 + 2\mathbb{E}(\mathbf{1}\{Z_k < 0\})\mathbb{E}\left(\mathbb{E}(X_{\mathrm{inv}}^D \mid Z_k)\right) \\
&= \left(p + \frac{1}{2}\right)^2 (n-1)^2 + p^2 + 2p(n-1)\left(p + \frac{1}{2}\right).
\end{aligned}
$$

We conclude that

$$\begin{aligned}\mathrm{Var}(\hat{X}^B_{\mathrm{inv}}) &= \mathrm{Var}(\hat{X}^D_{\mathrm{inv}}) + n(p-p^2) - p(2p+1)\frac{n(n-1)}{2} \\ &= \mathrm{Var}(\hat{X}^D_{\mathrm{inv}}) + O(n^2).\end{aligned}$$

The desired claim follows from Theorem 2.2. □

8.5. **Proof of Lemma 5.6a).** For the groups B_n and D_n, the calculation follows the same approach as on S_n. Recall that now, $Z_1, \dots, Z_n \sim \mathrm{GR}(p)$. On D_n, we have by (16b), (17b) that

$$\mathrm{Cov}(X^D_{\mathrm{inv}}, X^D_{\mathrm{des}}) = \sum_{i<j}\sum_{k=1}^{n-1} \mathrm{Cov}\left(\mathbf{1}\{Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}\right) \tag{23a}$$

$$+ \sum_{i<j}\sum_{k=1}^{n-1} \mathrm{Cov}\left(\mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_k > Z_{k+1}\}\right) \tag{23b}$$

$$+ \sum_{i<j} \mathrm{Cov}(\mathbf{1}\{Z_i > Z_j\} + \mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\}). \tag{23c}$$

The contribution of (23a) is $(n-1)/4$, in analogy with Lemma 2.6a). In (23b), we first demonstrate the cancellation in the non-exceptional case when $\{i-1, i, j-1, j\}$ form a set of distinct numbers. In that case, we have

$$\begin{aligned}&\mathrm{Cov}(\mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_i > Z_{i+1}\}) + \mathrm{Cov}(\mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_{i-1} > Z_i\}) \\ &= \mathbb{E}(\mathbf{1}\{-Z_i > Z_j\}\mathbf{1}\{Z_i > Z_{i+1}\}) + \mathbb{E}(\mathbf{1}\{-Z_i > Z_j\}\mathbf{1}\{Z_{i-1} > Z_i\}) - 2(p/2) \\ &= \mathbb{E}(\mathbf{1}\{-Z_i > Z_j\}\mathbf{1}\{Z_i > Z_{i+1}\}) + \mathbb{E}(\mathbf{1}\{-Z_i > Z_j\}\mathbf{1}\{Z_{i+1} > Z_i\}) - p \\ &= \mathbb{E}(\mathbf{1}\{-Z_i > Z_j\}) - p = 0,\end{aligned}$$

and likewise,

$$\mathrm{Cov}(\mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_j > Z_{j+1}\}) + \mathrm{Cov}(\mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_{j-1} > Z_j\}) = 0.$$

However, this cancellation occurs not only in the non-exceptional cases, but also in the aggregation of the exceptional cases (E1) – (E6) from the proof of Lemma 2.6a), except for the covariances resulting from the clash of $j = i+1$ and $k = i$. To be precise, (E2) and (E3) cancel mutually. (E4) and (E5) give two clashes and a canceling pair. (E6) consists of another canceling pair. All of this can be checked from the computation of $\mathbb{E}(X^+X^-)$ in the proof of Lemma 5.3. From there, we also obtain $\forall i = 1, \dots, n-1$:

$$\mathrm{Cov}(\mathbf{1}\{-Z_i > Z_{i+1}\}, \mathbf{1}\{Z_i > Z_{i+1}\}) = \frac{p^2}{2} + p^2 q - \frac{p}{2},$$

which interestingly vanishes in the unbiased case. Overall,

$$(23b) = (n-1)\left(\frac{p^2}{2} + p^2 q - \frac{p}{2}\right).$$

Finally, consider (23c). Obviously, this double-indexed sum involves exactly the pairs (i, j) with $i = 1$ or $i = 2$. We get

$$\begin{aligned}
(23c) &= \sum_{j=3}^{n} \operatorname{Cov}(\mathbf{1}\{Z_1 > Z_j\} + \mathbf{1}\{-Z_1 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\}) \\
&\quad + \sum_{j=3}^{n} \Big[\operatorname{Cov}(\mathbf{1}\{Z_2 > Z_j\} + \mathbf{1}\{-Z_2 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\}) \\
&\quad + \underbrace{\operatorname{Cov}(\mathbf{1}\{-Z_1 > Z_2\}, \mathbf{1}\{-Z_2 > Z_1\})}_{=p^2} + \underbrace{\operatorname{Cov}(\mathbf{1}\{Z_1 > Z_2\}, \mathbf{1}\{-Z_2 > Z_1\})}_{=0}\Big] \\
&= \underbrace{\sum_{j=3}^{n} \Big[\operatorname{Cov}(\mathbf{1}\{Z_1 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\}) + \operatorname{Cov}(\mathbf{1}\{-Z_1 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\})\Big]}_{=0} \\
&\quad + \underbrace{\sum_{j=3}^{n} \Big[\operatorname{Cov}(\mathbf{1}\{Z_2 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\}) + \operatorname{Cov}(\mathbf{1}\{-Z_2 > Z_j\}, \mathbf{1}\{-Z_2 > Z_1\})\Big]}_{=0} + p^2.
\end{aligned}$$

Therefore, we obtain the overall result on D_n, which is

$$\operatorname{Cov}(X_{\text{inv}}^D, X_{\text{des}}^D) = (n-1)\left(\frac{p^2}{2} + p^2 q - \frac{p}{2} + \frac{1}{4}\right) + p^2.$$

Next, we show that this result is obtained on B_n as well. By (17a) and (16a), we have

$$\operatorname{Cov}(X_{\text{inv}}^B, X_{\text{des}}^B) = (23a) + (23b)$$

$$+ \sum_{i=1}^{n} \sum_{k=1}^{n-1} \operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_k > Z_{k+1}\}) \tag{24a}$$

$$+ \sum_{i<j} \operatorname{Cov}(\mathbf{1}\{Z_i > Z_j\} + \mathbf{1}\{-Z_i > Z_j\}, \mathbf{1}\{Z_1 < 0\}) \tag{24b}$$

$$+ \sum_{i=1}^{n} \operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_1 < 0\})\,. \tag{24c}$$

We now see that (24a) and (24b) vanish. In (24a), the inner sum only involves $k = i - 1$ and $k = i$, therefore,

$$\begin{aligned}
(24a) &= \sum_{i=1}^{n} \Big[\operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_{i-1} > Z_i\}) + \operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_i > Z_{i+1}\})\Big] \\
&= \sum_{i=1}^{n} \Big[\operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_{i-1} > Z_i\}) + \operatorname{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_i > Z_{i-1}\})\Big] \\
&= \mathbb{E}(\mathbf{1}\{Z_1 < 0\}) - 2p/2 = 0.
\end{aligned}$$

This cancellation even applies to the border terms $i = 1$ and $i = n$. We also get

$$(24b) = \sum_{j=2}^{n} \Big[\mathrm{Cov}(\mathbf{1}\{Z_1 > Z_j\}, \mathbf{1}\{Z_1 < 0\}) + \mathrm{Cov}(\mathbf{1}\{-Z_1 > Z_j\}, \mathbf{1}\{Z_1 < 0\})\Big]$$

$$= \underbrace{\sum_{j=2}^{n} \Big[\mathbb{P}(Z_1 < 0, Z_1 > Z_j) - p/2\Big]}_{= -pq/2} + \underbrace{\sum_{j=2}^{n} \Big[\mathbb{P}(Z_1 < 0, -Z_1 > Z_j) - p^2\Big]}_{= pq/2} = 0.$$

Finally,

$$(24c) = \sum_{i=1}^{n} \mathrm{Cov}(\mathbf{1}\{Z_i < 0\}, \mathbf{1}\{Z_1 < 0\}) = \mathrm{Var}(\mathbf{1}\{Z_1 < 0\}) = p - p^2,$$

giving the overall result

$$\mathrm{Cov}(X^B_{\mathrm{inv}}, X^B_{\mathrm{des}}) = (n-1)\left(\frac{p^2}{2} + p^2 q - \frac{p}{2} + \frac{1}{4}\right) + (p - p^2),$$

so again, $\mathrm{Corr}(X^B_{\mathrm{inv}}, X^B_{\mathrm{des}})$ vanishes in the limit. □

8.6. **Proof of Lemma 5.6b).** For the groups B_n and D_n, with $Z_k \sim \mathrm{GR}(p)$ and the modifications (16a), (16b), the calculation is more extensive but its procedure is the same as in the proof of Lemma 2.6b). On D_n, we first have

$$\mathrm{Cov}(\hat{X}^D_{\mathrm{inv}}, X^D_{\mathrm{des}}) = \sum_{j=1}^{n} \sum_{k=1}^{n-1} (n-j)\mathrm{Cov}(U_j \mathbf{1}\{Z_j > 0\}, \mathbf{1}\{Z_k > Z_{k+1}\}) \tag{25a}$$

$$+ (n-j)\mathrm{Cov}((2-U_j)\mathbf{1}\{Z_j < 0\}, \mathbf{1}\{Z_k > Z_{k+1}\}) \tag{25b}$$

$$+ (j-1)\mathrm{Cov}(2pU_j - U_j + 1, \mathbf{1}\{Z_k > Z_{k+1}\}). \tag{25c}$$

$$+ \sum_{j=1}^{n} (n-j)\mathrm{Cov}(U_j \mathbf{1}\{Z_j > 0\}, \mathbf{1}\{-Z_2 > Z_1\}) \tag{25d}$$

$$+ (n-j)\mathrm{Cov}((2-U_j)\mathbf{1}\{Z_j < 0\}, \mathbf{1}\{-Z_2 > Z_1\}) \tag{25e}$$

$$+ (j-1)\mathrm{Cov}(2pU_j - U_j + 1, \mathbf{1}\{-Z_2 > Z_1\}). \tag{25f}$$

In the first three rows (25a) – (25c), there is cancellation if $j \notin \{1, n\}$ due to previously used arguments. Only $j = 1$ is relevant in (25a), (25b) and only $j = n$ is relevant in (25c). We have $\mathbb{E}(U_j \mathbf{1}\{Z_j > 0\})\mathbb{E}(\mathbf{1}\{Z_j > Z_{j+1}\}) = q/4$, and the joint density of Z_j and Z_{j+1} is

$$f_p(x,y) := f_p(x) f_p(y) = \begin{cases} p^2, & x, y < 0 \\ pq, & x > 0, y < 0 \\ pq, & x < 0, y > 0 \\ q^2, & x, y > 0 \end{cases}.$$

By Fubini's Theorem, we obtain

$$\mathbb{E}(U_1 \mathbf{1}\{Z_1 > 0\} \mathbf{1}\{Z_1 > Z_2\}) = \int_{[-1,1]^2} |x| \mathbf{1}\{x > 0\} \mathbf{1}\{x > y\} f_p(x,y) \mathrm{d}(x,y)$$

$$
\begin{aligned}
&= \int_{[0,1]^2} q^2 x \mathbf{1}\{x > y\} \mathrm{d}(x,y) + \int_{[0,1]\times[-1,0]} pqx \mathbf{1}\{x > y\} \mathrm{d}(x,y) \\
&= q^2 \int_0^1 x^2 dx + pq \int_0^1 x \mathrm{d}x = \frac{q^2}{3} + \frac{pq}{2},
\end{aligned}
$$

and likewise,

$$
\begin{aligned}
\mathbb{E}\big((2-U_1)\mathbf{1}\{Z_1 < 0\}\mathbf{1}\{Z_1 > Z_2\}\big) &= \int_{-1}^0 \int_{-1}^1 (2-|x|)\mathbf{1}\{x > y\} f_p(x,y) \mathrm{d}(x,y) \\
&= p^2 \int_{-1}^0 (2+x)(1+x)\mathrm{d}x = \frac{5}{6}p^2, \\
\mathbb{E}\big((2p-1)U_n \mathbf{1}\{Z_{n-1} > Z_n\}\big) &= (2p-1) \int_{[-1,1]^2} |x| \mathbf{1}\{x < y\} f_p(x,y) \mathrm{d}(x,y) \\
&= (2p-1)\left(\frac{p^2}{3} + \frac{q^2}{6} + \frac{pq}{2}\right) = (2p-1)\frac{p+1}{6}.
\end{aligned}
$$

Therefore,

$$
(25a) + (25b) + (25c) = (n-1)\left(-\frac{p^3}{3} + \frac{5p^2}{3} - \frac{4p}{3} + \frac{1}{6}\right)
$$

has linear order. The remaining three rows (25d) – (25f) also have no more than linear order, since the summands are nonzero only for $j = 1$ or $j = 2$. □

References

[1] Arslan, İ., Işlak, Ü., and Pehlivan, C. On unfair permutations. *Statistics & Probability Letters 141* (2018), 31–40.

[2] Bender, E. A. Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A 15*, 1 (1973), 91–111.

[3] Björner, A., and Brenti, F. *Combinatorics of Coxeter groups*, vol. 231. Springer Science & Business Media, 2006.

[4] Brück, B., and Röttger, F. A central limit theorem for the two-sided descent statistic on Coxeter groups. *Electronic Journal of Combinatorics 29*, 1 (2022), P1.1.

[5] Chang, J., Chen, X., and Wu, M. Central limit theorems for high dimensional dependent data. *arXiv preprint arXiv:2104.12929* (2021).

[6] Chatterjee, S., and Diaconis, P. A central limit theorem for a new statistic on permutations. *Indian Journal of Pure and Applied Mathematics 48*, 4 (2017), 561–573.

[7] Chernozhukov, V., Chetverikov, D., and Kato, K. Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics 41*, 6 (2013), 2786–2819.

[8] Chernozhukov, V., Chetverikov, D., and Kato, K. Central limit theorems and bootstrap in high dimensions. *The Annals of Probability 45*, 4 (2017), 2309–2352.

[9] Chernozhukov, V., Chetverikov, D., and Koike, Y. Nearly optimal central limit theorem and bootstrap approximations in high dimensions. *The Annals of Applied Probability 33*, 3 (2023), 2374–2425.

[10] Conger, M. A Refinement of the Eulerian Numbers, and the Joint Distribution of $\pi(1)$ and $\mathrm{Des}(\pi)$ in S_n. *arXiv preprint math/0508112* (2005).

[11] Conger, M., and Viswanath, D. Normal approximations for descents and inversions of permutations of multisets. *Journal of Theoretical Probability 20*, 2 (2007), 309–325.

[12] Das, D., and Lahiri, S. Central limit theorem in high dimensions: The optimal bound on dimension growth rate. *Transactions of the American Mathematical Society 374*, 10 (2021), 6991–7009.

[13] Dörr, P., and Kahle, T. Extreme values of permutation statistics. *arXiv preprint arXiv:2205.01426* (2022).
[14] Fang, X., and Koike, Y. High-dimensional central limit theorems by Stein's method. *The Annals of Applied Probability 31*, 4 (2021), 1660–1686.
[15] Fang, X., and Röllin, A. Rates of convergence for multivariate normal approximation with applications to dense graphs and doubly indexed permutation statistics. *Bernoulli 21*, 4 (2015), 2157–2189.
[16] Fulman, J. Stein's method and non-reversible markov chains. *Lecture Notes-Monograph Series* (2004), 69–77.
[17] Harper, L. H. Stirling behavior is asymptotically normal. *Ann. Math. Statist. 38* (1967), 410–414.
[18] Kahle, T., and Stump, C. Counting inversions and descents of random elements in finite Coxeter groups. *Mathematics of Computation 89*, 321 (2020), 437–464.
[19] Koike, Y. Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. *Japanese Journal of Statistics and Data Science 4*, 1 (2021), 257–297.
[20] Leadbetter, M. R., Lindgren, G., and Rootzén, H. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 1983.
[21] Meier, K., and Stump, C. Central limit theorems for generalized descents and generalized inversions in finite root systems. *preprint, arXiv:2202.05580* (2022).
[22] Pike, J. Convergence rates for generalized descents. *The Electronic Journal of Combinatorics* (2011), P236–P236.
[23] Pitman, J. Probabilistic bounds on the coefficients of polynomials with only real zeros. *Journal of Combinatorial Theory, Series A 77*, 2 (1997), 279–303.
[24] van der Vaart, A. W. *Asymptotic statistics*, vol. 3. Cambridge university press, 2000.

Department of Mathematics, Otto-von-Guericke University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
Email address: philip.doerr@ovgu.de

Department of Mathematics, Stockholm University, Albano hus 1, 10691 Stockholm, Sweden
Email address: johannes.heiny@math.su.se