CYCLIC AND ALTERNATING U-STATISTICS

SVANTE JANSON

Abstract. We define *cyclic U-statistics* as a variant of U-statistics based on variables $X_1, \dots, X_n$ that are assumed to be cyclically ordered. We also define *alternating U-statistics* where in the definition terms are summed with alternating sings (in three different ways). Only U-statistics of order 2 are considered. The definitions are inspired by special cases studied by Chebikin (2008) and Even-Zohar (2017) for random permutations.

We show limit theorems similar to well-known results for standard U-statistics, but with some differences between the different versions. In particular, we find both "nondegenerate" normal limits and "degenerate" non-normal limits.

1. Introduction

U-statistics were introduced by Hoeffding [10] as statistics of the form

$$\bar{U}_n = \bar{U}_n(f) = \bar{U}_n(f; X_1, \dots, X_n) := \sum f(X_{i_1}, \dots, X_{i_m}) \tag{1.1}$$

where m (the *order* of the U-statistic) and n are integers with $1 \leqslant m \leqslant n$, the sum is over all $(n)_m = n!/(n-m)!$ different m-tuples $i_1, \dots, i_m$ of distinct indices in $\{1, \dots, n\}$, $X_1, \dots, X_n$ is a sequence of random variables taking values in some measurable space $\mathcal{X}$, and $f : \mathcal{X}^m \to \mathbb{R}$ is a measurable function, called the *kernel* of the U-statistic. (Hoeffding [10] and many later authors include in the definition a normalization factor $1/(n)_m$; this is often convenient, but in the present paper we choose to omit such factors in the definitions.) The random variables X_i are usually assumed to be independent and identically distributed (i.i.d.), and this will be assumed in the present paper. We will also assume that the kernel f is square integrable in the sense $\mathbb{E}\,|f(X_1, \dots, X_m)|^2 < \infty$, which we write as $f \in L^2 = L^2(\mathcal{X}^m)$.

In the definition (1.1), the order of the variables $X_1, \dots, X_n$ does not matter; in other words, the indices $1, \dots, n$ are used for labelling but their order does not matter and any other labels could be used. Another definition of U-statistics where the order of the variables matters is obtained by summing only over increasing sequences $i_1 < \dots < i_m$:

$$U_n = U_n(f) = U_n(f; X_1, \dots, X_n) := \sum_{i_1 < \dots < i_m} f(X_{i_1}, \dots, X_{i_m}). \tag{1.2}$$

Note that if f is symmetric, as is often assumed, the definitions (1.1) and (1.2) differ only by an unimportant factor $m!$. In fact, $\bar{U}_n$ in (1.1) remains the same if f is replaced by its symmetrization; hence we may as well assume that f in (1.1) is symmetric, and therefore (1.1) can be seen as a special case of the more general (1.2). (Although many applications use only the symmetric version (1.1), or equivalently (1.2) with a symmetric kernel f, there are also many applications that require the

Date: 14 October, 2025.

Supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

asymmetric version (1.2).) The literature on U-statistics and applications of them is enormous, and we will in the sequel only give a few relevant references.

In the present paper we consider four modifications of the definitions above, defined in the following two subsections. As in (1.1) and (1.2), we may include f and the variables X_i in the notation, but we often omit them when they are clear from the context. Our main results are theorems on the asymptotic distribution of these modifications, stated in Sections 4–6 and summarized for convenience in Section 7. The results are similar to the well-known results for classical U-statistics, which we state for comparison in Section 3, although the details in the limit theorems differ between the different versions. In particular, in the classical case $U_n(f)$ there is a well-known dichotomy of the kernels f into a nondegenerate case with variance of order n^3 and an asymptotically normal distribution, and a degenerate case with a variance of order n^2 only and with a non-normal limit distribution; all but one of the versions studied here exhibit the same two cases, but not necessarily for the same kernels; however, for one version (Theorem 5.1) there is no “nondegenerate” case. Furthermore, for some versions there is an exceptional third, futher degenerate and almost trivial case, with variance of order n and again a normal limit distribution. In the degenerate case, for all versions, the limit distribution can be expressed as a (possibly infinite) linear combination of centred squares of independent standard normal variables, where the coefficients are the eigenvalues of a certain integral operator with kernel derived from the kernel f of the U-statistic, although again the details differ between the differnt versions; see e.g. (3.6). (From an abstract point of view, the limit distribution is a Wiener chaos of order 2, see e.g. [12, in particular Theorem 6.1].)

Some simple examples are given in Section 8; this includes applications to the *writhe* and *alternating inversion number* of a uniformly random permutation, previously defined and studied in [6] and [2]. These two examples were the inspiration of the definitions and results in the present paper.

Some further results and open problems are given in Section 9.

Appendix A contains the proof of the well-known Theorem 3.1 for classical U-statistics, included for completeness; we also reuse parts of the appendix in other proofs. Appendix B collects some formulas for cumulants. Appendix C gives further calculations for one example from Section 8.

As mentioned above, in the degenerate cases, the asymptotic distribution is described by the eigenvalues of an integral operators. Consequently, some proofs and most examples require finding such eigenvalues; this is straightforward in our cases, but we include details for completeness.

1.1. Cyclic U-statistics.

For this version, we regard the indices $1, \dots, n$ as circularly ordered instead of linearly ordered. We regard the indices as elements of the cyclic group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$; we therefore assume again that $X_1, \dots, X_n$ are i.i.d. random variables, and extend the notation to X_i for all $i \in \mathbb{Z}$ by $X_i := X_j$ if $i \equiv j \pmod{n}$ and $j \in \{1, \dots, n\}$.

We consider only the case $m = 2$ and then define the *cyclic U-statistic*

$$U_n^\circ = U_n^\circ(f) := \sum_{i \in \mathbb{Z}_n} \sum_{1 \leqslant j < n/2} f(X_i, X_{i+j}). \tag{1.3}$$

Note that if we regard the elements of $\mathbb{Z}_n$ as lying on a circle in the natural way, then for any pair of distinct $i, j \in \mathbb{Z}_n$, the sum (1.3) contains the term $f(X_i, X_j)$ if the shortest path from i to j goes in the positive direction, and the term $f(X_j, X_i)$ if the shortest path goes in the negative direction; for even n, no terms $f(X_i, X_{i+n/2})$ appears at all.

One example of a cyclic U-statistic appears in a paper by Even-Zohar [6] where he studies the *writhe* of permutations and framed knots, see Example 8.5 for details. In particular, [6] studies the writhe of a uniformly random permutation and finds its asymptotic distribution. The writhe of a uniformly random permutation can be written as a cyclic U-statistic (1.3) (see Example 8.5 again), and this example is the motivation for the present paper, where we study general cyclic U-statistics (assuming only that $f \in L^2$) and prove general limit theorems; in particular, this gives an alternative proof of the limit theorem by Even-Zohar [6] (where the theorem is proved by quite different methods).

Remark 1.1. If $f(x, y)$ is a symmetric function, then we have

$$U_n^\circ = \begin{cases} U_n, & n \text{ is odd} \\ U_n - \sum_{i=1}^{n/2} f(X_i, X_{i+n/2}) & n \text{ is even,} \end{cases} \tag{1.4}$$

where it is easily seen that the sum appearing for even n is asymptotically negligible. (If $f \in L^2$, this sum has variance $O(n)$ since its terms are i.i.d., while U_n except in trivial cases has variance of order at least n^2.) Hence, we do not really obtain anything new for symmetric f. The main interest seems to be in the opposite case, when f is antisymmetric. We will in the sequel give special attention to the cases of symmetric and antisymmetric kernels.

Note that every f may be decomposed as a sum of a symmetric and an antisymmetric part; hence a U-statistic of order $m = 2$ (of any of the versions studied here) can be written as a sum of a symmetric and an antisymmetric U-statistic; We will see in Theorem 4.1 and Remark 4.4 that for the cyclic U-statistic U_n°, the two parts are asymptotically independent, and thus can be treated separately. However, for the other U-statistics considered here, including the classical U_n, this decomposition is of limited value since the two parts typically are dependent, also asymptotically. For an interesting example of this, largely taken from [14], see Example 8.2. △

Remark 1.2. A variation of the definition (1.3) is the more symmetrical

$$\widetilde{U}_n^\circ(f) := \sum_{i \in \mathbb{Z}_n} \sum_{1 \leqslant j < n/2} f(X_i, X_{i+j}) - \sum_{i \in \mathbb{Z}_n} \sum_{1 \leqslant j < n/2} f(X_i, X_{i-j}). \tag{1.5}$$

However, replacing i by $i + j$ in the second sum yields

$$\widetilde{U}_n^\circ(f) = \sum_{i \in \mathbb{Z}_n} \sum_{1 \leqslant j < n/2} \big(f(X_i, X_{i+j}) - f(X_{i+j}, X_i)\big) = U_n^\circ(g), \tag{1.6}$$

where $g(x, y) := f(x, y) - f(y, x)$. Hence, cyclic U-statistics of the form (1.5) are special cases of U_n°, and therefore need not be considered further. Note that the function g arising here always is antisymmetric; conversely, if f is antisymmetric, then (1.6) implies $U_n^\circ(f) = \widetilde{U}_n^\circ(\frac{1}{2} f)$. Consequently, the version $\widetilde{U}_n^\circ$ is equivalent to the special case of U_n° for antisymmetric kernels only. △

1.2. **Alternating U-statistics.** One of the results in [6] is that the writhe of a uniformly random permutation has the same distribution as the *bi-alternating inversion number*, which is defined in [6] in analogy to the *alternating inversion number* defined in [2], see Example 8.6; it is also shown in [6] that the alternating and bi-alternating inversion numbers of a uniformly random permutation, although similarly defined, have quite different asymptotic distributions.

These numbers are defined as versions of the usual inversion number of a permutation. It is well-known that the usual inversion number of a uniformly random permutation can be written as a U-statistic. Again we study corresponding modifications of general U-statistics. We consider again only the case of order $m = 2$, so that (1.2) becomes

$$U_n = U_n(f) := \sum_{1\leqslant i<j\leqslant n} f(X_i, X_j). \tag{1.7}$$

We then define the *alternating U-statistics*

$$U_n^{-+} = U_n^{-+}(f) := \sum_{1\leqslant i<j\leqslant n} (-1)^{i+1} f(X_i, X_j), \tag{1.8}$$

$$U_n^{+-} = U_n^{+-}(f) := \sum_{1\leqslant i<j\leqslant n} (-1)^{j} f(X_i, X_j), \tag{1.9}$$

$$U_n^{--} = U_n^{--}(f) := \sum_{1\leqslant i<j\leqslant n} (-1)^{i+j} f(X_i, X_j). \tag{1.10}$$

We also call U_n^{--} *bi-alternating.* (As in [6] for a special case, see Example 8.5.)

Remark 1.3. Define

$$f^*(x, y) := f(y, x). \tag{1.11}$$

Then, by replacing X_i by X_{n+1-i}, it follows that

$$\begin{aligned} U_n^{+-}(f) &\overset{\mathrm{d}}{=} \sum_{1\leqslant i<j\leqslant n} (-1)^{j} f(X_{n+1-i}, X_{n+1-j}) = (-1)^n \sum_{1\leqslant k<\ell\leqslant n} (-1)^{k+1} f(X_\ell, X_k) \\ &= (-1)^n U_n^{-+}(f^*). \end{aligned} \tag{1.12}$$

Consequently, up to a trivial change of sign and replacing f by f^*, U_n^{+-} is the same as U_n^{-+}, and thus it suffices to consider the latter. On the other hand, as noted in [6] for the example in Example 8.6, we will see that U^{--} in general is quite different. △

2. Preliminaries

2.1. **Some notation.** We assume throughout that $X_1, X_2, \dots$ are i.i.d. random variables with values in some measurable space $\mathcal{X}$, and let ν be the common distribution of X_i. We let X denote any random variable with this distribution. Thus $(\mathcal{X}, \nu)$ is a probability space, which we for simplicity also denote by $\mathcal{X}$.

We define $\widehat{\mathcal{X}} := \mathcal{X} \times [0, 1]$, where (as always below), $[0, 1]$ is equipped with the Lebesgue measure ℓ.

We assume also that $f : \mathcal{X}^2 \to \mathbb{R}$ is a given function such that $f \in L^2(\mathcal{X}^2)$, i.e.,

$$\mathbb{E}\,|f(X_1, X_2)|^2 = \int_{\mathcal{X}\times\mathcal{X}} |f(x, y)|^2 \,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y) < \infty. \tag{2.1}$$

We define

$$\mu := \mathbb{E}\, f(X_1, X_2). \tag{2.2}$$

We will mainly consider real-valued functions, but to apply functional analysis it will sometimes be convenient to also consider complex-valued fuctions. When it is necessary to distinguish them, we use $L^2_{\mathbb{C}}(\mathcal{X})$ for the complex Hilbert space of complex-valued functions g on $\mathcal{X}$ such that $\int_{\mathcal{X}} |g|^2 \,\mathrm{d}\nu < \infty$, and $L^2_{\mathbb{R}}(\mathcal{X})$ for the subspace of real-valued functions (this is a real Hilbert space).

All functions are assumed to be measurable. We will sometimes omit "a.s." or "a.e." when this is obvious.

For a function $g : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$, let T_g denote the integral operator on $L^2(\mathcal{X})$ defined by

$$T_g\varphi(x) := \int_{\mathcal{X}} g(x,y)\varphi(y) \,\mathrm{d}\nu(y), \qquad \varphi \in L^2(\mathcal{X}). \tag{2.3}$$

We will only consider the case when $g \in L^2(\mathcal{X} \times \mathcal{X})$; it is well-known that then, for every $h \in L^2(\mathcal{X})$, the integral in (2.3) converges for a.e. x and that $T_g h \in L^2(\mathcal{X})$, so that T_g is well-defined, and furthermore that T_g is a Hilbert–Schmidt operator on $L^2(\mathcal{X})$, and thus in particular compact (and bounded). (See e.g. [3, Proposition II.4.7 and Exercise IX.2.19].) We will also use the notation (2.3) for other measure spaces, in particular $\mathcal{X}^2 = \mathcal{X} \times \mathcal{X}$ and $\widehat{\mathcal{X}} = \mathcal{X} \times [0,1]$.

Recall that a bounded operator T on a Hilbert space H is *self-adjoint* (also called *Hermitian*, or *symmetric*) if $\langle Th, k\rangle = \langle h, Tk\rangle$ for all $h, k \in H$. It is easily seen that if g is real-valued and symmetric (i.e., $g(x,y) = g(y,x)$), then T_g is self-adjoint.

Eigenvalues of operators are always counted with multiplicities; sets of eigenvalues are thus in general really multisets.

We will occasionally also use tensor notation for functions and operators. If g and h are (real- or complex-valued) functions defined on measure spaces $\mathcal{Y}$ and $\mathcal{Z}$, then $g \otimes h$ denotes the function $(y,z) \mapsto g(y)h(z)$ defined on $\mathcal{Y} \times \mathcal{Z}$. It is well known that if $(\varphi_\alpha)_{\alpha\in\mathcal{A}}$ and $(\psi_\beta)_{\beta\in\mathcal{B}}$ are orthonormal bases in $L^2(\mathcal{Y})$ and $L^2(\mathcal{Z})$, then $(\varphi_\alpha \otimes \psi_\beta)_{\alpha\in\mathcal{A},\beta\in\mathcal{B}}$ is an orthonormal basis in $L^2(\mathcal{Y} \times \mathcal{Z})$. If $g \in L^2(\mathcal{Y} \times \mathcal{Y})$, and $h \in L^2(\mathcal{Z} \times \mathcal{Z})$, then $g \otimes h$ can be regarded as a kernel on $\mathcal{Y} \times \mathcal{Z}$, and it is easily seen that

$$T_{g\otimes h}(\varphi \otimes \psi) = (T_g\varphi) \otimes (T_h\psi), \qquad \varphi \in L^2(\mathcal{Y}), \psi \in L^2(\mathcal{Z}). \tag{2.4}$$

We write also $T_g \otimes T_h := T_{g\otimes h}$. (This is a special case of tensor products of linear operators on Hilbert, or more general, spaces, but we have no need for the general theory.) It follows from (2.4) that if φ is an eigenfunction of T_g with eigenvalue λ and ψ is an eigenfunction of T_h with eigenvalue ρ, then $\varphi \otimes \psi$ is an eigenfunction of $T_g \otimes T_h = T_{g\otimes h}$ with eigenvalue $\lambda\rho$.

For a function f on $\mathcal{X} \times \mathcal{X}$, we define its *symmetric* and *antisymmetric parts* by, recalling (1.11),

$$f_{\mathsf{s}} := \tfrac12(f + f^*) \qquad \text{and} \qquad f_{\mathsf{a}} := \tfrac12(f - f^*). \tag{2.5}$$

Thus $f = f_{\mathsf{s}} + f_{\mathsf{a}}$. We also define (as another form of symmetrization) the symmetric function $\widehat{f}$ on $\widehat{\mathcal{X}} \times \widehat{\mathcal{X}}$ by (recall that $\widehat{\mathcal{X}} := \mathcal{X} \times [0,1]$)

$$\widehat{f}\bigl((x,t),(y,u)\bigr) := \begin{cases} f(x,y), & t < u, \\ f(y,x), & t > u. \end{cases} \tag{2.6}$$

(For completeness, we may define $\widehat{f}\bigl((x,t),(y,u)\bigr) := 0$ when $t = u$; this case has measure 0 and is therefore irrelevant.) Note that, with $\widehat{\nu} := \nu \times \ell$,

$$\int_{\widehat{\mathcal{X}}^2} |\widehat{f}(\widehat{x},\widehat{y})|^2 \,\mathrm{d}\widehat{\nu}(\widehat{x}) \,\mathrm{d}\widehat{\nu}(\widehat{y})$$

$$= \int_{[0,1]^2} \int_{\mathcal{X}^2} |\widehat{f}\big((x,t),(y,u)\big)|^2 \,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y)\,\mathrm{d}t\,\mathrm{d}u$$
$$= \int_{[0,1]^2} \int_{\mathcal{X}^2} |f(x,y)|^2 \,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y)\,\mathrm{d}t\,\mathrm{d}u = \int_{\mathcal{X}^2} |f(x,y)|^2 \,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y). \tag{2.7}$$

In some cases (when studying U^{-+} and U^{+-}) we need also the corresponding antisymmetric function on $\widehat{\mathcal{X}} \times \widehat{\mathcal{X}}$ defined by

$$\breve{f}\big((x,t),(y,u)\big) := \begin{cases} f(x,y), & t < u, \\ -f(y,x), & t > u. \end{cases} \tag{2.8}$$

For a real number x, we let $\lfloor x \rfloor$ be x rounded down to the nearest integer. The complex unit is denoted i. (This should not be confused with i, often used to denote indices.) $\mathbf{1}\{\mathcal{E}\}$ denotes the indicator function of an event $\mathcal{E}$. The sign function sgn is given by

$$\operatorname{sgn}(x) := \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases} \tag{2.9}$$

We let $\xrightarrow{\mathrm{d}}$, $\xrightarrow{\mathrm{p}}$, and $\xrightarrow{\mathrm{a.s.}}$ denote convergence of random variables in distribution, in probability, and almost surely, respectively.

Given a sequence $(a_n)_n$, we let $Y_n = O_{L^2}(a_n)$ mean that Y_n are random variables such that $\|Y_n\|_{L^2} := (\mathbb{E}\,[|Y_n|^2])^{1/2} = O(a_n)$. Similarly, $Y_n = o_{L^2}(a_n)$ means that $\|Y_n\|_{L^2} = o(a_n)$, in other words that $Y_n/a_n \to 0$ in L^2.

Unspecified limits are as $n \to \infty$.

2.2. **Hoeffding's decomposition.** As said above, we assume $f \in L^2(\mathcal{X}^2)$. The basis of our work (as for many previous results for U-statistics) is the orthogonal decomposition introduced (in the symmetric case) by Hoeffding [10]. In the case $m = 2$ treated here, the orthogonal decomposition is:

$$f(x,y) = f_\varnothing + f_1(x) + f_2(y) + f_{12}(x,y) \tag{2.10}$$

where

$$f_\varnothing := \mathbb{E}\, f(X_1,X_2) = \int_{\mathcal{X}\times\mathcal{X}} f(x,y)\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y) = \mu, \tag{2.11}$$
$$f_1(x) := \mathbb{E}\, f(x,X) - f_\varnothing = \int_{\mathcal{X}} f(x,y)\,\mathrm{d}\nu(y) - f_\varnothing, \tag{2.12}$$
$$f_2(y) := \mathbb{E}\, f(X,y) - f_\varnothing = \int_{\mathcal{X}} f(x,y)\,\mathrm{d}\nu(x) - f_\varnothing, \tag{2.13}$$
$$f_{12}(x,y) := f(x,y) - f_1(x) - f_2(y) - f_\varnothing. \tag{2.14}$$

Equivalently, (2.12)–(2.14) can be written

$$f_1(X_1) = \mathbb{E}\,[f(X_1,X_2) \mid X_1] - \mathbb{E}\, f(X_1,X_2), \tag{2.15}$$
$$f_2(X_2) = \mathbb{E}\,[f(X_1,X_2) \mid X_2] - \mathbb{E}\, f(X_1,X_2), \tag{2.16}$$
$$f_{12}(X_1,X_2) = f(X_1,X_2) - \mathbb{E}\,[f(X_1,X_2) \mid X_1] - \mathbb{E}\,[f(X_1,X_2) \mid X_2] + \mathbb{E}\, f(X_1,X_2). \tag{2.17}$$

Of course, (2.14) makes (2.10) trivial, but the point is that the four terms in the sum in (2.10) are orthogonal in $L^2(\mathcal{X}^2)$, which is easily verified from the definitions (2.11)–(2.17), which imply

$$\mathbb{E} f_1(X) = \mathbb{E} f_2(X) = 0, \tag{2.18}$$

$$\mathbb{E} f_{12}(x, X) = \mathbb{E} f_{12}(X, y) = \mathbb{E} f_{12}(X_1, X_2) = 0. \tag{2.19}$$

Lemma 2.1. *Let (i,j) and (k,l) be two pairs of indices with $i \neq j$, $k \neq l$, and $\{i,j\} \neq \{k,l\}$ (i.e., $(i,j) \neq (k,l)$ and $(i,j) \neq (l,k)$). Then $f_{12}(X_i, X_j)$ and $f_{12}(X_k, X_l)$ are uncorrelated and thus $\mathbb{E}\big[f_{12}(X_i,X_j)f_{12}(X_k,X_l)\big] = 0$.*

Proof. A simple consequence of (2.19) and the standing assumption that (X_i) are i.i.d. ☐

Remark 2.2. If f is symmetric, then $f_1 = f_2$, and furthermore, f_{12} is symmetric. On the other hand, if f is antisymmetric, then $f_\varnothing = 0$, $f_1 = -f_2$, and f_{12} is antisymmetric. △

2.3. **Three distributions.** The limit distributions below will, apart from normal distributions, be given by (possibly infinite) linear combinations of independent copies of the following three random variables.

(i) If $\zeta \in N(0,1)$, then ζ^2 has a $\chi^2(1) = \Gamma(\frac{1}{2}, 2)$ distribution. We will use the centred variable $\zeta^2 - 1$, which has mean 0, variance

$$\operatorname{Var}\big(\zeta^2 - 1\big) = 2, \tag{2.20}$$

and characteristic function

$$\mathbb{E}\, e^{\mathrm{i}t(\zeta^2-1)} = e^{-\mathrm{i}t}(1-2\mathrm{i}t)^{-1/2}, \qquad t \in \mathbb{R}. \tag{2.21}$$

(ii) Lévy's *stochastic area*, which we denote by η, is the stochastic integral

$$\eta := \int_0^1 B_1(x)\,\mathrm{d}B_2(x) - \int_0^1 B_2(x)\,\mathrm{d}B_1(x) \tag{2.22}$$

where $B_1(x)$ and $B_2(x)$ are two independent Brownian motions. See e.g. [17], [24], [21, Theorem II.43 and its Corollary]. For us this background is not important; we only need that the stochastic area is a random variable η with the characteristic function

$$\mathbb{E}\, e^{\mathrm{i}t\eta} = \frac{1}{\cosh(t)}, \qquad t \in \mathbb{R}. \tag{2.23}$$

The stochastic area has variance, from (2.22) or from (2.23),

$$\operatorname{Var} \eta = 1, \tag{2.24}$$

and density $\frac{1}{2\cosh(\pi x/2)}$, $-\infty < x < \infty$.

(iii) Thirdly, we define ϑ by $\vartheta := \int_0^1 B_1(x)\,\mathrm{d}B_2(x)$, i.e., "the first half of the stochastic area (2.22)". This random variable has the characteristic function, see e.g. [24, (1)] (or as a consequence of the calculations in Example 8.2 and (2.31) below).

$$\mathbb{E}\, e^{\mathrm{i}\vartheta} = \frac{1}{\cosh^{1/2}(t)}, \tag{2.25}$$

and hence variance

$$\operatorname{Var} \vartheta = \tfrac{1}{2}. \tag{2.26}$$

Remark 2.3. The characteristic function (2.23) of the difference of the two stochastic integrals in (2.22) thus equals the product of their characteristic functions, which both are (2.25); however, the two integrals are *not independent*. Their joint characteristic function is given by, see [24, (1)],

$$\mathbb{E}\Bigl[\exp\Bigl(\mathrm{i}s\int_0^1 B_1(x)\,\mathrm{d}B_2(x)+\mathrm{i}t\int_0^1 B_2(x)\,\mathrm{d}B_1(x)\Bigr)\Bigr] \\ = \Bigl(\cosh^2\Bigl(\frac{s-t}{2}\Bigr)+\Bigl(\frac{s+t}{s-t}\Bigr)^2\sinh^2\Bigl(\frac{s-t}{2}\Bigr)^2\Bigr)^{-1/2}. \tag{2.27}$$

△

Remark 2.4. We will frequently use sums $\sum_{r=1}^R \lambda_r(\zeta_r^2-1)$ where $(\lambda_r)_1^R$ is a finite or infinite sequence of real numbers with $\sum_{r=1}^R \lambda_r^2<\infty$, and $\zeta_r\in N(0,1)$ are independent. Note that this sum converges in L^2 (and a.s.), and is thus well defined also when $R=\infty$. Furthermore, it is easily seen from (2.21) that the sum $\sum_{r=1}^R \lambda_r(\zeta_r^2-1)$ has a characteristic function whose square extends to a meromorphic function in the complex plane with poles (counted with multiplicity) at the points $1/(2\mathrm{i}\lambda_r)$ and nowhere else. Thus the distribution of the sum determines the coefficients $(\lambda_r)_1^R$ (up to order).

By the representations in Lemma 2.5 below, the same holds for sums $\sum_{r=1}^R \lambda_r\eta_r$ and $\sum_{r=1}^R \lambda_r\vartheta_r$.

It follows also that the distribution of the sum is *not* normal, since its characteristic function is not entire. This alternatively follows by the Lévy–Cramér theorem [4]. △

Obviously, if ϑ_1 and ϑ_2 are i.i.d. with the distribution (2.25), then $\vartheta_1+\vartheta_2$ and $\vartheta_1-\vartheta_2$ have the distribution (2.23). The following lemma shows further relations between variables in (i)–(iii) above.

Lemma 2.5. *let $\zeta_{k,j}\in N(0,1)$ be independent. Then*

$$\widetilde{\vartheta}:=\sum_{k=-\infty}^{\infty}\frac{1}{(2k-1)\pi}\bigl(\zeta_{k,1}^2-1\bigr)\overset{\mathrm{d}}{=}\sum_{k=1}^{\infty}\frac{1}{(2k-1)\pi}\bigl(\zeta_{k,1}^2-\zeta_{k,2}^2\bigr) \tag{2.28}$$

has the distribution of ϑ in (2.25), *and*

$$\widetilde{\eta}:=\sum_{k=1}^{\infty}\frac{1}{(2k-1)\pi}\bigl(\zeta_{k,1}^2+\zeta_{k,2}^2-\zeta_{k,3}^2-\zeta_{k,4}^2\bigr) \tag{2.29}$$

has the distribution of the stochastic area η in (2.23).

Proof. All sums converge in L^2, cf. Remark 2.4. We obtain the second equality (in distribution) in (2.28) by combining in the first sum the terms for k and $1-k$.

The difference $\zeta_{k,1}^2-\zeta_{k,2}^2$ has the characteristic function, see (2.21),

$$\mathbb{E}\,e^{\mathrm{i}t(\zeta_{k,1}^2-\zeta_{k,2}^2)}=(1-2\mathrm{i}t)^{-1/2}(1+2\mathrm{i}t)^{-1/2}=(1+4t^2)^{-1/2}, \tag{2.30}$$

and consequently, $\widetilde{\vartheta}$ has the characteristic function

$$\mathbb{E}\,e^{\mathrm{i}t\widetilde{\vartheta}}=\prod_{k=1}^{\infty}\Bigl(1+\frac{4t^2}{(2k-1)^2\pi^2}\Bigr)^{-1/2}=\bigl(\cosh(t)\bigr)^{-1/2}, \tag{2.31}$$

where the last equality is well-known, see e.g. [20, (4.36.2)]. Since $\widetilde{\eta} \overset{\mathrm{d}}{=} \widetilde{\vartheta} + \widetilde{\vartheta}'$ where $\widetilde{\vartheta}'$ is an independent copy of $\widetilde{\vartheta}$, it follows that $\widetilde{\eta}$ has the characteristic function

$$\mathbb{E}\, e^{\mathrm{i}t\widetilde{\eta}} = \bigl(\mathbb{E}\, e^{\mathrm{i}t\widetilde{\vartheta}}\bigr)^2 = \frac{1}{\cosh(t)}, \tag{2.32}$$

which agrees with (2.23). □

Formulas for the cumulants of these variables are given in Appendix B.

3. Background: classical U-statistics

As a background, we summarize in the following theorem some known result on the asymptotic distribution of U-statistics of the standard type (1.2) in the special case of order $m = 2$. The general case with arbitrary (fixed) $m \geqslant 2$ is similar with mainly notational complications; the only essential difference is that degeneracies of higher order may occur and then the limiting distributions are much more complicated (although such cases are rarely seen in applications). We restrict ourselves to $m = 2$ because this is the case relevant for the cyclic and alternating U-statistics discussed in the present paper, and also because it may be easier to see the general ideas in this somewhat simpler case. (We do not know any reference where all these results are collected and presented for the case $m = 2$.) For completeness, we give a proof in Appendix A.

For further results and for the general case with arbitrary m, we refer to, for example, [10; 11; 22; 5] for the symmetric case (1.1), and [12, Chapter 11.1–2] for the general (asymmetric) case (1.2). For the strong law of large numbers, see Section 9.2.

In the theorem, note in particular the dichotomy between the *nondegenerate* case with $\sigma^2 > 0$ and then variance of order n^3 and asymptotically normal distribution (see (ii)), and the *degenerate* case in (iii)–(v) with $\sigma^2 = 0$ and then variance of smaller order n^2 and a non-normal limit distribution.

Theorem 3.1. *With notations and assumptions as in Section 2, the following holds.*

(i) *We have*

$$\mathbb{E}\, U_n(f) = \tbinom{n}{2}\mu \tag{3.1}$$

and, as $n \to \infty$, we have the weak law of large numbers

$$\frac{1}{\binom{n}{2}} U_n(f) \xrightarrow{\mathrm{p}} \mu. \tag{3.2}$$

(ii) *As $n \to \infty$,*

$$n^{-3}\operatorname{Var}[U_n(f)] \to \sigma^2 := \tfrac{1}{3}\bigl(\mathbb{E}\,[f_1(X)^2] + \mathbb{E}\,[f_2(X)^2] + \mathbb{E}\,[f_1(X)f_2(X)]\bigr) \tag{3.3}$$

and

$$n^{-3/2}\left(U_n(f) - \tbinom{n}{2}\mu\right) \xrightarrow{\mathrm{d}} N(0,\sigma^2). \tag{3.4}$$

Furthermore, $\sigma^2 > 0$ unless $f_1(X) = f_2(X) = 0$ a.s.

(iii) *If $f_1(X) = f_2(X) = 0$ a.s., and thus $\sigma^2 = 0$, then*

$$\operatorname{Var}[U_n(f)] = \tbinom{n}{2}\operatorname{Var}[f(X_1,X_2)] = \tfrac{1}{2}n^2\operatorname{Var}[f(X_1,X_2)] + O(n). \tag{3.5}$$

Moreover, there exists a finite or infinite sequence of real numbers $(\lambda_r)_1^R$ *such that*

$$n^{-1}\bigl(U_n(f)-\tbinom{n}{2}\mu\bigr)\xrightarrow{\mathrm{d}} W:=\sum_{r=1}^{R}\tfrac12\lambda_r(\zeta_r^2-1), \tag{3.6}$$

where $(\zeta_r)_1^R$ *are independent standard normal variables. The coefficients* $(\lambda_r)_1^R$ *are the nonzero eigenvalues (with multiplicities) of the self-adjoint integral operator* $T_{\widehat{f-\mu}}$ *on* $L^2(\mathcal{X}\times[0,1],\nu\times\ell)$ *(where* ℓ *is Lebesgue measure) defined as in* (2.3) *using* (2.6). *We have*

$$\operatorname{Var} W=\tfrac12\sum_{r=1}^{R}\lambda_r^2=\tfrac12\operatorname{Var}[f(X_1,X_2)]<\infty. \tag{3.7}$$

(iv) *In the special case of* (iii) *where furthermore* f *is symmetric, the coefficients* $(\lambda_r)_1^R$ *in* (3.6) *are the nonzero eigenvalues (with multiplicities) of the self-adjoint integral operator* $T_{f-\mu}$ *on* $L^2(\mathcal{X},\nu)$.

(v) *In the special case of* (iii) *where furthermore* f *is antisymmetric, then also*

$$n^{-1}U_n(f)\xrightarrow{\mathrm{d}} W:=\sum_{q=1}^{Q_+}\lambda_q^{\mathsf{a}}\eta_q, \tag{3.8}$$

where $(\eta_q)_1^{Q_+}$ *are independent random variables with the stochastic area distribution* (2.23), *and the coefficients* $(\lambda_q^{\mathsf{a}})_1^{Q_+}$ *are the positive numbers such that the imaginary number* $\mathrm{i}\lambda_q^{\mathsf{a}}$ *is an eigenvalue of the anti-self-adjoint operator* T_f *on* $L^2_{\mathbb{C}}(\mathcal{X},\nu)$. *We have*

$$\operatorname{Var} W=\sum_{q=1}^{Q_+}(\lambda_q^{\mathsf{a}})^2=\tfrac12\operatorname{Var}[f(X_1,X_2)]. \tag{3.9}$$

As said above, the proof is given in Appendix A.

Remark 3.2. As a sanity check, we note that if f is symmetric, then the nonzero eigenvalues of the operators $T_{\widehat{f-\mu}}$ in (iii) and $T_{f-\mu}$ in (iv) are the same, so the conclusions agree. In fact, if f is symmetric, the (2.6) yields

$$\widehat{f}\bigl((x,t),(y,u)\bigr)=f(x,y). \tag{3.10}$$

Letting $\mathbf{1}$ denote the function on $[0,1]^2$ that is constant 1, we thus have, using the tensor notation in (2.4), $\widehat{f}=f\otimes\mathbf{1}$ and consequently

$$T_{\widehat{f}}=T_f\otimes T_{\mathbf{1}}. \tag{3.11}$$

$T_{\mathbf{1}}$ is the integral operator $T_{\mathbf{1}}g(t)=\int_0^1 g(u)\,\mathrm{d}u$; this is the projection onto the constant functions and has a single nonzero eigenvalue 1. Consequently, $T_{\widehat{f}}=T_f\otimes T_{\mathbf{1}}$ has the same nonzero eigenvalues as T_f. (In this simple case, this can also easily be seen directly from (3.10); the eigenfunctions of $T_{\widehat{f}}$ with nonzero eigenvalues are the functions of the form $\varphi(x,t)=\varphi_1(x)$ where φ_1 is an eigenfunction of T_f with the same eigenvalue.) $\triangle$

4. Cyclic U-statistics

We next give the corresponding result for cyclic U-statistics.

Theorem 4.1. *With notations and assumptions as in Section 2, the following holds for the cyclic U-statistic $U_n^\circ(f)$ in* (1.3).

(i) *We have*

$$\mathbb{E}\left[U_n^\circ(f)\right] = n\lfloor \tfrac{n}{2} \rfloor \mu = \tfrac{n^2}{2}\mu + O(n), \tag{4.1}$$

and, as $n \to \infty$, we have the weak law of large numbers

$$\frac{2}{n^2} U_n^\circ(f) \xrightarrow{\mathrm{p}} \mu. \tag{4.2}$$

(ii) *As $n \to \infty$,*

$$\begin{aligned} n^{-3} \operatorname{Var}[U_n^\circ(f)] \to \sigma^2 &:= \tfrac14 \operatorname{Var}\bigl[f_1(X) + f_2(X)\bigr] \\ &= \tfrac14 \bigl(\mathbb{E}\left[f_1(X)^2\right] + \mathbb{E}\left[f_2(X)^2\right] + 2\,\mathbb{E}\left[f_1(X) f_2(X)\right]\bigr) \end{aligned} \tag{4.3}$$

and

$$n^{-3/2}\Bigl(U_n^\circ(f) - \tfrac{n^2}{2}\mu\Bigr) \xrightarrow{\mathrm{d}} N(0,\sigma^2), \tag{4.4}$$

Furthermore, $\sigma^2 > 0$ unless $f_1(X) + f_2(X) = 0$ a.s.

(iii) *If $f_1(X) + f_2(X) = 0$ a.s., and thus $\sigma^2 = 0$, then*

$$\operatorname{Var}[U_n^\circ(f)] = n\lfloor \tfrac{n}{2} \rfloor \operatorname{Var}[f_{12}(X_1,X_2)] = \tfrac12 n^2 \operatorname{Var}[f_{12}(X_1,X_2)] + O(n). \tag{4.5}$$

Moreover, there exist finite or infinite sequences of real numbers $(\lambda_r^{\mathsf{s}})_1^R$ and $(\lambda_q^{\mathsf{a}})_1^{Q_+}$ such that

$$n^{-1}\left(U_n^\circ(f) - \mathbb{E}\left[U_n^\circ(f)\right]\right) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \tfrac12 \lambda_r^{\mathsf{s}}(\zeta_r^2 - 1) + \sum_{q=1}^{Q_+} \lambda_q^{\mathsf{a}} \eta_q \tag{4.6}$$

where $(\zeta_r)_1^R$ are standard normal variables and $(\eta_q)_1^{Q_+}$ have the stochastic area distribution (2.23)*, and all are independent. The coefficients $(\lambda_r^{\mathsf{s}})_1^R$ in* (4.6) *are the nonzero eigenvalues (with multiplicities) of the self-adjoint integral operator $T_{f_{12\mathsf{s}}}$ on $L^2(\mathcal{X},\nu)$, where, recalling* (2.5)*, $f_{12\mathsf{s}} := (f_{12} + f_{12}^*)/2$ is the symmetric part of f_{12}. Similarly, the coefficients $(\lambda_q^{\mathsf{a}})_1^{Q_+}$ are the positive numbers such that the imaginary number $\mathrm{i}\lambda_q^{\mathsf{a}}$ is an eigenvalue of the anti-self-adjoint operator $T_{f_{12\mathsf{a}}}$ on $L^2_{\mathbb{C}}(\mathcal{X},\nu)$, where $f_{12\mathsf{a}} := (f_{12} - f_{12}^*)/2$ is the antisymmetric part of f_{12}. We have*

$$\sum_{r=1}^{R} (\lambda_r^{\mathsf{s}})^2 = \operatorname{Var}[f_{12\mathsf{s}}(X_1,X_2)], \tag{4.7}$$

$$\sum_{q=1}^{Q_+} (\lambda_q^{\mathsf{a}})^2 = \tfrac12 \operatorname{Var}[f_{12\mathsf{a}}(X_1,X_2)], \tag{4.8}$$

and

$$\operatorname{Var} W = \tfrac12 \sum_{r=1}^{R} (\lambda_r^{\mathsf{s}})^2 + \sum_{q=1}^{Q_+} (\lambda_q^{\mathsf{a}})^2 = \tfrac12 \operatorname{Var}[f_{12}(X_1,X_2)]. \tag{4.9}$$

Proof. We start by closely following the proof in Appendix A of the corresponding classical result for the usual U-statistics in Theorem 3.1.

Note that the cyclic U-statistic U_n° in (1.3) is a sum of $n\lfloor\frac{n}{2}\rfloor = \frac12 n^2 + O(n)$ terms. This shows (4.1).

We substitute the decomposition (2.10) into the definition (1.3). Note that each $i \in \mathbb{Z}_n$ occurs in $\lfloor\frac{n}{2}\rfloor$ terms in the double sum (1.3), and so does every value of $i+j \in \mathbb{Z}_n$. Hence,

$$U_n^\circ(f) = \mathbb{E}\left[U_n^\circ(f)\right] + \left\lfloor\frac{n}{2}\right\rfloor \sum_{i=1}^n f_1(X_i) + \left\lfloor\frac{n}{2}\right\rfloor \sum_{i=1}^n f_2(X_i) + \sum_{i\in\mathbb{Z}_n} \sum_{1\leqslant j<n/2} f_{12}(X_i, X_{i+j}). \tag{4.10}$$

Each of the first two sums is a sum of i.i.d. random variables with mean zero and finite variance; hence these sums are $O_{L^2}(n^{1/2})$, and the corresponding terms are $O_{L^2}(n^{3/2})$. Furthermore, the final double sum is a sum of $O(n^2)$ identically distributed terms that are uncorrelated by Lemma 2.1 and have mean zero, and thus the double sum is $O_{L^2}(n) = o_{L^2}(n^{3/2})$.

In the rest of the proof we replace f by $f-\mu$ (which does not change f_1, f_2, or f_{12}); hence we may and do assume without loss of generality that $\mu = 0$. Thus $\mathbb{E}\left[U_n^\circ(f)\right] = 0$ by (4.1).

(i) and (ii): By (4.10) and the comments after it,

$$n^{-3/2} U_n^\circ(f) = n^{-3/2}\frac{n}{2}\sum_{i=1}^n \bigl(f_1(X_i) + f_2(X_i)\bigr) + o_{L^2}(1). \tag{4.11}$$

Since the variables X_i are i.i.d., (4.11) implies immediately both (4.3) and, by the classical central limit theorem (together with the Cramér–Slutsky theorem [8, Theorem 5.11.4]), (4.4). Any of these implies (4.2), and (4.1) was proved above.

(iii): Since $f_1(X_i) + f_2(X_i) = 0$ a.s., the first two sums in (4.10) cancel. Hence, using also our simplifying assumption $\mu = 0$, we now have $U_n^\circ(f) = U_n^\circ(f_{12})$, and we may simplify the notation by assuming $f = f_{12}$. Then, Lemma 2.1 shows that (1.3) is a sum of $n\lfloor\frac{n}{2}\rfloor$ uncorrelated, identically distributed, terms, and (4.5) follows.

It will be convenient to consider even n, so we first note that for any $n \geqslant 1$, by (1.3) and some bookkeeping,

$$U_{2n+1}^\circ(f) = U_{2n}^\circ(f) + \sum_{i=1}^n f(X_i, X_{i+n}) + \sum_{i=1}^n f(X_{2n+1}, X_i) + \sum_{i=n+1}^{2n} f(X_i, X_{2n+1}). \tag{4.12}$$

(As a check, note that the total number of terms is $(2n+1)n$ in U_{2n+1}° and $2n(n-1)$ in U_{2n}°, and that every term in U_{2n}° appears also in U_{2n+1}°.) Each of the three sums in (4.12) is, by Lemma 2.1, a sum of n uncorrelated variables, and it follows that it is $O_{L^2}(n^{1/2})$. Hence it suffices to prove (4.6) for even n; the case of odd n then follows.

We consider thus U_{2n}°, where we assume $n \geqslant 2$. Define

$$\widetilde{X}_i := (X_i, X_{i+n}), \qquad i = 1, \dots, n. \tag{4.13}$$

Then $(\widetilde{X}_i)_1^n$ is an i.i.d. sequence of random variables in $\mathcal{X}^2$. Define the function F on $\mathcal{X}^4 = \mathcal{X}^2 \times \mathcal{X}^2$ by

$$F\bigl((x_1, x_2), (y_1, y_2)\bigr) := f(x_1, y_1) + f(y_1, x_2) + f(x_2, y_2) + f(y_2, x_1). \tag{4.14}$$

It now follows from the definitions (1.3) and (1.2) that

$$U^{\circ}_{2n}(f) = U_n(F; \widetilde{X}_1, \dots, \widetilde{X}_n). \tag{4.15}$$

Consequently, we may use the classical result Theorem 3.1 for the usual U-statistics, applied to F and $(\widetilde{X}_i)$. Recall that we have $f_1 = f_2 = 0$ and that we have assumed $\mu = 0$, which clearly implies also (using that $(X_i)_i$ are i.i.d.)

$$F_{\varnothing} = \mathbb{E}\, F(\widetilde{X}_1, \widetilde{X}_2) = \mathbb{E}\, F\big((X_1, X_2), (X_3, X_4)\big) = 0. \tag{4.16}$$

Furthermore, (2.12) applied to F yields, using (4.14),

$$\begin{aligned} F_1(x_1, x_2) &= \mathbb{E}\, F\big((x_1, x_2), (X_3, X_4)\big) \\ &= \mathbb{E}\, f(x_1, X_3) + \mathbb{E}\, f(X_3, x_2) + \mathbb{E}\, f(x_2, X_4) + \mathbb{E}\, f(X_4, x_1) \\ &= f_1(x_1) + f_2(x_2) + f_1(x_2) + f_2(x_1) = 0 \end{aligned} \tag{4.17}$$

and similarly $F_2(x_1, x_2) = 0$. (This follows also from (4.15) and (4.5), which show that $\operatorname{Var}[U_n(F)] = \operatorname{Var}[U^{\circ}_{2n}(f)] = O(n^2)$, together with (3.3).)

Hence, Theorem 3.1(iii) applies, and shows that (3.6) holds for $U_n(F)$; it remains to find the eigenvalues λ_r of $T_{\widehat{F}}$, where $T_{\widehat{F}}$ is the integral operator on $L^2(\mathcal{X}^2 \times [0,1])$ with kernel, by (2.6) and (4.14),

$$\begin{aligned} \widehat{F}\big((x_1, x_2, t), (y_1, y_2, u)\big) &= \big(f(x_1, y_1) + f(y_1, x_2) + f(x_2, y_2) + f(y_2, x_1)\big)\mathbf{1}\{t < u\} \\ &\quad + \big(f(y_1, x_1) + f(x_1, y_2) + f(y_2, x_2) + f(x_2, y_1)\big)\mathbf{1}\{t > u\} \\ &= \widehat{f}\big((x_1, t), (y_1, u)\big) + \widehat{f}\big((y_1, t), (x_2, u)\big) + \widehat{f}\big((x_2, t), (y_2, u)\big) + \widehat{f}\big((y_2, t), (x_1, u)\big) \\ &= \widehat{f}\big((x_1, t), (y_1, u)\big) + \widehat{f^*}\big((x_2, t), (y_1, u)\big) + \widehat{f}\big((x_2, t), (y_2, u)\big) + \widehat{f^*}\big((x_1, t), (y_2, u)\big). \end{aligned} \tag{4.18}$$

We pause for a general observation on this type of kernels. Let

$$L^2_0(\mathcal{X} \times [0,1]) := \Big\{h \in L^2(\mathcal{X} \times [0,1]) : \int_{\mathcal{X}} h(x,t)\, \mathrm{d}\nu(x) = 0 \text{ for a.e. } t \in [0,1]\Big\} \tag{4.19}$$

and let, for $j = 1, 2$, M_j be the subspace of $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$ consisting of functions of the type $g(x_1, x_2, t) = h(x_j, t)$ for some $h \in L^2_0(\mathcal{X} \times [0,1])$.

If $g_j \in M_j$ for $j = 1, 2$, then with obvious notation,

$$\begin{aligned} &\int_{\mathcal{X} \times \mathcal{X} \times [0,1]} g_1(x_1, x_2, t) g_2(x_1, x_2, t)\, \mathrm{d}\nu(x_1)\, \mathrm{d}\nu(x_2)\, \mathrm{d}t \\ &\qquad = \int_{\mathcal{X} \times \mathcal{X} \times [0,1]} h_1(x_1, t) h_2(x_2, t)\, \mathrm{d}\nu(x_1)\, \mathrm{d}\nu(x_2)\, \mathrm{d}t = 0. \end{aligned} \tag{4.20}$$

Thus M_1 and M_2 are orthogonal subspaces of $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$. Let $M_1 \oplus M_2$ by their direct sum; this is also a subspace of $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$.

Lemma 4.2. *Let $i, j \in \{1, 2\}$. Suppose that g is a function in $L^2\big((\mathcal{X} \times \mathcal{X} \times [0,1])^2\big)$ of the form $g\big((x_1, x_2, t), (y_1, y_2, u)\big) = h(x_i, y_j, t, u)$ where $\int_{\mathcal{X}} h(x, y, t, u)\, \mathrm{d}\nu(x) = 0$ for a.e. $y \in \mathcal{X}$ and $t, u \in [0,1]$. Then T_g maps $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$ into the subspace M_i. Furthermore, T_g maps M_{3-j} to 0.*

Proof. Simple consequences of the definitions and Fubini's theorem. $\square$

To continue the proof of Theorem 4.1, we see from (4.18) that $\widehat{F}$ is a sum of 8 terms, each of them of the type in Lemma 4.2. Hence, $T_{\widehat{F}}$ maps $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$ into $M_1 \oplus M_2$, and thus all eigenfunctions for a nonzero eigenvalue belong to $M_1 \oplus M_2$. Hence, to find the nonzero eigenvalues, it suffices to consider the restriction of $T_{\widehat{F}}$ to $M_1 \oplus M_2$.

Let $\psi_1 + \psi_2 \in M_1 \oplus M_2$, with $\psi_j \in M_j$, and let (with a minor abuse of notation) ψ_j denote also the corresponding function in $L^2_0 := L^2_0(\mathcal{X} \times [0,1])$. Then, by (4.18) and Lemma 4.2 (which also shows that some terms vanish),

$$T_{\widehat{F}}(\psi_1 + \psi_2)(x_1, x_2, t) = T_{\widehat{f}}\psi_1(x_1, t) + T_{\widehat{f^*}}\psi_1(x_2, t) + T_{\widehat{f}}\psi_2(x_2, t) + T_{\widehat{f^*}}\psi_2(x_1, t), \tag{4.21}$$

where the four terms on the right-hand side belong to M_1, M_2, M_2, M_1, respectively. Since $\psi_1 + \psi_2$ is an eigenfunction with eigenvalue λ if and only if the left-hand side of (4.21) equals $\lambda\psi_1(x_1, t) + \lambda\psi_2(x_2, t)$, it follows by separating both sides of (4.21) into their components in M_1 and M_2 (or, equivalently, by separating terms depending on x_1 from terms depending on x_2) that $\psi_1 + \psi_2$ is an eigenfunction with eigenvalue $\lambda \neq 0$ if and only if

$$\begin{cases} T_{\widehat{f}}\psi_1 + T_{\widehat{f^*}}\psi_2 = \lambda\psi_1, \\ T_{\widehat{f^*}}\psi_1 + T_{\widehat{f}}\psi_2 = \lambda\psi_2. \end{cases} \tag{4.22}$$

By adding and subtracting these equations, we obtain the equivalent system

$$\begin{cases} (T_{\widehat{f}} + T_{\widehat{f^*}})(\psi_1 + \psi_2) = \lambda(\psi_1 + \psi_2), \\ (T_{\widehat{f}} - T_{\widehat{f^*}})(\psi_1 - \psi_2) = \lambda(\psi_1 - \psi_2). \end{cases} \tag{4.23}$$

Let, for $0 \neq \lambda \in \mathbb{C}$ and an operator T on a vector space, $E_\lambda(T)$ denote the eigenspace $\{h : Th = \lambda h\}$. The map $\psi_1 + \psi_2 \mapsto (\psi_1 + \psi_2, \psi_1 - \psi_2)$ is a bijection of $M_1 \oplus M_2$ onto $L^2_0 \times L^2_0$, and (4.23) shows that this bijection maps the eigenspace $E_\lambda(T_{\widehat{F}})$ onto $E_\lambda(T_{\widehat{f}+\widehat{f^*}}) \oplus E_\lambda(T_{\widehat{f}-\widehat{f^*}})$. In particular, the dimensions agree, which shows that the multiset of nonzero eigenvalues of $T_{\widehat{F}}$ equals the union of the multisets of nonzero eigenvalues of $T_{\widehat{f}+\widehat{f^*}}$ and $T_{\widehat{f}-\widehat{f^*}}$. We analyze these separately.

First, recalling (1.11) and (2.5), $f + f^* = 2f_{\mathsf{s}}$ is symmetric, and thus, by (2.6), $\widehat{f + f^*}(x, y, t, u) = 2\widehat{f}_{\mathsf{s}}(x, y, t, u) = 2f_{\mathsf{s}}(x, y)$. Hence, the corresponding eigenvalues are 2 times the eigenvalues λ^{s}_r of $T_{f_{\mathsf{s}}}$. (Cf. Remark 3.2.) The contribution from the eigenvalues of $T_{\widehat{f}+\widehat{f^*}}$ to the limit (in distribution) (3.6) of $n^{-1}U_n(F)$ is thus

$$\sum_{r=1}^{R} \tfrac{1}{2}(2\lambda^{\mathsf{s}}_r)(\zeta_r^2 - 1) = \sum_{r=1}^{R} \lambda^{\mathsf{s}}_r(\zeta_r^2 - 1). \tag{4.24}$$

On the other hand, $f - f^* = 2f_{\mathsf{a}}$ is antisymmetric. Its eigenvalues on the positive imaginary axis are $(2\mathrm{i}\lambda^{\mathsf{a}}_q)_1^{Q_+}$, and thus it follows from Lemma A.2 that the contribution from the eigenvalues of $T_{\widehat{f}-\widehat{f^*}}$ to the limit (3.6) is

$$\sum_{q=1}^{Q_+} 2\lambda^{\mathsf{a}}_q\eta_q. \tag{4.25}$$

It follows from (3.6) also that the contributions in (4.24) amd (4.25) are independent.

Consequently, recalling (4.15), (3.6) for $U_n(F)$ implies that

$$\frac{1}{n}U^{\circ}_{2n}(f) = \frac{1}{n}U_n(F) \xrightarrow{\mathrm{d}} \sum_{r=1}^{R} \lambda^{\mathsf{s}}_r(\zeta_r^2 - 1) + 2\sum_{q=1}^{Q_+} \lambda^{\mathsf{a}}_q \eta_q. \tag{4.26}$$

This shows that $\frac{1}{2n}U^{\circ}_{2n}(f)$ has the limit W in (4.6). In other words, (4.6) holds for even n, which as said above implies the general case.

Finally, (4.7) follows from (A.28) applied to f_{s}, and (4.8) follows from (A.37) applied to f_{a}. The first equality in (4.9) follows from (4.6); the second follows from (4.7) and (4.8) since f_{s} and f_{a} are orthogonal. □

Remark 4.3. If f is symmetric, then we obtain the same asymptotic results for U°_n as in Theorem 3.1 for U_n. This is nothing new, since, as noted in Remark 1.1, in this case $U^{\circ}_n = U_n$ for odd n, and $U^{\circ}_n = U_n + O_{L^2}(n^{1/2})$ for even n.

On the other hand, if f is antisymmetric, then $f_2(x) = -f_1(x)$, and thus U°_n is always of the degenerate type, while U_n is nondegenerate unless $f_1(x) = 0$. In the latter case, when f is antisymmetric and $f = f_{12}$, we again find the same asymptotic results for U°_n and U_n, this time less obviously. △

Remark 4.4. Note that, rather surprisingly, (4.6) shows that in the degenerate case (iii), the contributions to $U^{\circ}_n(f)$ from the symmetric and antisymmetric parts of f decouple, so that W is a sum of two independent components. Equivalently, by the Cramér–Wold device and applying the theorem to $sf_{\mathsf{s}} + tf_{\mathsf{a}}$ for $s, t \in \mathbb{R}$, $n^{-1}U^{\circ}_n(f_{\mathsf{s}})$ and $n^{-1}U^{\circ}_n(f_{\mathsf{a}})$ converge jointly in distribution to the two independent sums in (4.6).

There is no such decoupling for the standard U-statistic U_n, or for any of the alternating U-statistics, as will be seen in Example 8.2. Hence, the decoupling for U°_n seems to be an effect of the larger (cyclical) symmetry of U°_n. △

5. Bi-alternating U-statistics

We next give the corresponding result for the bi-alternating U-statistic U^{--}. The result is similar to Theorems 3.1 and 4.1, but the alternating signs in the definition (1.10) lead to cancellations and as a result there is no case corresponding to the nondegenerate case for standard or cyclic U-statistics; the main case corresponds to the degenerate case in the previous theorems. There is also a rather uninteresting new case (v), included for completeness, with an even smaller variance $O(n)$ and $U^{--}_n(f)$ reduced to a sum of i.i.d. variables.

Theorem 5.1. *With notations and assumptions as in Section 2, the following holds for the bi-alternating U-statistic $U^{--}_n(f)$ in* (1.10).

(i) *We have*

$$\mathbb{E}\,[U^{--}_n(f)] = -\lfloor \tfrac{n}{2} \rfloor \mu = O(n), \tag{5.1}$$

and, as $n \to \infty$, we have the weak law of large numbers

$$\frac{1}{\binom{n}{2}} U^{--}_n(f) \xrightarrow{\mathrm{p}} 0. \tag{5.2}$$

(ii) *We have*

$$\operatorname{Var}[U^{--}_n(f)] = \tfrac{1}{2}n^2 \operatorname{Var}[f_{12}(X_1, X_2)] + O(n). \tag{5.3}$$

Moreover, there exists a finite or infinite sequence of real numbers $(\lambda_r)_1^R$ *such that*

$$n^{-1}\left(U_n^{--}(f) - \mathbb{E}\left[U_n^{--}(f)\right]\right) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \tfrac{1}{2}\lambda_r(\zeta_r^2 - 1) \tag{5.4}$$

where $(\zeta_r)_1^R$ *are independent standard normal variables. The coefficients* $(\lambda_r)_1^R$ *in* (5.4) *are the nonzero eigenvalues (with multiplicities) of the self-adjoint integral operator* $T_{\widehat{f_{12}}}$ *on* $L^2(\mathcal{X} \times [0,1], \nu \times \ell)$*. We have*

$$\operatorname{Var} W = \tfrac{1}{2}\sum_{r=1}^{R} \lambda_r^2 = \tfrac{1}{2}\operatorname{Var}[f_{12}(X_1, X_2)]. \tag{5.5}$$

(iii) *In the special case of* (ii) *where furthermore* f *is symmetric, the coefficients* $(\lambda_r)_1^R$ *in* (5.4) *are the nonzero eigenvalues (with multiplicities) of the self-adjoint integral operator* $T_{f_{12}}$ *on* $L^2(\mathcal{X}, \nu)$.

(iv) *In the special case of* (ii) *where furthermore* f *is antisymmetric, then also*

$$n^{-1}U_n^{--}(f) \xrightarrow{\mathrm{d}} W := \sum_{q=1}^{Q_+} \lambda_q^{\mathsf{a}} \eta_q, \tag{5.6}$$

where $(\eta_q)_1^{Q_+}$ *are independent random variables with the stochastic area distribution* (2.23)*, and the coefficients* $(\lambda_q^{\mathsf{a}})_1^{Q_+}$ *are the positive numbers such that the imaginary number* $\mathrm{i}\lambda_q^{\mathsf{a}}$ *is an eigenvalue of the anti-self-adjoint operator* $T_{f_{12}}$ *on* $L^2_{\mathbb{C}}(\mathcal{X}, \nu)$*. We have*

$$\operatorname{Var} W = \sum_{q=1}^{Q_+} (\lambda_q^{\mathsf{a}})^2 = \tfrac{1}{2}\operatorname{Var}[f_{12}(X_1, X_2)]. \tag{5.7}$$

(v) *If* $f_{12} = 0$*, then*

$$\operatorname{Var}[U_{2n}^{--}(f)] = 2n\sigma_{\mathsf{e}}^2 + O(1), \tag{5.8}$$

$$\operatorname{Var}[U_{2n+1}^{--}(f)] = (2n+1)\sigma_{\mathsf{o}}^2 + O(1), \tag{5.9}$$

where

$$\sigma_{\mathsf{e}}^2 := \tfrac{1}{2}\bigl(\operatorname{Var}[f_1(X)] + \operatorname{Var}[f_2(X)]\bigr), \tag{5.10}$$

$$\sigma_{\mathsf{o}}^2 := \tfrac{1}{2}\operatorname{Var}[f_1(X) + f_2(X)]. \tag{5.11}$$

Furthermore,

$$(2n)^{-1/2}\bigl(U_{2n}^{--}(f) - \mathbb{E}\left[U_{2n}^{--}(f)\right]\bigr) \xrightarrow{\mathrm{d}} N(0, \sigma_{\mathsf{e}}^2), \tag{5.12}$$

$$(2n+1)^{-1/2}\bigl(U_{2n+1}^{--}(f) - \mathbb{E}\left[U_{2n+1}^{--}(f)\right]\bigr) \xrightarrow{\mathrm{d}} N(0, \sigma_{\mathsf{o}}^2). \tag{5.13}$$

Proof. We follow the proofs of Theorems 3.1 and 4.1, with some differences. First, (5.1) follows immediately from (1.10) and

$$\sum_{1 \leqslant i < j \leqslant n} (-1)^{i+j} = \sum_{j=1}^{n} (-1)^j \sum_{i=1}^{j-1} (-1)^i = \sum_{j=1}^{n} (-1)^{j+1} \mathbf{1}\{j \text{ is even}\} = -\left\lfloor \frac{n}{2} \right\rfloor. \tag{5.14}$$

We substitute the decomposition (2.10) into the definition (1.10), and obtain by simple calculations

$$U_n^{--}(f) = \mathbb{E}\,[U_n^{--}(f)] - \sum_{i=1}^{n} \mathbf{1}\{n-i \text{ is odd}\} f_1(X_i) - \sum_{j=1}^{n} \mathbf{1}\{j \text{ is even}\} f_2(X_j) + \sum_{1\leqslant i<j\leqslant n} (-1)^{i+j} f_{12}(X_i, X_j). \tag{5.15}$$

Each of the first two sums is a sum of $\lfloor \frac{n}{2} \rfloor$ i.i.d. random variables with mean zero and finite variance; hence these sums are $O_{L^2}(n^{1/2})$. The final double sum is a sum of $\binom{n}{2}$ terms that are uncorrelated by Lemma 2.1 and have mean zero, and up to sign have the same distribution; hence the double sum has variance $\binom{n}{2} \operatorname{Var}[f_{12}(X_1, X_2)]$. It is also easily seen that the double sum is orthogonal to the two other sums, and (5.3) follows.

The weak law of large numbers (5.2) is a consequence of (5.1) and (5.3).

(ii): As in the other proofs, we replace f by $f - \mu$. In the rest of the proof we thus may assume that $\mu = 0$, and thus $\mathbb{E}\, U_n^{--} = 0$. By (5.15), then

$$U_n^{--}(f) = U_n^{--}(f_{12}) + O_{L^2}(n^{1/2}), \tag{5.16}$$

so it suffices to consider $U^{--}(f_{12})$, and we may, without loss of generality, for (notational) simplicity assume $f = f_{12}$.

Again it will be convenient to consider even n, and we note that (1.10) implies

$$U_{2n+1}^{--}(f) = U_{2n}^{--}(f) + \sum_{i=1}^{2n} (-1)^{i+1} f(X_i, X_{2n+1}). \tag{5.17}$$

The sum in (5.17) is, by Lemma 2.1 and our assumption $f = f_{12}$, a sum of $2n$ uncorrelated variables with means 0, and it follows that it is $O_{L^2}(n^{1/2})$. Hence it suffices to prove (5.4) for even n.

We consider thus U_{2n}^{--}, where we assume $n \geqslant 2$. We now define (cf. (4.13) for the cyclic U-statistic)

$$\widetilde{X}_i := (X_{2i-1}, X_{2i}), \qquad i = 1, \dots, n. \tag{5.18}$$

Again, $(\widetilde{X}_i)_1^n$ is an i.i.d. sequence of random variables in $\mathcal{X}^2$. We now define the function F on $\mathcal{X}^4 = \mathcal{X}^2 \times \mathcal{X}^2$ by

$$F\big((x_1, x_2), (y_1, y_2)\big) := f(x_1, y_1) - f(x_1, y_2) - f(x_2, y_1) + f(x_2, y_2), \tag{5.19}$$

and it follows from the definitions (1.10) and (1.2) that

$$U_{2n}^{--}(f) = U_n(F; \widetilde{X}_1, \dots, \widetilde{X}_n) - \sum_{i=1}^{n} f(X_{2i-1}, X_{2i}). \tag{5.20}$$

The final sum in (5.20) is $O_{L^2}(n^{1/2})$, as a sum of n i.i.d. variables with zero mean, and is thus negligible in (5.4). Consequently, we may use Theorem 3.1 for the usual U-statistics, applied to F and $(\widetilde{X}_i)$. Recall that we have assumed $f = f_{12}$ and thus $f_1 = f_2 = 0 = \mu$, which clearly implies also

$$F_\varnothing = \mathbb{E}\, F(\widetilde{X}_1, \widetilde{X}_2) = \mathbb{E}\, F\big((X_1, X_2), (X_3, X_4)\big) = 0. \tag{5.21}$$

Furthermore, (2.12)–(2.13) applied to F yield, similarly to (4.17),

$$F_1(x_1, x_2) = F_2(x_1, x_2) = 0. \tag{5.22}$$

(Again, this follows also from (5.20) and (5.3), which show that $\operatorname{Var}[U_n(F)] = O(n^2)$, together with (3.3).)

Hence, Theorem 3.1(iii) applies, and shows that (3.6) holds for $U_n(F)$. It remains to find the eigenvalues λ_r of $T_{\widehat{F}}$, where $T_{\widehat{F}}$ now is the integral operator on $L^2(\mathcal{X}^2 \times [0,1])$ with kernel, by (2.6) and (5.19),

$$\begin{aligned}\widehat{F}\bigl((x_1,x_2,t),(y_1,y_2,u)\bigr) &= \bigl(f(x_1,y_1) - f(x_1,y_2) - f(x_2,y_1) + f(x_2,y_2)\bigr)\mathbf{1}\{t<u\}\\ &\quad + \bigl(f(y_1,x_1) - f(y_1,x_2) - f(y_2,x_1) + f(y_2,x_2)\bigr)\mathbf{1}\{t>u\}\\ &= \widehat{f}\bigl((x_1,t),(y_1,u)\bigr) - \widehat{f}\bigl((x_1,t),(y_2,u)\bigr) - \widehat{f}\bigl((x_2,t),(y_1,u)\bigr) + \widehat{f}\bigl((x_2,t),(y_2,u)\bigr).\end{aligned} \tag{5.23}$$

It follows again from Lemma 4.2 that $T_{\widehat{F}}$ maps $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$ into $M_1 \oplus M_2$, and thus all eigenfunctions for a nonzero eigenvalue belong to $M_1 \oplus M_2$.

Let $\psi_1 + \psi_2 \in M_1 \oplus M_2$, with $\psi_j \in M_j$, and let again (with a minor abuse of notation) ψ_j denote also the corresponding function in $L^2_0 := L^2_0(\mathcal{X} \times [0,1])$. Then, by (5.23) and Lemma 4.2,

$$T_{\widehat{F}}(\psi_1+\psi_2)(x_1,x_2,t) = T_{\widehat{f}}\psi_1(x_1,t) - T_{\widehat{f}}\psi_2(x_1,t) - T_{\widehat{f}}\psi_1(x_2,t) + T_{\widehat{f}}\psi_2(x_2,t), \tag{5.24}$$

and it follows that $\psi_1 + \psi_2$ is an eigenfunction with eigenvalue $\lambda \neq 0$ if and only if

$$\begin{cases} T_{\widehat{f}}\psi_1 - T_{\widehat{f}}\psi_2 = \lambda\psi_1,\\ -T_{\widehat{f}}\psi_1 + T_{\widehat{f}}\psi_2 = \lambda\psi_2.\end{cases} \tag{5.25}$$

These equations imply $\lambda(\psi_1 + \psi_2) = 0$, and thus $\psi_2 = -\psi_1$; furthermore in this case the system simplifies to $2T_{\widehat{f}}\psi_1 = \lambda\psi_1$. Hence, if the nonzero eigenvalues of $T_{\widehat{f}}$ are $(\lambda_r)_1^R$, then the nonzero eigenvalues of $T_{\widehat{F}}$ are $(2\lambda_r)_1^R$.

Consequently, (5.20) and Theorem 3.1(iii) applied to F show that

$$\frac{1}{n}U^{--}_{2n}(f) = \frac{1}{n}U_n(F) + o_{L^2}(1) \xrightarrow{\mathrm{d}} \sum_{r=1}^{R} \lambda_r(\zeta_r^2 - 1). \tag{5.26}$$

This shows that $\frac{1}{2n}U^{--}_{2n}(f)$ has the limit W in (5.4). In other words, (5.4) holds for even n, and thus in the general case.

The first equality in (5.5) follows from (5.4); the second follows from (A.28) applied to $\widehat{f_{12}}$ together with (2.7) applied to f_{12}.

(iii) and (iv): These follow from (ii) as in the proof of Theorem 3.1. Alternatively, we may note that (ii) shows that $\frac{1}{n}U^{--}_n(f)$ has the same limit (in distribution) as $\frac{1}{n}U_n(f_{12})$, and thus (iii) and (iv) follow directly from the corresponding parts (iv) and (v) in Theorem 3.1.

(v): When $f_{12} = 0$ and $\mu = 0$, (5.15) simplifies to

$$U^{--}_n(f) = -\sum_{k=1}^{n/2} f_1(X_{2k-1}) - \sum_{k=1}^{n/2} f_2(X_{2k}) \tag{5.27}$$

when n is even, and

$$U^{--}_n(f) = -\sum_{k=1}^{(n-1)/2} \bigl(f_1(X_{2k}) + f_2(X_{2k})\bigr) \tag{5.28}$$

when n is odd. The summands in (5.27)–(5.28) are independent, and (5.8)–(5.9) follow directly; furthermore, (5.12)–(5.13) follow from the central limit theorem. □

Remark 5.2. As noted above, there is for the bi-alternating U-statistic U_n^{--} no case similar to the nondegenerate cases in Theorems 3.1 and 4.1 with a variance of order n^3. In fact, apart from (5.1) and (v), we can summarize Theorem 5.1 by saying that (as noted in the proof above), $\frac{1}{n}U_n^{--}(f-\mu)$ has the same asymptotic distribution as $\frac{1}{n}U_n(f_{12})$. However, these variables are not the same for finite n; in fact, they are asymptotically uncorrelated, as is easily seen using Lemma 2.1. Moreover, they converge to two *independent* copies of the same W; this may be seen by adapting the method in the proof to show that for any constants $s,t\in\mathbb{R}$,

$$s\cdot\frac{1}{n}U_n^{--}(f-\mu)+t\cdot\frac{1}{n}U_n(f_{12})\overset{\mathrm{d}}{\longrightarrow}s\sum_{r=1}^{R}\tfrac12\lambda_r(\zeta_r^2-1)+t\sum_{r=1}^{R}\tfrac12\lambda_r(\tilde\zeta_r^2-1), \tag{5.29}$$

where $\zeta_r,\tilde\zeta_r$ are independent standard normal variables. We omit the details. △

The following connection with U_n° was noted (and used) by [6] (in a special case).

Proposition 5.3. *If f is antisymmetric, and n is odd, then*

$$U_n^{\circ}(f)\overset{\mathrm{d}}{=}U_n^{--}(f). \tag{5.30}$$

Proof. More precisely, we show that

$$U_n^{\circ}(f;X_2,X_4,\dots,X_{2n})=U_n^{--}(X_1,X_2,\dots,X_n), \tag{5.31}$$

where the indices are interpreted modulo n as in Section 4; then (5.30) follows since $(X_i)_1^n$ are i.i.d. Note that since n is odd, $i\mapsto 2i$ is a bijection of the index set $\mathbb{Z}_n$ onto itself; hence the left-hand side of (5.31) contains all variables $X_1,\dots,X_n$, but in different order.

It follows from the definitions, and the assumption that f is antisymmetric, that both sides of (5.31) are sums containing $\binom{n}{2}$ term of the type $\pm f(X_i,X_j)$, one for each unordered pair $\{i,j\}$ with $i\neq j$. We only have to verify that the signs agree. On the left-hand side, we have one term $f(X_{2i},X_{2i+2j})$ for every $i\in[n]$ and $j\in[(n-1)/2]$, where $[n]:=\{1,\dots,n\}$. Letting $k\equiv 2i \pmod n$ and $l\equiv 2(i+j)\pmod n$ be the representatives with $k,l\in[n]$, then either $k<l$ and $l-k=2j$ is even, or $l<k$ and $k-l=n-2j$ is odd; conversely, every such pair (k,l) corresponds to a unique pair $(i,j)\in[n]\times[(n-1)/2]$. Since f is antisymmetric, a term $f(X_k,X_l)$ with $l<k$ equals $-f(X_l,X_k)$, and thus we see that the left-hand side of (5.30) contains $f(X_k,X_l)$ for $k<l$ with $k-l$ even, and (interchanging k and l), $-f(X_k,X_l)$ for $k<l$ with $k-l$ odd; this is the same as U_n^{--} in (1.10). □

Remark 5.4. Proposition 5.3 does not hold for even n, simply because U_n° and U_n^{--} then are sums of different numbers of terms and thus, even when $f=f_{12}$, they have in general different variances (using Lemma 2.1). However, for antisymmetric f, (5.30) holds approximatively with an error $O_{L^2}(n^{1/2})$ also for even n as a consequence of (4.12) and (5.17). Note also that (5.30) fails in general for symmetric f, even if $f=f_{12}$; for an example let $f(x,y):=(x-p)(y-p)$ and $X\in\mathrm{Be}(p)$ with $\frac12<p<1$. △

6. Singly alternating U-statistics

We turn to U_n^{-+} and U_n^{+-} in (1.8)–(1.9). We note first that by arguing as in (4.12) and (5.17), we see that

$$\begin{aligned} U_n^{-+}(f) &= \sum_{1\leqslant i<j\leqslant n} (-1)^{i+1} f(X_i, X_j) = \sum_{2\leqslant i<j\leqslant n+1} (-1)^{i+1} f(X_i, X_j) + O_{L^2}(n^{1/2}) \\ &\overset{\mathrm{d}}{=} -U_n^{-+}(f) + O_{L^2}(n^{1/2}). \end{aligned} \tag{6.1}$$

The error term will be negligible in most of our asymptotic results below; in particular, (6.1) implies that any distributional limit found for $n^{-3/2}U_n^{-+}$ or $n^{-1}U_n^{-+}$ has to be symmetric. The same holds for U_n^{+-} by the same argument, or by (1.12). (Note that U_n, U_n°, and U_n^{--} can have asymmetric asymptotic distributions, see Example 8.1.)

Theorem 6.1. *With notations and assumptions as in Section 2, the following holds.*

(i) *We have*

$$\mathbb{E}\,[U_n^{+-}(f)] = (-1)^n \left\lfloor \frac{n}{2} \right\rfloor \mu. \tag{6.2}$$

(ii) *As $n \to \infty$,*

$$n^{-3}\operatorname{Var}[U_n^{+-}(f)] \to \sigma^2 := \tfrac{1}{3}\,\mathbb{E}\,[f_2(X)^2] \tag{6.3}$$

and

$$n^{-3/2}\left(U_n^{+-}(f) - \mathbb{E}\,[U_n^{+-}(f)]\right) \xrightarrow{\mathrm{d}} N(0,\sigma^2). \tag{6.4}$$

(iii) *If $f_2(X) = 0$ a.s., and thus $\sigma^2 = 0$, then*

$$\operatorname{Var}[U_n^{+-}(f)] = \tfrac{1}{2} n^2 \operatorname{Var}[f_{12}(X_1, X_2)] + O(n). \tag{6.5}$$

Moreover, there exists a finite or infinite sequence of real numbers $(\lambda_r)_1^R$ such that

$$n^{-1}\left(U_n^{+-}(f) - \mathbb{E}\,[U_n^{+-}(f)]\right) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \tfrac{1}{2}\lambda_r(\zeta_r^2 - 1), \tag{6.6}$$

where $(\zeta_r)_1^R$ are independent standard normal variables. The coefficients $(\lambda_r)_1^R$ in (6.6) are the nonzero eigenvalues (with multiplicities) of the self-adjoint operator on $(L^2(\mathcal{X} \times [0,1], \nu \times \ell))^2$ given in block form by

$$\frac{1}{2}\begin{pmatrix} -T_{\widehat{f_{12}}} & T_{\widetilde{f_{12}}} \\ -T_{\widetilde{f_{12}}} & T_{\widehat{f_{12}}} \end{pmatrix}. \tag{6.7}$$

We have

$$\operatorname{Var} W = \tfrac{1}{2}\sum_{r=1}^{R} \lambda_r^2 = \tfrac{1}{2}\operatorname{Var}[f_{12}(X_1, X_2)] < \infty. \tag{6.8}$$

(iv) *In the special case of (iii) where furthermore f is symmetric, then also*

$$n^{-1}\left(U_n^{+-}(f) - \mathbb{E}\,[U_n^{+-}(f)]\right) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \lambda_r \vartheta_r, \tag{6.9}$$

where $(\vartheta_r)_1^R$ *are i.i.d. with the distribution* (2.25)*, and the coefficients* $(\lambda_r)_1^R$ *are the nonzero eigenvalues (with multiplicities) of the integral operator* T_f *on* $L^2(\mathcal{X},\nu)$*. We have*

$$\operatorname{Var} W = \tfrac{1}{2}\sum_{r=1}^{R}\lambda_r^2 = \tfrac{1}{2}\operatorname{Var}[f_{12}(X_1,X_2)]. \tag{6.10}$$

(v) *In the special case of* (iii) *where furthermore* f *is antisymmetric, then also*

$$n^{-1}U_n^{+-}(f) \xrightarrow{\mathrm{d}} W := \sum_{q=1}^{Q_+}\lambda_q^{\mathsf{a}}\eta_q, \tag{6.11}$$

where $(\eta_q)_1^{Q_+}$ *are independent random variables with the stochastic area distribution* (2.23)*, and the coefficients* $(\lambda_q^{\mathsf{a}})_1^{Q_+}$ *are the positive numbers such that the imaginary number* $\mathrm{i}\lambda_q^{\mathsf{a}}$ *is an eigenvalue of the anti-self-adjoint operator* T_f *on* $L^2_{\mathbb{C}}(\mathcal{X},\nu)$*. We have*

$$\operatorname{Var} W = \sum_{q=1}^{Q_+}(\lambda_q^{\mathsf{a}})^2 = \tfrac{1}{2}\operatorname{Var}[f_{12}(X_1,X_2)]. \tag{6.12}$$

(vi) *If* $f_2 = f_{12} = 0$*, then*

$$\operatorname{Var}[U_n^{+-}(f)] = n\sigma_1^2 + O(1) \tag{6.13}$$

where

$$\sigma_1^2 := \tfrac{1}{2}\operatorname{Var}[f_1(X)]. \tag{6.14}$$

Furthermore,

$$n^{-1/2}\bigl(U_n^{+-}(f) - \mathbb{E}\,[U_n^{+-}(f)]\bigr) \xrightarrow{\mathrm{d}} N(0,\sigma_1^2). \tag{6.15}$$

Proof. We follow the proofs of Theorems 4.1 and 5.1, again with some differences; we omit some details that are the same as above. First, (i) follows immediately from (1.9) and

$$\sum_{1\leqslant i<j\leqslant n}(-1)^j = \sum_{i=1}^{n}\sum_{j=i+1}^{n}(-1)^j = \sum_{i=1}^{n}(-1)^n\mathbf{1}\{n-i\text{ is odd}\} = (-1)^n\left\lfloor\frac{n}{2}\right\rfloor. \tag{6.16}$$

We substitute the decomposition (2.10) into the definition (1.9), and obtain by simple calculations

$$\begin{aligned} U_n^{+-}(f) = \mathbb{E}\,[U_n^{+-}(f)] + (-1)^n\sum_{i=1}^{n}\mathbf{1}\{n-i\text{ is odd}\}f_1(X_i) + \sum_{j=1}^{n}(-1)^j(j-1)f_2(X_j)\\ + \sum_{1\leqslant i<j\leqslant n}(-1)^j f_{12}(X_i,X_j). \end{aligned} \tag{6.17}$$

The first sum in (6.17) has variance $O(n)$, the second has variance $O(n^3)$ and the final (double) sum has variance $O(n^2)$.

As in the other proofs, we replace f by $f-\mu$. In the rest of the proof we thus may assume that $\mu = 0$, and thus $\mathbb{E}\,U_n^{+-} = 0$.

(ii): The first and third sums in (6.17) can be ignored, and the remaining second sum is a sum of independent variables. Hence, a simple calculation yields (6.3), and (6.4) follows by the central limit theorem.

(iii): We obtain (6.5) from (6.17). For (6.6), it follows from (6.17) that it suffices to consider $U^{+-}(f_{12})$, and thus we may, without loss of generality, for simplicity assume $f = f_{12}$. Again it will be convenient to consider even n, and we note that (since $f = f_{12}$) (1.9) implies

$$U^{+-}_{2n+1}(f) = U^{+-}_{2n}(f) - \sum_{i=1}^{2n} f(X_i, X_{2n+1}) = U^{+-}_{2n}(f) + O_{L^2}(n^{1/2}). \tag{6.18}$$

Hence it suffices to prove (6.6) for even n. We consider thus $U^{+-}_{2n}(f)$, where we assume $n \geqslant 2$. We use again the definition (5.18) of $\widetilde{X}$, so that $(\widetilde{X}_i)_1^n$ is an i.i.d. sequence of random variables in $\mathcal{X}^2$. We now define the function F on $\mathcal{X}^4 = \mathcal{X}^2 \times \mathcal{X}^2$ by

$$F\bigl((x_1,x_2),(y_1,y_2)\bigr) := -f(x_1,y_1) + f(x_1,y_2) - f(x_2,y_1) + f(x_2,y_2), \tag{6.19}$$

and it follows from the definitions (1.9) and (1.2) that

$$\begin{aligned} U^{+-}_{2n}(f) &= U_n(F;\widetilde{X}_1,\dots,\widetilde{X}_n) + \sum_{i=1}^{n} f(X_{2i-1},X_{2i}) \\ &= U_n(F;\widetilde{X}_1,\dots,\widetilde{X}_n) + O_{L^2}(n^{1/2}). \end{aligned} \tag{6.20}$$

Consequently, we may use Theorem 3.1 applied to F and $(\widetilde{X}_i)$; we have again $F_\varnothing = F_1 = F_2 = 0$, so Theorem 3.1(iii) applies.

It remains to find the eigenvalues λ_r of the integral operator $T_{\widehat{F}}$ on $L^2(\mathcal{X}^2 \times [0,1])$ which has kernel, by (6.19) and recalling both (2.6) and (2.8),

$$\begin{aligned} \widehat{F}\bigl((x_1,x_2,t),(y_1,y_2,u)\bigr) &= \bigl(-f(x_1,y_1) + f(x_1,y_2) - f(x_2,y_1) + f(x_2,y_2)\bigr)\mathbf{1}\{t<u\} \\ &\quad + \bigl(-f(y_1,x_1) + f(y_1,x_2) - f(y_2,x_1) + f(y_2,x_2)\bigr)\mathbf{1}\{t>u\} \\ &= -\widehat{f}\bigl((x_1,t),(y_1,u)\bigr) + \check{f}\bigl((x_1,t),(y_2,u)\bigr) - \check{f}\bigl((x_2,t),(y_1,u)\bigr) + \widehat{f}\bigl((x_2,t),(y_2,u)\bigr). \end{aligned} \tag{6.21}$$

It follows again from Lemma 4.2 that $T_{\widehat{F}}$ maps $L^2(\mathcal{X} \times \mathcal{X} \times [0,1])$ into $M_1 \oplus M_2$, and thus all eigenfunctions for a nonzero eigenvalue belong to $M_1 \oplus M_2$.

Let $\psi_1 + \psi_2 \in M_1 \oplus M_2$, with $\psi_j \in M_j$, and let again ψ_j denote also the corresponding function in $L^2_0(\mathcal{X} \times [0,1])$. Then, by (6.21) and Lemma 4.2,

$$T_{\widehat{F}}(\psi_1+\psi_2)(x_1,x_2,t) = -T_{\widehat{f}}\psi_1(x_1,t) + T_{\check{f}}\psi_2(x_1,t) - T_{\check{f}}\psi_1(x_2,t) + T_{\widehat{f}}\psi_2(x_2,t), \tag{6.22}$$

and it follows that $\psi_1+\psi_2$ is an eigenfunction with eigenvalue $\lambda \neq 0$ if and only if

$$\begin{cases} -T_{\widehat{f}}\psi_1 + T_{\check{f}}\psi_2 = \lambda\psi_1, \\ -T_{\check{f}}\psi_1 + T_{\widehat{f}}\psi_2 = \lambda\psi_2. \end{cases} \tag{6.23}$$

Hence the nonzero eigenvalues of $T_{\widehat{F}}$ are the nonzero eigenvalues of the operator $\begin{pmatrix} -T_{\widehat{f}} & T_{\check{f}} \\ -T_{\check{f}} & T_{\widehat{f}} \end{pmatrix}$ on $L^2_0(\widehat{\mathcal{X}}) \times L^2_0(\widehat{\mathcal{X}})$. These are the same as the nonzero eigenvalues on $L^2(\widehat{\mathcal{X}}) \times L^2(\widehat{\mathcal{X}})$, since both $T_{\widehat{f}}$ and $T_{\check{f}}$ map $L^2(\widehat{\mathcal{X}})$ into $L^2_0(\widehat{\mathcal{X}})$. (Cf. Lemma 4.2.) We denote these eigenvalues by $2\lambda_r$, as in the statement, and obtain by (6.20) and (3.6)

$$n^{-1}U^{+-}_{2n}(f) = n^{-1}U_n(F;\widetilde{X}_1,\dots,\widetilde{X}_n) + o_{\mathrm{p}}(1) \xrightarrow{\mathrm{d}} \sum_{r=1}^{R} \lambda_r(\zeta_r^2 - 1), \tag{6.24}$$

which proves (6.6) for even n, and thus, by (6.18), in general.

Regard $\{1,2\}$ as a measure space with mass 1 at each of the two points. Then, the operator in (6.7) can be regarded as the operator T_G on $L^2(\mathcal{X}\times[0,1]\times\{1,2\})$ with kernel G given by the block form

$$G := \tfrac12\begin{pmatrix} -\widehat{f} & \breve{f} \\ -\breve{f} & \widehat{f}\end{pmatrix}. \tag{6.25}$$

It follows that, using (2.7) and its counterpart for $\breve{f}$,

$$\int_{(\widehat{\mathcal{X}}\times\{1,2\})^2} |G|^2 = \frac{2}{4}\int_{\widehat{\mathcal{X}}^2}|\widehat{f}|^2 + \frac{2}{4}\int_{\widehat{\mathcal{X}}^2}|\breve{f}|^2 = \int_{\mathcal{X}^2}|f|^2. \tag{6.26}$$

Hence, (6.8) follows from (6.6) and (A.28).

(iv): Since f is symmetric and $f_2(X)=0$ a.s., we also have $f_1(X)=0$ a.s., and thus (since we assume $\mu=0$) $f=f_{12}$. Furthermore, the definitions (2.6) and (2.8) yield

$$\widehat{f}\big((x,t),(y,u)\big) = f(x,y), \tag{6.27}$$

$$\breve{f}\big((x,t),(y,u)\big) = f(x,y)\operatorname{sgn}(u-t). \tag{6.28}$$

Hence, if we define the symmetric function H_{s} on $([0,1]\times\{1,2\})^2$ by the block form

$$H_{\mathsf{s}}\big((t,\alpha),(u,\beta)\big) := \frac12\begin{pmatrix} -1 & \operatorname{sgn}(u-t) \\ -\operatorname{sgn}(u-t) & 1\end{pmatrix}, \qquad t,u\in[0,1];\alpha,\beta\in\{1,2\}, \tag{6.29}$$

then we can regard the kernel G in (6.25) as the tensor product $f\otimes H_{\mathsf{s}}$ in the natural way, and thus

$$T_G = T_f\otimes T_{H_{\mathsf{s}}}, \tag{6.30}$$

where both T_f and $T_{H_{\mathsf{s}}}$ are self-adjoint. It follows, by the same argument as in the proof of Lemma A.2 in a similar case, that if T_f has the nonzero eigenvalues $\{\lambda_r : r\in\mathcal{R}\}$ and H_{s} has the nonzero eigenvalues $\{\rho_s : s\in\mathcal{S}\}$, then T_G has the nonzero eigenvalues $\{\lambda_r\rho_s : r\in\mathcal{R}, s\in\mathcal{S}\}$. The eigenvalues ρ_s are given by Lemma 6.4 below. Hence, the limit in (6.6) is

$$\begin{aligned} W &= \sum_{r=1}^{R}\sum_{k=-\infty}^{\infty}\tfrac12\lambda_r\frac{2}{(2k-1)\pi}(\zeta_{r,k}^2-1) = \sum_{r=1}^{R}\lambda_r\sum_{k=-\infty}^{\infty}\frac{1}{(2k-1)\pi}(\zeta_{r,k}^2-1) \\ &=: \sum_{r=1}^{R}\lambda_r\vartheta_r, \end{aligned} \tag{6.31}$$

where ϑ_r are i.i.d. and by Lemma 2.5 have the distribution (2.25).

Finally, (6.10) follows from (6.9) and (2.26) together with (6.8).

(v): Since f is antisymmetric, we again see that $f_2(X)=0$ a.s. implies $f_1(X)=0$ a.s., and thus $f=f_{12}$. The definitions (2.6) and (2.8) now yield

$$\widehat{f}\big((x,t),(y,u)\big) = f(x,y)\operatorname{sgn}(u-t). \tag{6.32}$$

$$\breve{f}\big((x,t),(y,u)\big) = f(x,y). \tag{6.33}$$

Hence, if we now define the antisymmetric function H_{a} on $([0,1]\times\{1,2\})^2$ by the block form

$$H_{\mathsf{a}}\bigl((t,\alpha),(u,\beta)\bigr) := \frac{1}{2}\begin{pmatrix} -\operatorname{sgn}(u-t) & 1 \\ -1 & \operatorname{sgn}(u-t)\end{pmatrix}, \qquad t,u\in[0,1];\alpha,\beta\in\{1,2\} \tag{6.34}$$

then we have again (6.30), where now both T_f and $T_{H_{\mathsf{a}}}$ are anti-self-adjoint, and thus have imaginary eigenvalues. It follows that if T_f has the nonzero eigenvalues $\{\mathrm{i}\lambda'_q : q\in\mathcal{Q}\}$ and H_{a} has the nonzero eigenvalues $\{\mathrm{i}\rho_s : s\in\mathcal{S}\}$, then T_G has the nonzero eigenvalues $\{-\lambda'_q\rho_s : q\in\mathcal{Q}, s\in\mathcal{S}\}$. The eigenvalues $\mathrm{i}\rho_s$ are given by Lemma 6.5 below. Hence, the limit in (6.6) is, cf. (6.31),

$$W = \sum_{q\in\mathcal{Q}}\sum_{k=-\infty}^{\infty}\lambda'_q\frac{-1}{(2k-1)\pi}(\zeta_{r,k}^2-1) =: \sum_{q\in\mathcal{Q}}\lambda'_q\vartheta_q, \tag{6.35}$$

where $(\vartheta_q)_{q\in\mathcal{Q}}$ are i.i.d. and by Lemma 2.5 have the distribution (2.25). Furthermore, since f is real, the nonzero eigenvalues of T_f are symmetric with respect to the real axis, and are thus $(\pm\mathrm{i}\lambda^{\mathsf{a}}_q)_{q=1}^{Q_+}$. Hence, we can rewrite (6.35) as

$$W = \sum_{q=1}^{Q_+}(\lambda^{\mathsf{a}}_q\vartheta_q-\lambda^{\mathsf{a}}_q\vartheta'_q) = \sum_{q=1}^{Q_+}\lambda^{\mathsf{a}}_q(\vartheta_q-\vartheta'_q) = \sum_{q=1}^{Q_+}\lambda^{\mathsf{a}}_q\eta_q, \tag{6.36}$$

where all ϑ_q and ϑ'_q are i.i.d. with the distribution (2.25), and thus $\eta_q := \vartheta_q-\vartheta'_q$ are i.i.d. with the distribution (2.23) by (2.25), see also Lemma 2.5.

Finally, (6.12) follows from (6.11) and (2.24) together with (6.8).

(vi): Follows from (6.17) and the central limit theorem. ☐

Remark 6.2. As noted after (6.1), the limits in distribution in Theorem 6.1 have to be symmetric random variables. This is obvious in (ii) and (iv)–(vi), but in (iii), it implies that the set of eigenvalues $(\lambda_r)_1^R$ has to be symmetric, i.e., $(\lambda_r)_1^R$ equals $(-\lambda_r)_1^R$ up to order. This can also be seen from (6.7): the measure-preserving bijection $(\widehat{x},\widehat{y})\mapsto(\widehat{y},\widehat{x})$ of $\widehat{X}\times\widehat{X}$ onto itself induces a unitary equivalence of the operator (6.7) with its negative. As a consequence, the limit in (6.6) can also be written

$$W = \sum_{\lambda_r>0}\tfrac{1}{2}\lambda_r(\zeta_r^2-\tilde{\zeta}_r^2), \tag{6.37}$$

where $\zeta_r,\tilde{\zeta}_r$ are independent standard normal variables. △

Theorem 6.3. *With notations and assumptions as in Section 2, the following holds.*

(i) *We have*

$$\mathbb{E}\,U_n^{-+}(f) = \Bigl\lfloor\frac{n}{2}\Bigr\rfloor\mu. \tag{6.38}$$

(ii) *All conclusions of Theorem 6.1*(ii)–(vi) *hold also for* $U^{-+}(f)$ *instead of* $U^{+-}(f)$, *provided* f_2 *is replaced by* f_1 *and conversely.*

Proof. This follows from (1.12) and Theorem 6.1 (applied to f^*) together with the following observations. First, $f_1^*=f_2$, $f_2^*=f_1$, and $(f^*)_{12}=(f_{12})^*$. Secondly, the factor $(-1)^n$ in (1.12) does not matter in (ii), since the limits are symmetric, see Remark 6.2. Thirdly, the operator (6.7) and its counterpart for f_{12}^* are unitarily equivalent and thus have the same eigenvalues, with the unitary equivalence induced by the measure-preserving bijection $\bigl((x,t),(y,u)\bigr) \mapsto \bigl((y,1-u),(x,1-t)\bigr)$ of $\widehat{\mathcal{X}}^2$ onto itself, since

$$\widehat{f^*_{12}}\bigl((x,t),(y,u)\bigr) = \widehat{f_{12}}\bigl((y,1-u),(x,1-t)\bigr) \tag{6.39}$$

and similarly for $\widecheck{f^*_{12}}$ and $\widecheck{f_{12}}$. ☐

Lemma 6.4. *Let H_{s} be the symmetric function on* $([0,1]\times\{1,2\})^2$ *given by* (6.29). *Then the eigenvalues of $T_{H_{\mathsf{s}}}$, all simple, are*

$$\pm\frac{2}{(2k-1)\pi}, \qquad k=1,2,3,\dots. \tag{6.40}$$

As a sanity check we note that if the eigenvalues in (6.40) are enumerated $(\lambda_r)_1^\infty$, then

$$\sum_r \lambda_r^2 = \frac{4}{\pi^2}\cdot 2\sum_{k=1}^\infty \frac{1}{(2k-1)^2} = 1 = \int_{([0,1]\times\{1,2\})^2} |H_{\mathsf{s}}|^2 \tag{6.41}$$

since $|H_{\mathsf{s}}| = \frac12$ and $[0,1]\times\{1,2\}$ has measure 2; this agrees with (A.28).

Proof. An eigenfunction of $T_{H_{\mathsf{s}}}$, with eigenvalue λ, is a pair (φ_1,φ_2) of functions on $[0,1]$ such that

$$2\lambda\varphi_1(t) = -\int_0^1 \varphi_1(u)\,\mathrm{d}u - \int_0^t \varphi_2(u)\,\mathrm{d}u + \int_t^1 \varphi_2(u)\,\mathrm{d}u, \tag{6.42}$$

$$2\lambda\varphi_2(t) = \int_0^t \varphi_1(u)\,\mathrm{d}u - \int_t^1 \varphi_1(u)\,\mathrm{d}u + \int_0^1 \varphi_2(u)\,\mathrm{d}u. \tag{6.43}$$

Suppose $\lambda\neq 0$. It then follows, as in the proof of Lemma A.3, that φ_1 and φ_2 are continuously differentiable, and differentiation yields

$$2\lambda\varphi_1'(t) = -2\varphi_2(t), \tag{6.44}$$

$$2\lambda\varphi_2'(t) = 2\varphi_1(t). \tag{6.45}$$

Let $\omega := 1/\lambda$. Then the system (6.44)–(6.45) becomes

$$\varphi_1' = -\omega\varphi_2, \tag{6.46}$$

$$\varphi_2' = \omega\varphi_1. \tag{6.47}$$

It follows that $\varphi_1'' = -\omega^2\varphi_1$, and thus, for some constants a and b, using also (6.46) again,

$$\varphi_1(t) = a\cos(\omega t) + b\sin(\omega t), \tag{6.48}$$

$$\varphi_2(t) = a\sin(\omega t) - b\cos(\omega t), \tag{6.49}$$

Furthermore, taking $t=0$ in (6.42) and (6.43), and integrating using (6.44)–(6.45),

$$\begin{aligned}2\lambda\varphi_1(0) = 2\lambda\varphi_2(0) &= -\int_0^1 \varphi_1(t)\,\mathrm{d}t + \int_0^1 \varphi_2(t)\,\mathrm{d}t \\ &= -\lambda\bigl(\varphi_2(1)-\varphi_2(0)\bigr) - \lambda\bigl(\varphi_1(1)-\varphi_1(0)\bigr),\end{aligned} \tag{6.50}$$

which implies $\varphi_1(0)=\varphi_2(0)$ and thus $a=-b$, and then $\varphi_1(1)+\varphi_2(1)=0$ and thus, by adding (6.48) and (6.49),

$$\cos(\omega) = 0. \tag{6.51}$$

Hence, for some $k \in \mathbb{Z}$,

$$\omega = (k + \tfrac{1}{2})\pi. \tag{6.52}$$

Conversely, it follows that for each such ω, we obtain an eigenfunction with eigenvalue $\lambda = 1/\omega$ by (6.48)–(6.49) with $b = -a$. This shows that eigenvalues are (6.40); we see also that these eigenvalues are simple.

For completeness we note that 0 is not an eigenvalue, since (6.42)–(6.43) with $\lambda = 0$ imply that the right-hand sides do not depend on t, and thus $\varphi_1(t) = \varphi_2(t) = 0$ a.e. □

Lemma 6.5. *Let H_{a} be the antisymmetric function on $([0,1] \times \{1,2\})^2$ given by* (6.34)*. Then the eigenvalues of $T_{H_{\mathsf{a}}}$, all simple, are*

$$\pm\frac{2\mathrm{i}}{(2k-1)\pi}, \qquad k = 1, 2, 3, \dots. \tag{6.53}$$

Proof. We argue as in the proof of Lemma 6.4. An eigenfunction of $T_{H_{\mathsf{a}}}$, with eigenvalue λ, is now a pair (φ_1, φ_2) of functions on $[0,1]$ such that

$$2\lambda\varphi_1(t) = \int_0^t \varphi_1(u)\,\mathrm{d}u - \int_t^1 \varphi_1(u)\,\mathrm{d}u + \int_0^1 \varphi_2(u)\,\mathrm{d}u, \tag{6.54}$$

$$2\lambda\varphi_2(t) = -\int_0^1 \varphi_1(u)\,\mathrm{d}u - \int_0^t \varphi_2(u)\,\mathrm{d}u + \int_t^1 \varphi_2(u)\,\mathrm{d}u. \tag{6.55}$$

Suppose $\lambda \neq 0$. It then follows, that φ_1 and φ_2 are continuously differentiable, and differentiation yields

$$2\lambda\varphi_1'(t) = 2\varphi_1(t), \tag{6.56}$$

$$2\lambda\varphi_2'(t) = -2\varphi_2(t). \tag{6.57}$$

Let $\omega := -\mathrm{i}/\lambda$. Then the system (6.56)–(6.57) becomes

$$\varphi_1' = \mathrm{i}\omega\varphi_1, \tag{6.58}$$

$$\varphi_2' = -\mathrm{i}\omega\varphi_2. \tag{6.59}$$

Thus, for some constants a and b,

$$\varphi_1(t) = ae^{\mathrm{i}\omega t}, \tag{6.60}$$

$$\varphi_2(t) = be^{-\mathrm{i}\omega t}. \tag{6.61}$$

Furthermore, taking $t = 0$ in (6.54) and (6.55), and integrating using (6.56)–(6.57),

$$\begin{aligned} 2\lambda\varphi_1(0) = 2\lambda\varphi_2(0) &= -\int_0^1 \varphi_1(t)\,\mathrm{d}t + \int_0^1 \varphi_2(t)\,\mathrm{d}t \\ &= -\lambda\big(\varphi_1(1) - \varphi_1(0)\big) - \lambda\big(\varphi_2(1) - \varphi_2(0)\big), \end{aligned} \tag{6.62}$$

which implies first $\varphi_1(0) = \varphi_2(0)$ and thus $a = b$, and then $\varphi_1(1) + \varphi_2(1) = 0$ and thus, by adding (6.60) and (6.61),

$$2\cos(\omega) = e^{\mathrm{i}\omega} + e^{-\mathrm{i}\omega} = 0. \tag{6.63}$$

Hence, for some $k \in \mathbb{Z}$,

$$\omega = (k + \tfrac{1}{2})\pi. \tag{6.64}$$

Conversely, it follows that for each such ω, we obtain an eigenfunction with eigenvalue $\lambda = -\mathrm{i}/\omega$ by (6.60)–(6.61) with $b = a$. This shows that eigenvalues are (6.53); we see also that these eigenvalues are simple.

By the same argument as in the proof of Lemma 6.4, 0 is not an eigenvalue. $\square$

Remark 6.6. Lemmas 6.4 and 6.5 show that the self-adjoint operators T_{H_s} and $\mathrm{i}T_{H_\mathsf{a}}$ have the same eigenvalues, and thus are unitarily equivalent. Operators with the same eigenvalues appear also in Example 8.2 and Lemma A.3 below. A unitary equivalence between any two of these is given by mapping eigenfunctions to eigenfunctions with the same eigenvalue but, in spite of the simple explicit forms of the eigenfunctions found in the proofs (and the great similarties between the proofs above and below), we do not see a simple explicit form of the unitary equivalences except for the case of Lemma 6.4 and Example 8.2. $\triangle$

7. A short summary

Comparing Theorems 3.1, 4.1, 5.1, 6.1, and 6.3 we see strong similarities but also differences. The nondegenerate cases are similar, with variances of order n^3 and normal limits; the proofs show that in these case, the dominating terms are linear combinations of $f_1(X_i)$ and $f_2(X_i)$, but the details differ because the linear combinations that appear are different for the different U-statistics. For U^{-+} and U^{+-} this is due to partial cancellations caused by the alternating signs, and for U^{--} this cancellation is (almost) complete so that the nondegenerate case does not occur at all. The asymptotic variances in (3.3), (4.3), and (6.3) are in general different, but note that in the special case when f is antisymmetric, and thus $f_2 = -f_1$, and further $f_1 \neq 0$, it follows that $U_n(f)$ and $U_n^{+-}(f)$ (and $U_n^{-+}(f)$) have the same asymptotic variance and thus the same asymptotic distribution, while $U_n^{\circ}(f)$ is of the degenerate type, and has the same asymptotic distribution as $U_n^{--}(f)$.

For the degenerate cases the general pattern is again similar, but details differ in more subtle and nonobvious ways. To see this clearer, we collect in the corollaries below the results for the degenerate case when f is symmetric or antisymmetric; we further assume for simplicity $f = f_{12}$, i.e., $\mu = 0 = f_1 = f_2$, since the degenerate cases always reduce to this case. Note that the variables $f_{12}(X_i, X_j)$ are orthogonal by Lemma 2.1; hence, when $f = f_{12}$, alternating signs do not change the variance of the U-statistic and do not cause any cancellation, although they may affect the asymptotic distribution. We let $\zeta_r \in N(0,1)$, η_q, and ϑ_r be independent copies of the variables in Section 2.3.

Corollary 7.1. *Suppose that f is symmetric and that $f = f_{12}$. Let the nonzero eigenvalues of T_f be $(\lambda_r)_1^R$. Then*

$$n^{-1}U_n(f) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \tfrac{1}{2}\lambda_r(\zeta_r^2 - 1). \tag{7.1}$$

The same result holds for U_n° and U_n^{--}. Furthermore,

$$n^{-1}U_n^{+-}(f) \xrightarrow{\mathrm{d}} W := \sum_{r=1}^{R} \lambda_r \vartheta_r. \tag{7.2}$$

The same result holds for U_n^{-+}.

The limit distributions in (7.1) and (7.2) are different by Remark 2.4 (unless f is constant and thus both limits are 0).

Corollary 7.2. *Suppose that f is symmetric and that $f = f_{12}$. Let the nonzero eigenvalues of T_f with positive imaginary part be $(\mathrm{i}\lambda^{\mathsf{a}}_q)_1^{Q_+}$. Then*

$$n^{-1}U_n(f) \xrightarrow{\mathrm{d}} W := \sum_{q=1}^{Q_+} \lambda^{\mathsf{a}}_q \eta_q. \tag{7.3}$$

The same result holds for U°_n, U^{--}_n, U^{+-}_n, and U^{-+}_n.

Recall from Remark 5.2 that the fact that two of our U-statistics have the same limit distribution does not imply that they converge jointly to the same random variable. On the other hand, Remark 1.1 shows that this happens in the case of $U_n(f)$ and $U^{\circ}_n(f)$ for symmetric f. See further Section 9.1.

Consider now the general case $f = f_{12}$, without any symmetry assumption. In this case, we have seen in Theorem 4.1(iii) and Remark 4.4 that the contributions to $U^{\circ}_n(f)$ from the symmetric and antisymmetric parts are asymptotically independent; hence the asymptotic distribution in the general case follows from the special cases in Corollaries 7.1 and 7.2. However, as is shown in Example 8.2, this does not hold for U_n, U^{--}_n, U^{+-}_n, or U^{-+}_n.

8. Examples

Examples with nondegenerate, and thus normal, limits are straightforward, so we concentrate on limits of the more complicated degenerate type. Again, we let $\zeta \in N(0,1)$, η, and ϑ, with or without subscripts, be independent copies of the variables in Section 2.3.

Example 8.1. Consider first the simple example where $f(x,y) = xy$ and X is real-valued with finite variance $\sigma^2_X > 0$. Let $\mu_X := \mathbb{E}\,X$. Then the Hoeffding decomposition (2.10)–(2.14) is given by $\mu = \mu_X^2$, $f_1(x) = f_2(x) = \mu_X(x - \mu_X)$, and $f_{12}(x,y) = (x-\mu_X)(y-\mu_X)$. Hence, if $\mu_X \neq 0$, we have the nondegenerate case with variance of order n^3 and normal limits for U_n, U°_n, U^{-+}_n, and U^{+-}_n. Recall that U^{--}_n never has this behaviour; in this example Theorem 5.1(ii) shows that $U^{--}_n(f)$ has the same asymptotic distribution as $U^{--}_n(f_{12})$, which is equivalent to replacing X by $X - \mu_X$.

Suppose now that $\mu_X = 0$. The integral operator T_f is

$$T_f g(x) = x \int y\, g(y)\,\mathrm{d}\nu(y) = x\,\mathbb{E}\,[Xg(X)]. \tag{8.1}$$

This operator has a single nonzero eigenvalue σ^2_X with eigenvector $\varphi(x) = x$; hence Theorem 3.1 yields

$$n^{-1}U_n \xrightarrow{\mathrm{d}} \tfrac{1}{2}\sigma^2_X(\zeta^2 - 1), \tag{8.2}$$

where $\zeta \in N(0,1)$. Theorems 4.1 and 5.1 yield the same result for U°_n and U^{--}_n (cf. Remarks 1.1 and 5.2), while Theorems 6.1(iv) and 6.3 yield

$$n^{-1}U^{+-}_n \xrightarrow{\mathrm{d}} \sigma^2_X\vartheta \tag{8.3}$$

and the same for U^{-+}_n, with ϑ as in (2.25). These results also follow from Corollary 7.1. △

Example 8.2. Let $X = (\xi_1, \xi_2)$ be a random vector in $\mathcal{X} = \mathbb{R}^2$, with ξ_1 and ξ_2 independent and both having variance 1 and symmetric distributions, i.e., $\xi_j \overset{\mathrm{d}}{=} -\xi_j$. (For example, ξ_1 and ξ_2 may both be $N(0,1)$, or uniformly distributed on ± 1.) Let

$$f\bigl((x_1, x_2), (x_1', x_2')\bigr) := x_1 x_2'. \tag{8.4}$$

Then, writing $X_n = (\xi_{n1}, \xi_{n2})$,

$$U_n(f) = \sum_{1 \leqslant i < j \leqslant n} \xi_{i1}\xi_{j2} = \sum_{j=1}^{n} \xi_{j2} \sum_{i=1}^{j} \xi_{i1}. \tag{8.5}$$

Note that $U_n^{+-}(f)$ is given by the same sum with ξ_{j2} replaced by $(-1)^j \xi_{j2}$. Since we assume $(-1)^j \xi_{j2} \overset{\mathrm{d}}{=} \xi_{j2}$, and all ξ_{ik} are independent, it follows that $U_n^{+-}(f)$ has the same distribution as $U_n(f)$. The same argument applies also to $U_n^{-+}(f)$, now replacing ξ_{i1} by $(-1)^i \xi_{i1}$, and to $U_n^{--}(f)$ (doing both). Thus, for any $n \geqslant 1$,

$$U_n(f) \overset{\mathrm{d}}{=} U_n^{--}(f) \overset{\mathrm{d}}{=} U_n^{+-}(f) \overset{\mathrm{d}}{=} U_n^{-+}(f). \tag{8.6}$$

It is shown in [14, p. 83] that

$$n^{-1} U_n(f) \xrightarrow{\mathrm{d}} \int_0^1 B_1(t)\,\mathrm{d}B_2(t) =: \vartheta, \tag{8.7}$$

where $B_k(t)$ are independent Brownian motions and thus ϑ is as in Section 2.3. (This is a consequence of (8.5) and Donsker's theorem applied to $\sum_{i=1}^n \xi_{i1}$ and $\sum_{i=1}^n \xi_{i2}$; see [14] for the nontrivial technical details.) By (8.6), we have the same asymptotic distribution (8.7) for $U_n^{--}(f)$, $U_n^{+-}(f)$, and $U_n^{-+}(f)$.

We can also obtain this limit from Theorem 3.1(iii) (or Theorem 5.1(ii)) above; note that $\mu = f_1 = f_2 = 0$ so $f = f_{12}$. It follows from the definitions (2.6) and (2.3) that $T_{\widehat{f}}$ is the integral operator on $L^2(\mathbb{R}^2 \times [0,1])$ given by

$$\begin{aligned} T_{\widehat{f}}\varphi(x_1, x_2, t) &= \int_t^1 \mathbb{E}\bigl[x_1 \xi_2 \varphi(\xi_1, \xi_2, u)\bigr]\,\mathrm{d}u + \int_0^t \mathbb{E}\bigl[\xi_1 x_2 \varphi(\xi_1, \xi_2, u)\bigr]\,\mathrm{d}u \\ &= x_1 \int_t^1 \mathbb{E}\bigl[\xi_2 \varphi(\xi_1, \xi_2, u)\bigr]\,\mathrm{d}u + x_2 \int_0^t \mathbb{E}\bigl[\xi_1 \varphi(\xi_1, \xi_2, u)\bigr]\,\mathrm{d}u. \end{aligned} \tag{8.8}$$

Consequently, an eigenfunction with a nonzero eigenvalue λ has to be of the form $\varphi(x_1, x_2, t) = x_1 \psi_1(t) + x_2 \psi_2(t)$. Substitution in (8.8) then yields

$$\lambda x_1 \psi_1(t) + \lambda x_2 \psi_2(t) = x_1 \int_t^1 \psi_2(u)\,\mathrm{d}u + x_2 \int_0^t \psi_1(u)\,\mathrm{d}u \tag{8.9}$$

and thus

$$\lambda \psi_1(t) = \int_t^1 \psi_2(u)\,\mathrm{d}u, \qquad \lambda \psi_2(t) = \int_0^t \psi_1(u)\,\mathrm{d}u. \tag{8.10}$$

Consequently,

$$\lambda \psi_1'(t) = -\psi_2(t), \qquad \lambda \psi_2'(t) = \psi_1(t). \tag{8.11}$$

It is easily seen that (8.11), with the boundary values $\psi_1(1) = \psi_2(0) = 0$ given by (8.10), is solved by, with $\omega := 1/\lambda$,

$$\psi_1(t) = C \cos(\omega t), \qquad \psi_2(t) = C \sin(\omega t), \tag{8.12}$$

where we must have

$$\cos(\omega) = 0. \tag{8.13}$$

Hence $\omega = (k + \frac{1}{2})\pi$, $k \in \mathbb{Z}$. The nonzero eigenvalues $\lambda = 1/\omega$ are thus $\{\frac{2}{(2k+1)\pi} : k \in \mathbb{Z}\}$, and (3.6) yields

$$n^{-1}U_n(f) \xrightarrow{\mathrm{d}} \sum_{k=-\infty}^{\infty} \frac{1}{(2k+1)\pi}(\zeta_k^2 - 1), \tag{8.14}$$

which by Lemma 2.5 has the same distribution as ϑ in (2.25), which proves (8.7). As noted above, we have the same limit (8.7) also for $U_n^{+-}(f)$ and $U_n^{-+}(f)$. In principle, this can be shown as above from Theorem 6.1, but that would require studying the more complicated integral operator (6.7).

Consider now the symmetric and antisymmetric parts; for convenience we consider

$$2f_{\mathsf{s}}\big((x_1, x_2), (x_1', x_2')\big) = x_1x_2' + x_2x_1', \tag{8.15}$$

$$2f_{\mathsf{a}}\big((x_1, x_2), (x_1', x_2')\big) = x_1x_2' - x_2x_1'. \tag{8.16}$$

Arguing as after (8.8) above, we see that it suffices to consider the subspace of linear functions $\{ax_1 + by_1 : a, b \in \mathbb{C}\} \subset L^2_{\mathbb{C}}(\mathcal{X}, \nu)$. This subspace is two-dimensional, and it is easy to see, in analogy to (8.8) but simpler, that in this subspace, $T_{2f_{\mathsf{s}}}$ and $T_{2f_{\mathsf{a}}}$ act by the matrices $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$; the nonzero eigenvalues are thus ± 1 and $\pm\mathrm{i}$, respectively. Consequently, Theorem 3.1(iv) and (v) yield

$$2n^{-1}U_n(f_{\mathsf{s}}) = n^{-1}U_n(2f_{\mathsf{s}}) \xrightarrow{\mathrm{d}} \tfrac{1}{2}(\zeta_1^2 - \zeta_2^2), \tag{8.17}$$

$$2n^{-1}U_n(f_{\mathsf{a}}) = n^{-1}U_n(2f_{\mathsf{a}}) \xrightarrow{\mathrm{d}} \eta. \tag{8.18}$$

This was also shown in [14, p. 83], representing the limits as $\int_0^1 B_1(t)\,\mathrm{d}B_2(t) \pm \int_0^1 B_2(t)\,\mathrm{d}B_1(t)$ in analogy to (8.7) above. See also (2.27), which implies that if we denote these limits in (8.17)–(8.18) by W_{s} and W_{a}, then, as noted in [14], their joint characteristic function is

$$\mathbb{E}\left[\exp(\mathrm{i}sW_{\mathsf{s}} + \mathrm{i}tW_{\mathsf{a}})\right] = \Big(\cosh^2(t) + s^2\frac{\sinh^2(t)}{t^2}\Big)^{-1/2}. \tag{8.19}$$

In particular, (8.17)–(8.18) hold jointly, but the limits W_{s} and W_{a} are *not* independent.

By Corollary 7.1, (8.17) holds also for $U_n^{\circ}(f_{\mathsf{s}})$ and $U_n^{--}(f_{\mathsf{s}})$, while, by (7.2),

$$2n^{-1}U_n^{+-}(f_{\mathsf{s}}) \xrightarrow{\mathrm{d}} \vartheta_1 - \vartheta_2 \overset{\mathrm{d}}{=} \eta \tag{8.20}$$

and the same for $U_n^{-+}(f_{\mathsf{s}})$. Similarly, by Corollary 7.2, (8.18) holds also for $U_n^{\circ}(f_{\mathsf{a}})$, $U_n^{--}(f_{\mathsf{a}})$, $U_n^{+-}(f_{\mathsf{a}})$, and $U_n^{-+}(f_{\mathsf{a}})$. Note that (8.17) and (8.20) show that $U_n(f_{\mathsf{s}})$ and $U_n^{+-}(f_{\mathsf{s}})$ have different limits in distribution; in particular, (8.6) cannot be extended to f_{s}.

We have so far ignored $U_n^{\circ}(f)$, but armed with these results for f_{s} and f_{a}, we obtain from (4.6) and Lemma 2.5

$$\begin{aligned} n^{-1}U_n^{\circ}(f) &\xrightarrow{\mathrm{d}} \tfrac{1}{4}(\zeta_1^2 - \zeta_2^2) + \tfrac{1}{2}\eta \\ &\overset{\mathrm{d}}{=} \tfrac{1}{4}(\zeta_1^2 - \zeta_2^2) + \sum_{k=1}^{\infty} \frac{1}{2(2k-1)\pi}\big(\zeta_{k,1}^2 + \zeta_{k,2}^2 - \zeta_{k,3}^2 - \zeta_{k,4}^2\big) \end{aligned} \tag{8.21}$$

This differs, by (2.28) and the uniqueness assertion in Remark 2.4, from the limit ϑ found in (8.6)–(8.7) for U_n, U_n^{--}, U_n^{+-}, and U_n^{-+}. (Note that (8.21) contains some coefficients that are rational, and some that are rational multiples of $1/\pi$.) It follows similarly that the decoupling of the contributions from the symmetric and antisymmetric parts that is seen in Theorem 4.1(iii) is unique to U_n°, and does not hold for the other U-statistics considered here. $\triangle$

KOLLA!

Remark 8.3. The formula (8.19) for the joint asymptotic distribution of $U_n(f_{\mathsf{s}})$ and $U_n(f_{\mathsf{a}})$ can also be obtained from Theorem 3.1 applied to $sf_{\mathsf{s}}+tf_{\mathsf{a}}$, see Appendix C. $U_n^{--}(f_{\mathsf{s}})$ and $U_n^{--}(f_{\mathsf{a}})$ have the same joint asymptotic distribution, by the same sign-change argument as for (8.6), but note that this argument does not apply to $U_n^{+-}(f_{\mathsf{s}})$ and $U_n^{+-}(f_{\mathsf{a}})$, as is shown by (8.20). Theorem 6.1 applied to $sf_{\mathsf{s}}+tf_{\mathsf{a}}$ shows that $U_n^{+-}(f_{\mathsf{s}})$ and $U_n^{+-}(f_{\mathsf{a}})$ have a joint asymptotic distribution, which in principle can be found by arguments similar to Appendix C (but for the more complicated operator (6.7)); we have not pursued this and leave it as an open problem $\triangle$

Remark 8.4. In this paper, we generally assume that $X_1,\dots,X_n$ are i.i.d. random variables. However, the definitions (1.1)–(1.3) and (1.7)–(1.10) make sense for any deterministic or random sequence $X_1,\dots,X_n$. One interesting instance of this is to let $\boldsymbol{\sigma}=(X_1,\dots,X_n)$ be a permutation of $\{1,\dots,n\}$. If further $f(x,y):=\mathbf{1}\{x>y\}$; then $U_n(f;\boldsymbol{\sigma})$ is the number of inversions in $\boldsymbol{\sigma}$; we will in the following two examples consider the equivalent (and more symmetric)

$$f(x,y):=\operatorname{sgn}(x-y)=2\mathbf{1}\{x>y\}-1. \tag{8.22}$$

We furthermore take $\boldsymbol{\sigma}$ to be a uniformly random permutation in the symmetric group $\mathfrak{S}_n$. It is well-known that $\boldsymbol{\sigma}$ can be constructed as the ranks of a sequence $X_1,\dots,X_n$ of i.i.d. random variables with, say, a uniform distribution on $[0,1]$. Since f only cares about the order relations, it follows that then $U_n(f;\boldsymbol{\sigma})=U_n(f;X_1,\dots,X_n)$, and similary for the other U-statistics; hence we are back to the case of i.i.d. X_i. $\triangle$

Example 8.5 (writhe)**.** Even-Zohar [6] defines the *writhe* of a permutation $\boldsymbol{\sigma}\in\mathfrak{S}_{2n+1}$ as, in our notation in Remark 8.4, $U_{2n+1}^{\circ}(f;\boldsymbol{\sigma})$, where $f(x,y):=\operatorname{sgn}(x-y)$ as in (8.22), and studied this in the case of a uniformly random permutation $\boldsymbol{\sigma}\in\mathfrak{S}_{2n+1}$. This was motivated by the study of a model for random knots; see [6] for details. (Only permutations of odd lengths appear in this model.)

The main result of [6] finds the asymptotic distribution of the writhe as $n\to\infty$; this is proved using the method of moments, together with a lengthy (but interesting) combinatorial calculation of the moments. (The proof actually uses the equivalence with U_{2n+1}^{--} in Proposition 5.3, which was given in [6] for this case; the moment calculations there are done for U_{2n+1}^{--}.)

As said in Remark 8.4, the distribution of the writhe equals the distribution of $U_{2n+1}^{\circ}(f;X_1,\dots,X_{2n+1})$ where $X_i\in U(0,1)$ are i.i.d. Consequently, we may apply Theorem 4.1. The kernel (8.22) is alternating, and thus $\mu=0$. Furthermore,

$$f_1(x)=\mathbb{E}\left[\operatorname{sgn}(x-X)\right]=2x-1, \tag{8.23}$$

$$f_2(x)=-f_1(x)=1-2x, \tag{8.24}$$

$$f_{12}(x,y)=\operatorname{sgn}(x-y)-2x+2y, \tag{8.25}$$

and we find $\operatorname{Var}[f_{12}(X_1,X_2)] = \frac{1}{3}$. We apply Theorem 4.1(iii), with $f_{12\mathsf{s}} = 0$ and $f_{12\mathsf{a}} = f_{12}$, and it remains to find the eigenvalues of $T_{f_{12\mathsf{a}}} = T_{f_{12}}$. Note that the eigenvalues of T_f are given in Lemma A.3, but here we consider $T_{f_{12}}$.

Suppose that φ is an eigenfunction of $T_{f_{12}}$ with eigenvalue $\mathrm{i}\lambda$. (All eigenvalues are imaginary, since f_{12} is antisymmetric.) Then

$$\mathrm{i}\lambda\varphi(x) = \int_0^x \varphi(y)\,\mathrm{d}y - \int_x^1 \varphi(y)\,\mathrm{d}y - 2x\int_0^1 \varphi(y)\,\mathrm{d}y + 2\int_0^1 y\varphi(y)\,\mathrm{d}y. \tag{8.26}$$

Suppose that $\lambda \neq 0$. It follows as in the proof of Lemma A.3 that φ is continuously differentiable, and then

$$\mathrm{i}\lambda\varphi'(x) = 2\varphi(x) - 2\int_0^1 \varphi(y)\,\mathrm{d}y. \tag{8.27}$$

Let $\omega := -2/\lambda$. Then (8.27) implies $\varphi''(x) = \mathrm{i}\omega\varphi'(x)$, and thus

$$\varphi(x) = Ae^{\mathrm{i}\omega x} + B \tag{8.28}$$

for some constants A and B. Simple calculus shows that then (8.26) holds if and only if $B = 0$ and

$$e^{\mathrm{i}\omega} = 1, \tag{8.29}$$

and thus

$$\omega = 2\pi k, \qquad k \in \mathbb{Z}. \tag{8.30}$$

Consequently, if $\omega \neq 0$ is as in (8.30), then $e^{\mathrm{i}\omega x}$ is an eigenfunction with eigenvalue $\mathrm{i}\lambda = -2\mathrm{i}/\omega = -\mathrm{i}/(\pi k)$, and the nonzero eigenvalues of $T_{f_{12}}$ are $\{\frac{\mathrm{i}}{\pi k} : 0 \neq k \in \mathbb{Z}\}$. In the notation of Theorem 4.1 we thus have $(\lambda_q^{\mathsf{a}})_1^{Q_+} = (\frac{1}{\pi k})_1^{\infty}$, and thus (4.6) yields

$$\frac{1}{n}U_n^{\circ}(f) \xrightarrow{\mathrm{d}} \sum_{k=1}^{\infty} \frac{1}{\pi k}\eta_k, \tag{8.31}$$

where η_k are i.i.d. with the stochastic area distribution (2.23). This is equivalent to the limit theorem in [6, Corollary 2] (there proved by the method of moments) with the limit represented as in [6, Section 5.3]. For properties and other descriptions of the limit, see [6]. △

Example 8.6 (alternating inversion number)**.** We continue to consider uniformly random permutations in $\mathfrak{S}_n$ as in Remark 8.4 and Example 8.5, using the kernel f in (8.22). Let $X_1,\dots,X_n$ be i.i.d. with $X_i \in U[0,1]$ as in Remark 8.4. Then, as noted above, $U_n(f;\boldsymbol{\sigma}) = U_n(f;X_1,\dots,X_n)$ is, up to a trivial linear transformation, the classical inversion number, and $U_n^{\circ}(f;\boldsymbol{\sigma}) = U_n^{\circ}(f;X_1,\dots,X_n)$ is the writhe studied in [6]. Furthermore, [6] also defines the *alternating inversion number* as $U_n^{-+}(f;\boldsymbol{\sigma})$ and the *bi-alternating inversion number* as $U_n^{--}(f;\boldsymbol{\sigma})$. The alternating inversion number $U_n^{-+}(f;\boldsymbol{\sigma})$ was (up to the same trivial linear transformation) earlier introduced and studied by [2], who showed that it has the same distribution as the "inversion number" $U_n(f;\boldsymbol{\sigma})$. (This is easily seen by regarding the random permutation $\boldsymbol{\sigma}$ as a random linear order on $[n]$, and constructing it recursively, by inserting a new element n in a random position relative to the previous $n-1$ elements; the number of new inversions is uniformly random in $\{0,\dots,n-1\}$, and it follows by induction that $U_n(f;\boldsymbol{\sigma})$ and $U_n^{+-}(f;\boldsymbol{\sigma})$ have the same distribution. And so has $U_n^{-+}(f;\boldsymbol{\sigma})$ by (1.12) since these variables are symmetric.)

As a consequence, $U_n^{+-}(f;\boldsymbol{\sigma})$ and $U_n^{-+}(f;\boldsymbol{\sigma})$ have the same asymptotic normal distribution as $U_n(f;\boldsymbol{\sigma})$. From the perspective of our theorems, this follows because f is antisymmetric, as noted in Section 7.

In contrast, as shown by [6], the bi-alternating inversion number $U_n^{--}(f;\boldsymbol{\sigma})$ has a different limit distribution, of the degenerate type. As noted above, [6] showed that $U_{2n+1}^{--}(f;\boldsymbol{\sigma}) \overset{\mathrm{d}}{=} U_{2n+1}^{\circ}(f;\boldsymbol{\sigma})$, which generalizes to Proposition 5.3; [6] also shows $U_{2n}^{--}(f;\boldsymbol{\sigma}) \overset{\mathrm{d}}{=} U_{2n+1}^{--}(f;\boldsymbol{\sigma})$, which in contrast seems to be a very special property for this f. As a consequence, the bi-alternating inversion number has the same asymptotic distribution as the writhe in (8.31), which was found by different methods in [6]. $\triangle$

9. Further results and open problems

9.1. **Joint convergence.** We may also consider joint convergence of the different U-statistics $U_n(f)$, $U_n^{\circ}(f)$, $U_n^{-+}(f)$, $U_n^{+-}(f)$, and $U_n^{--}(f)$ for the same kernel f. In the nondegenerate cases, this is straightforward, since the proofs above in all cases approximate the U-statistic by a linear combination of $f_1(X_i)$ and $f_2(X_i)$, and the central limit theorem implies that these linear combinations converge jointly (after normalization by $n^{-3/2}$) to some jointly normal limits.

Also in the degenerate cases, or when some U-statistics are nondegenerate and some degenerate (and we thus normalize them differently), the methods above make it in principle possible to study also asymptotic joint distributions of U_n, U^{-+}, U^{+-}, and U^{--}, see Remark 5.2 for a simple example; furthermore, it seems possible to include also U_n° by considering $\widetilde{X}_i := (X_{2i-1}, X_{2i}, X_{2i-1+n/2}, X_{2i+n/2})$ (for n divisible by 4). We leave such extensions to the reader.

Note that in all results above showing convergence in distribution to some limit, both in nondegenerate and degenerate cases, the theorems and proofs also show convergence of first and second moments. Hence, if U_n' and U_n'' denote two of the U-statistics in this paper, and $\tilde{U}_n'$ and $\tilde{U}_n''$ are the corresponding normalized variables, then the squares $|\tilde{U}_n'|^2$ and $|\tilde{U}_n''|^2$ are uniformly integrable, see e.g. [8, Theorem 5.5.9]. It follows, by the Cauchy–Schwarz inequality, that also the product $\tilde{U}_n'\tilde{U}_n''$ is uniformly integrable, and thus if $\tilde{U}_n$ and $\tilde{U}_n''$ have limits in distribution jointly, then the covariance of their limits is the limit of their covariances. In particular, when the limits are jointly normal, we can easily find their joint distribution.

In the case when both U_n' and U_n'' have limits of the degenerate type, and we thus may replace f by f_{12} (up to negligible terms), it follows easily from Lemma 2.1 that, except in the case (U_n, U_n°), the two different U-statistics have covariance of order $o(n^2)$ because of cancellations caused by the alternating signs. Hence, except for (U_n, U_n°), any joint limits have to be uncorrelated. (In particular, the two limits cannot be the same random variable, unless they are 0.) We conjecture that, more strongly, in these cases there is joint convergence to independent limits, but we leave this as an open problem. On the other hand, (U_n, U_n°) is different: if we further assume that f is symmetric, then $U_n(f) - U_n^{\circ}(f)$ is negligible, see Remark 1.1, and thus $U_n(f)$ and $U_n^{\circ}(f)$ jointly converge, after normalization, to the same limit.

9.2. **Strong law of large numbers.** We stated in (3.2) the weak law of large numbers for classical U-statistics. There is also a well-known corresponding strong law of large numbers, see [11] (the symmetric case) and [13] (the general case):

$$\frac{1}{\binom{n}{2}} U_n(f) \xrightarrow{\text{a.s.}} \mu. \tag{9.1}$$

This extends to the alternating U-statistics in the following, less interesting form; recall that by (5.1) and (6.2), the expectations are $O(n)$.

Theorem 9.1. *We have, as $n \to \infty$, for any $f \in L^2$,*

$$\frac{1}{\binom{n}{2}} U_n^{--}(f) \xrightarrow{\text{a.s.}} 0. \tag{9.2}$$

The same holds for U_n^{-+} and U_n^{+-}.

Proof. We treat U_n^{+-}; the same argument works for U_n^{-+} and U_n^{--} with minor modifications (and some simplifications). We use (6.17) and treat the terms on the right-hand side separately. As just noted, the expectation $\mathbb{E}\,[U_n^{+-}(f)] = O(n)$, so the first term in (6.17) is $o(n^2)$. The second term has varianc $O(n)$, and it follows from Chebyshev's inequality and the Borel–Cantelli lemma that it is $o(n^2)$ a.s. For the third term, let

$$S_k := \sum_{j=1}^{k} (-1)^j f_2(X_j), \tag{9.3}$$

and note that by the law of large numbers, applied to even and odd indices separately, $S_k/k \xrightarrow{\text{a.s.}} 0$, i.e., $S_k = o(k)$ a.s. Then the third term in (6.17) can be written

$$\sum_{j=1}^{n}\sum_{i=1}^{j-1} (-1)^j f_2(X_j) = \sum_{i=1}^{n}\sum_{j=i+1}^{n} (-1)^j f_2(X_j) = \sum_{i=1}^{n} (S_n - S_i) = o(n^2) \quad \text{a.s.} \tag{9.4}$$

The double sum in (6.17) is $U_n^{+-}(f_{12})$. For even n we use (6.20), where the sum again has variance $O(n)$ and thus a.s. is $o(n^2)$, and $U_n(F)/\binom{n}{2} \xrightarrow{\text{a.s.}} 0$ by (9.1) since $\mathbb{E}\,[F(\widetilde{X}_1, \widetilde{X}_2)] = 0$ by (6.19). Finally, for odd n, the result follows from (6.18), where the sum again has variance $O(n)$. ∎

Note that this argument does not work for U_n°, since the definition (4.13) involves n explicitly, and we therefore cannot apply (9.1) to $U_n(F)$. Moreover, it is not clear that it is interesting to study the sequence $(U_n^{\circ}(f))_{n=1}^{\infty}$ as a stochastic process, since the point of the definition (1.3) of U_n° is that the indices are regarded as elements of $\mathbb{Z}/n\mathbb{Z}$. Nevertheless, out of mathematical curiosity, we might ask:

Problem 9.2. Does (9.1) hold for $U_n^{\circ}(f)$?

9.3. **Functional limit theorems.** As another aspect of regarding the sequence $(U_n(f))_{n=1}^{\infty}$ and its variants as stochastic processes, we may ask for functional limit theorems of Donsker-type. For the classical U-statistic $U_n(f)$, it is known that, extending (3.4), $n^{-1/2}\bigl(U_{\lfloor nt\rfloor}(f) - \frac{n^2}{2}t^2\mu\bigr)$, regarded as a stochastic process with continuous parameter $t \geqslant 0$, converges in $D[0,\infty)$, as $n \to \infty$, to a continuous centred Gaussian process; similarly, in the degenerate case, $n^{-1}\bigl(U_{\lfloor nt\rfloor}(f) - \mathbb{E}\,[U_{\lfloor nt\rfloor}(f)]\bigr)$ converges to a continuous process whose marginals are of the type (3.6); see e.g. [18], [19], [9], [12, Remark 11.11] (the symmetric case); [14, p. 83], [12, Remarks 11.11 and 11.25], [13, Theorem 3.2] (the general case). It seems likely that this too extends to the alternating U-statistics by arguing using (5.20) and (6.20), but we have not checked the details and leave this to the reader.

As in Section 9.2, and for the same reason, this argument does not apply to U_n°; moreover, it seems less interesting to consider functional limit theorems for U_n°.

9.4. **Moment convergence.** As noted in Section 9.1, in all results above showing convergence in distribution, we also have convergence of first and second moments.

For higher moments, it is known that for the classical normal limit in (3.4), and any $p \in (2,\infty)$, all moments and absolute moments of order $\leqslant p$ converge provided $\mathbb{E}\,|f(X_1,X_2)|^p < \infty$, see [13, Theorem 3.15]. It seems likely that this extends to the cyclic and alternating U-statistics considered here, using the methods in the proofs above, but we have not checked the details and leave this to the reader.

We conjecture that there also is a similar result showing moment convergence in the degenerate cases, but in this case we are not even aware of a general result for the classical U-statistic and moment convergence in (3.6).

Appendix A. Proof of Theorem 3.1

We give here a proof of Theorem 3.1, which contains some known result on the asymptotic distribution of U-statistics of the standard type (1.2) in the special case of order $m = 2$. We give a proof for completeness, and because we reuse parts of it for other proofs; we also find it instructive to give complete proofs in the case $m = 2$, which avoids some minor complications for larger m. For previous proofs and for the general case with arbitrary m, see, for example, [10; 7; 22; 5] for the symmetric case, and [12, Chapter 11.1–2] for the general (asymmetric) case.

Proof of Theorem 3.1. To prove Theorem 3.1, we note first that (3.1) is immediate from the definitions (1.2) and (2.2).

We have, by (1.2) and (2.10),

$$\begin{aligned} U_n(f) &= \sum_{1\leqslant i<j\leqslant n} f(X_i,X_j) \\ &= \tbinom{n}{2} f_\varnothing + \sum_{i=1}^n (n-i) f_1(X_i) + \sum_{j=1}^n (j-1) f_2(X_j) + \sum_{1\leqslant i<j\leqslant n} f_{12}(X_i,X_j) \\ &=: S_n^{(\varnothing)} + S_n^{(1)} + S_n^{(2)} + S_n^{(12)}. \end{aligned} \tag{A.1}$$

Here $S_n^{(\varnothing)} = \binom{n}{2}\mu = \mathbb{E}\,U_n(f)$. In the sequel, we may replace f by $f-\mu$; this does not affect f_1, f_2, or f_{12}. (Note also that when f is antisymmetric, $\mu = \mathbb{E}\,f(X_1,X_2) = -\mathbb{E}\,f(X_2,X_1) = -\mu$ and thus $\mu = 0$ so $f-\mu = f$ is still antisymmetric.) We may thus without loss of generality assume that $\mu = 0$, and hence $\mathbb{E}\,U_n(f) = 0$.

We next study the variances of the sums in (A.1). The random vectors $\bigl(f_1(X_i), f_2(X_i)\bigr)$ are i.i.d., with mean 0 and finite second moments. Hence,

$$\operatorname{Var}[S_n^{(1)}] = \sum_{i=1}^n (n-i)^2 \operatorname{Var}[f_1(X)] \sim \tfrac13 n^3 \operatorname{Var}[f_1(X)], \tag{A.2}$$

$$\operatorname{Var}[S_n^{(2)}] = \sum_{j=1}^n (j-1)^2 \operatorname{Var}[f_2(X)] \sim \tfrac13 n^3 \operatorname{Var}[f_2(X)], \tag{A.3}$$

$$\begin{aligned} \operatorname{Cov}[S_n^{(1)}, S_n^{(2)}] &= \sum_{i=1}^n (n-i)(i-1) \operatorname{Cov}[f_1(X), f_2(X)] \\ &\sim \tfrac16 n^3 \operatorname{Cov}[f_1(X), f_2(X)]. \end{aligned} \tag{A.4}$$

Consequently, recalling (3.3),

$$\begin{aligned}\operatorname{Var}[S_n^{(1)}+S_n^{(2)}] &= n^3\bigl(\tfrac13\operatorname{Var}[f_1(X)]+\tfrac13\operatorname{Var}[f_2(X)]+\tfrac26\operatorname{Cov}[f_1(X),f_2(X)]+o(1)\bigr)\\ &= n^3\bigl(\sigma^2+o(1)\bigr). \end{aligned}\tag{A.5}$$

Turning to $S_n^{(12)}$, we note that the terms $f_{12}(X_i,X_j)$ are identically distributed and have mean 0, and that they are orthogonal; this follows from (2.17), which implies that $\mathbb{E}\,[f_{12}(X_1,X_2)\mid X_1]=\mathbb{E}\,[f_{12}(X_1,X_2)\mid X_2]=0$, and thus, for example,

$$\begin{aligned}\mathbb{E}\,[f_{12}(X_1,X_2)f_{12}(X_1,X_3)] &= \mathbb{E}\bigl[\mathbb{E}\,[f_{12}(X_1,X_2)\mid X_1]\,\mathbb{E}\,[f_{12}(X_1,X_3)\mid X_1]\bigr]\\ &=0.\end{aligned}\tag{A.6}$$

Consequently,

$$\operatorname{Var}[S_n^{(12)}]=\sum_{1\leqslant i<j\leqslant n}\operatorname{Var}[f_{12}(X_i,X_j)]=\binom n2\operatorname{Var}[f_{12}(X_1,X_2)].\tag{A.7}$$

The variance of $S_n^{(12)}$ is thus $O(n^2)$, while $S_n^{(1)}+S_n^{(2)}$ typically has a larger variance of order n^3. Hence, the sum (A.1) is dominated by $S_n^{(1)}+S_n^{(2)}$, except in the case that $\sigma^2=0$ when these terms vanish (as we will see below), and therefore (A.1) reduces to $S_n^{(12)}$. This is the reason for the two different cases (ii) and (iii) in Theorem 3.1; the generic case (ii), i.e., assuming $\sigma^2>0$, the nondegenerate case, and the degenerate case (iii) with $\sigma^2=0$. We treat these cases separately below, after completing the proof of (i).

(i): We have already shown (3.1). Furthermore, (A.5) and (A.7) imply

$$\operatorname{Var}[U_n(f)]=O(n^3),\tag{A.8}$$

and thus (3.2) follows. (Actually, with convergence in L^2.)

(ii): We apply the standard central limit theorem for triangular arrays (see for example [8, Theorem 7.2.4] or [15, Theorem 5.12]) to $S_n^{(1)}+S_n^{(2)}$. It is easily verified that the triangular arrays $\bigl(n^{-3/2}(n-i)f_1(X_i)\bigr)_{i\leqslant n}$ and $\bigl(n^{-3/2}(i-1)f_2(X_i)\bigr)_{i\leqslant n}$ satisfy the Lindeberg condition, and thus so does the summed array $\bigl(n^{-3/2}((n-i)f_1(X_i)+(i-1)f_2(X_i))\bigr)_{i\leqslant n}$. Consequently, the central limit theorem yields, using (A.5),

$$n^{-3/2}(S_n^{(1)}+S_n^{(2)})\xrightarrow{\mathrm{d}}N(0,\sigma^2).\tag{A.9}$$

(If $\sigma^2=0$ then (A.9) still holds, as a trivial consequence of (A.5).) Furthermore, as noted above, (A.7) implies that $\operatorname{Var}[n^{-3/2}S_n^{(12)}]=n^{-3}\operatorname{Var}[S_n^{(12)}]\to0$, and thus $n^{-3/2}S_n^{(12)}\xrightarrow{\mathrm{p}}0$. Hence, (3.4) follows from (A.9) and (A.1) by the Cramér–Slutsky theorem [8, Theorem 5.11.4].

Finally, (3.3) can be written

$$3\sigma^2=\mathbb{E}\,[(f_1(X)+\tfrac12f_2(X))^2]+\tfrac34\,\mathbb{E}\,[f_2(X)^2],\tag{A.10}$$

which implies that $\sigma^2=0$ if and only if $f_1(X)=f_2(X)=0$ a.s., which completes the proof for the nondegenerate case (ii).

We turn to the degenerate case $\sigma^2=0$ in (iii)–(v). In this case we thus have $f_1(X_i)=f_2(X_i)=0$ a.s., and consequently $S_n^{(1)}=S_n^{(2)}=0$ and, by (A.1) again,

$$U_n(f)=S_n^{(12)}.\tag{A.11}$$

We first treat the symmetric case (iv).

(iv): We apply Lemma A.1 below, noting that (A.29) holds by (2.12) and the assumption that $f_1(X) = 0$ a.s. (and our simplifying assumption $f_\varnothing = \mu = 0$ in this proof). This shows that $f(x, y) = f_{12}(x, y)$ has an orthogonal expansion (A.27), for some $R \leqslant \infty$ and some orthonormal sequence of functions $\varphi_r \in L^2_{\mathbb{R}}(\mathcal{X})$, which furthermore satisfy (A.30), which is equivalent to $\mathbb{E}\,\varphi_r(X) = 0$. The orthonormality means that

$$\mathbb{E}\left[\varphi_r(X)\varphi_q(X)\right] = \int_{\mathcal{X}} \varphi_r(x)\varphi_q(x)\,\mathrm{d}\nu(x) = \delta_{rq}. \tag{A.12}$$

In other words, $(\varphi_r(X))_1^R$ is a sequence of uncorrelated random variables with mean 0 and variance 1.

Suppose first that $R < \infty$, so that the sum (A.27) is finite. Then (1.2), the symmetry of f, and (A.27) yield

$$\begin{aligned}
2U_n(f) &= 2\sum_{1\leqslant i<j\leqslant n} f(X_i, X_j) = \sum_{i,j=1}^n f(X_i, X_j) - \sum_{i=1}^n f(X_i, X_i)\\
&= \sum_{i,j=1}^n \sum_{r=1}^R \lambda_r\varphi_r(X_i)\varphi_r(X_j) - \sum_{i=1}^n\sum_{r=1}^R \lambda_r\varphi_r(X_i)^2\\
&= \sum_{r=1}^R \lambda_r\Bigl(\Bigl(\sum_{i=1}^n \varphi_r(X_i)\Bigr)^2 - \sum_{i=1}^n \varphi_r(X_i)^2\Bigr).
\end{aligned} \tag{A.13}$$

Now let $n \to \infty$. By the law of large numbers, for each r,

$$n^{-1}\sum_{i=1}^n \varphi_r(X_i)^2 \xrightarrow{\mathrm{p}} \mathbb{E}\left[\varphi_r(X)^2\right] = 1. \tag{A.14}$$

Furthermore, by the central limit theorem, since $\mathbb{E}\left[\varphi_r(X)\right] = 0$ and $\mathbb{E}\left[\varphi_r(X)^2\right] = 1$ as remarked above,

$$n^{-1/2}\sum_{i=1}^n \varphi_r(X_i) \xrightarrow{\mathrm{d}} \zeta_r \in N(0,1). \tag{A.15}$$

Moreover, since the variables $\varphi_r(X)$ are uncorrelated, the limit in (A.15) holds jointly for all $r \leqslant R$, with the limits ζ_r uncorrelated and thus independent. Combining (A.13)–(A.15) yields

$$\begin{aligned}
2n^{-1}U_n(f) &= \sum_{r=1}^R \lambda_r\Bigl(\Bigl(n^{-1/2}\sum_{i=1}^n \varphi_r(X_i)\Bigr)^2 - n^{-1}\sum_{i=1}^n \varphi_r(X_i)^2\Bigr)\\
&\xrightarrow{\mathrm{d}} \sum_{r=1}^R \lambda_r\bigl(\zeta_r^2 - 1\bigr).
\end{aligned} \tag{A.16}$$

This proves (3.6) when $R < \infty$.

If $R = \infty$, let, for $N \in \mathbb{N}$,

$$f_N(x, y) := \sum_{r=1}^N \lambda_r\varphi_r(x)\varphi_r(y). \tag{A.17}$$

Then the case just proven applies to f_N, and thus, for each fixed $N < \infty$, as $n \to \infty$,

$$2n^{-1}U_n(f_N) \xrightarrow{\mathrm{d}} \sum_{r=1}^{N} \lambda_r\bigl(\zeta_r^2 - 1\bigr). \tag{A.18}$$

As $N \to \infty$, the right-hand side converges to $\sum_{r=1}^{\infty} \lambda_r\bigl(\zeta_r^2 - 1\bigr)$ in L^2 and a.s., and in particular in distribution. Furthermore, by (A.7) applied to $U_n(f - f_N)$,

$$\begin{aligned}\mathbb{E}\left[(U_n(f) - U_n(f_N))^2\right] &= \operatorname{Var}\bigl[U_n(f - f_N)\bigr] = \tbinom{n}{2}\operatorname{Var}\bigl((f - f_N)(X_1, X_2)\bigr)\\ &= \tbinom{n}{2}\int_{\mathcal{X}^2}\Bigl(\sum_{N+1}^{\infty} \lambda_r\varphi_r(x)\varphi_r(y)\Bigr)^2 \mathrm{d}\nu(x)\,\mathrm{d}\nu(y)\\ &= \tbinom{n}{2}\sum_{N+1}^{\infty} \lambda_r^2. \end{aligned}\tag{A.19}$$

Hence,

$$\mathbb{E}\left[(n^{-1}U_n(f) - n^{-1}U_n(f_N))^2\right] \leqslant \sum_{N+1}^{\infty} \lambda_r^2 \to 0 \tag{A.20}$$

as $N \to \infty$, uniformly in n. The result (3.6) now follows from (A.18) and (A.20), see e.g. [1, Theorem 4.2].

Finally, by (3.6) and (2.20),

$$\operatorname{Var} W = \sum_{r=1}^{R}\bigl(\tfrac12\lambda_r\bigr)^2\operatorname{Var}\bigl(\zeta_r^2 - 1\bigr) = \tfrac12\sum_{r=1}^{R}\lambda_r^2 \tag{A.21}$$

and thus (3.7) holds by (A.28) in Lemma A.1.

(iii): Now consider the general degenerate case, where $f_1 = f_2 = 0$, and further as above without loss of generality $f_\varnothing = \mu = 0$, but no symmetry assumption is made. We use the following trick to reduce to the symmetric case. (See [12, Remark 11.21] for the case of general order m.)

Let $(Z_i)_1^\infty$ be an i.i.d. sequence of random variables, independent of $(X_i)_1^\infty$, with each Z_i uniformly distributed on $[0, 1]$. Consider the random variables $\widehat{X}_i := (X_i, Z_i)$ in $\widehat{\mathcal{X}} := \mathcal{X} \times [0, 1]$ and define the function $\widehat{f} : \widehat{\mathcal{X}}^2 \to \mathbb{R}$ by (2.6). Note that this definition makes $\widehat{f}$ a symmetric function on $\widehat{\mathcal{X}} \times \widehat{\mathcal{X}}$. Furthermore, if we condition on the sequence $(Z_i)_1^n$, and assume as we may that $Z_1, \dots, Z_n$ are distinct, then, letting π be the permutation of $\{1, \dots, n\}$ that makes $Z_{\pi(1)} < \cdots < Z_{\pi(n)}$,

$$U_n(\widehat{f}; \widehat{X}_1, \dots, \widehat{X}_n) = U_n(\widehat{f}; \widehat{X}_{\pi(1)}, \dots, \widehat{X}_{\pi(n)}) = U_n(f; X_{\pi(1)}, \dots, X_{\pi(n)}), \tag{A.22}$$

where the first equality holds by the symmetry of $\widehat{f}$ and the second by the definitions of $\widehat{f}$ (in (2.6)) and π. Consequently, conditioned on $(Z_i)_1^n$ we have

$$U_n(\widehat{f}; \widehat{X}_1, \dots, \widehat{X}_n) \overset{\mathrm{d}}{=} U_n(f; X_1, \dots, X_n), \tag{A.23}$$

and hence (A.23) holds also unconditionally. The result now follows from (iv) applied to $\widehat{f}$ and $(\widehat{X}_i)_1^\infty$; note that this case applies since the definition (2.6) implies that, with definitions analogous to (2.11)–(2.14),

$$\widehat{f}_\varnothing = \int_{[0,1]^2}\int_{X^2} \widehat{f}\bigl((x, t), (y, u)\bigr)\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y)\,\mathrm{d}t\,\mathrm{d}u = \mu = 0, \tag{A.24}$$

$$\widehat{f}_1(x,t) = \int_{\widehat{\mathcal{X}}} \widehat{f}\big((x,t),(y,u)\big)\,\mathrm{d}\nu(y)\,\mathrm{d}u = \begin{Bmatrix} f_1(x), & t < u \\ f_2(x), & t > u \end{Bmatrix} = 0 \tag{A.25}$$

and, by symmetry, $\widehat{f}_2(x,t) = \widehat{f}_1(x,t) = 0$. Furthermore, (3.7) for f follows from the same formula for $\widehat{f}$, since (2.7) shows that $\operatorname{Var}[\widehat{f}(\widehat{X}_1,\widehat{X}_2)] = \operatorname{Var}[f(X_1,X_2)]$.

(v): This is proved directly in [14, Theorem 2.1] by different methods, relating U_n in the asymmetric case to the stochastic area process. We give here a different proof, by combining the general result in (iii) with Lemma A.2 below, which finds the eigenvalues λ_r of $T_{\widehat{f}}$ and shows that the limit variable $W := \sum_{r=1}^{R} \frac{1}{2}\lambda_r(\zeta_r^2 - 1)$ in (3.6) also has the representation (3.8). The formula (3.9) follows from (3.7) and (A.37). □

Lemma A.1. *If $f \in L^2(\mathcal{X} \times \mathcal{X})$ is real and symmetric, then T_f defined by* (2.3)*, i.e.,*

$$T_f g(x) := \int_{\mathcal{X}} f(x,y)g(y)\,\mathrm{d}\nu(y), \tag{A.26}$$

is a self-adjoint Hilbert–Schmidt operator on $L^2_{\mathbb{C}}(\mathcal{X})$. Hence T_f is compact and has thus at most countably many nonzero eigenvalues, each of them real and each having a finite-dimensional eigenspace. Let $(\lambda_r)_{r=1}^R$ (where $0 \leqslant R \leqslant \infty$) be an enumeration of the nonzero eigenvalues (with multiplicities) of T_f. It is then possible to find a corresponding orthonormal sequence of real-valued eigenfunctions $(\varphi_r)_1^R \in L^2_{\mathbb{R}}(\mathcal{X})$ such that $T_f\varphi_r = \lambda_r\varphi_r$ for every r.

For any such $(\lambda_r)_1^R$ and $(\varphi_r)_1^R$, $f(x,y)$ has a (finite or infinite) orthogonal expansion

$$f(x,y) = \sum_{r=1}^{R} \lambda_r\varphi_r(x)\varphi_r(y) \tag{A.27}$$

which converges in $L^2(\mathcal{X}^2)$ because

$$\sum_{r=1}^{R} \lambda_r^2 = \int_{\mathcal{X}^2} f(x,y)^2\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y) < \infty. \tag{A.28}$$

Moreover, if

$$\int_{\mathcal{X}} f(x,y)\,\mathrm{d}\nu(x) = 0, \qquad \text{for } \nu\text{-a.e. } y \in \mathcal{X}, \tag{A.29}$$

then

$$\int_{\mathcal{X}} \varphi_r(x)\,\mathrm{d}\nu(x) = 0, \qquad \text{for every } r \leqslant R. \tag{A.30}$$

Here and below, $r \leqslant R$ should be interpreted as $r < \infty$ when $R = \infty$.

Proof. Since $f \in L^2(\mathcal{X}^2)$, (A.26) defines a bounded and compact linear operator T_f on $L^2_{\mathbb{C}}(\mathcal{X})$; furthermore, T_f is a Hilbert-Schmidt operator. Since furthermore f is real and symmetric, T_f is a self-adjoint operator. By the spectral theorem for compact and self-adjoint linear operators on a Hilbert space (see e.g. [16, Theorem 28.3] or [23, Theorem 6.4-B]), T_f has a finite or countably infinite set of nonzero eigenvalues, each with finite multiplicity, so we may arrange the nonzero eigenvalues, with multiplicities, in a sequence $(\lambda_r)_1^R$ with $R \leqslant \infty$; we denote the index set also by $\mathcal{R} := \{r \in \mathbb{N} : r \leqslant R\}$. Moreover, the eigenvalues λ_r are real, and there exists a corresponding orthonormal sequence of eigenfunctions $(\varphi_r)_1^R$, and this may be extended to an orthonormal basis $(\varphi_r)_{r\in\mathcal{R}\cup\mathcal{N}}$ where $\mathcal{N}$ is a disjoint (possibly empty) index set such that $T_f(\varphi_r)=0$ for every $r\in\mathcal{N}$, i.e., each φ_r is an eigenfunction also for $r\in\mathcal{N}$, with eigenvalue $\lambda_r:=0$ when $r\in\mathcal{N}$.

Furthermore, since f is real, we may choose all φ_r to be in $L^2_{\mathbb{R}}(\mathcal{X})$. (By [23, Theorem 6.4-B] applied to $L^2_{\mathbb{R}}(\mathcal{X})$, or by noting that since also the eigenvalues are real, the real and imaginary parts of any eigenfunction of T_f are also eigenfunctions for the same eigenvalue; hence each eigenspace is spanned by the real functions in it.)

Fix any such sequence $(\varphi_r)_{r\in\mathcal{R}}$ and extension $(\varphi_r)_{r\in\mathcal{R}\cup\mathcal{N}}$. Since $(\varphi_r)_{r\in\mathcal{R}\cup\mathcal{N}}$ is an orthonormal basis in $L^2(\mathcal{X})$, it is easily seen (and well-known) that if we define $g\otimes h(x,y):=g(x)h(y)$ for functions $g,h\in L^2(\mathcal{X})$, then the set $\{\varphi_r\otimes\varphi_s : r,s\in\mathcal{R}\cup\mathcal{N}\}$ is an orthonormal basis in $L^2(\mathcal{X}\times\mathcal{X})$. Hence,

$$f=\sum_{r,s\in\mathcal{R}\cup\mathcal{N}}\langle f,\varphi_r\otimes\varphi_s\rangle\,\varphi_r\otimes\varphi_s, \tag{A.31}$$

where the sum converges in L^2. By Fubini's theorem and (A.26),

$$\begin{aligned}\langle f,\varphi_r\otimes\varphi_s\rangle&=\iint_{\mathcal{X}\times\mathcal{X}} f(x,y)\varphi_r(x)\varphi_s(y)\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y)=\int_{\mathcal{X}}\varphi_r(x)T_f(\varphi_s)(x)\,\mathrm{d}\nu(x)\\&=\int_{\mathcal{X}}\varphi_r(x)\lambda_s\varphi_s(x)\,\mathrm{d}\nu(x)=\lambda_s\langle\varphi_r,\varphi_s\rangle=\lambda_s\delta_{rs}.\end{aligned} \tag{A.32}$$

This vanishes unless $r=s\in\mathcal{R}$, and thus (A.31) simplifies to, using (A.32) again,

$$f=\sum_{r\in\mathcal{R}}\langle f,\varphi_r\otimes\varphi_r\rangle\,\varphi_r\otimes\varphi_r=\sum_{r\in\mathcal{R}}\lambda_r\,\varphi_r\otimes\varphi_r. \tag{A.33}$$

This is (A.27), and (A.28) follows because $\{\varphi_r\otimes\varphi_s\}$ is an orthonormal basis.

Finally, if (A.29) holds, then for every $r\in\mathcal{R}$, by (A.26) and Fubini's theorem,

$$\begin{aligned}\lambda_r\int_{\mathcal{X}}\varphi_r(x)\,\mathrm{d}\nu(x)&=\int_{\mathcal{X}}T_f(\varphi_r)(x)\,\mathrm{d}\nu(x)=\iint_{\mathcal{X}\times\mathcal{X}} f(x,y)\varphi_r(y)\,\mathrm{d}\nu(y)\,\mathrm{d}\nu(x)\\&=\int_{\mathcal{X}}\varphi_y(y)\int_{\mathcal{X}} f(x,y)\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y)=0.\end{aligned} \tag{A.34}$$

Since $\lambda_r\neq 0$ for $r\in\mathcal{R}$, (A.30) follows. ☐

Lemma A.2. *Let $f\in L^2(\mathcal{X}\times\mathcal{X})$ be real and antisymmetric. The the operator T_f on $L^2_{\mathbb{C}}(\mathcal{X})$ is anti-self-adjoint and has purely imaginary eigenvalues. Let $(\lambda^{\mathsf{a}}_q)_{q\in\mathcal{Q}_+}$ be an enumeration of the positive real numbers such that $\mathrm{i}\lambda^{\mathsf{a}}_q$ is an eigenvalue of T_f (counted with multiplicities). Then the multiset of nonzero eigenvalues $(\lambda_r)_1^R$ of the self-adjoint operator $T_{\hat f}$ on $L^2(\mathcal{X}\times[0,1])$ (counted with multiplicities) equals*

$$\Bigl\{\pm\frac{2}{(2k-1)\pi}\lambda^{\mathsf{a}}_q : q\in\mathcal{Q}_+, k\in\mathbb{N}\Bigr\}\quad\textit{with each pair } (q,k) \textit{ counted twice.} \tag{A.35}$$

As a consequence, if ζ_r are i.i.d. standard normal variables and η_q are independent random variables with the stochastic area distribution (2.23), *then*

$$\sum_{r=1}^{R}\tfrac12\lambda_r(\zeta_r^2-1)\overset{\mathrm{d}}{=}\sum_{q\in\mathcal{Q}_+}\lambda^{\mathsf{a}}_q\eta_q. \tag{A.36}$$

Furthermore,

$$\sum_{r=1}^{R}\lambda_r^2 = 2\sum_{q=1}^{Q_+}(\lambda^{\mathsf{a}}_q)^2 = \int_{\mathcal{X}^2} f(x,y)^2\,\mathrm{d}\nu(x)\,\mathrm{d}\nu(y) < \infty. \tag{A.37}$$

Proof. In the antisymmetric case, we can write (2.6) as

$$\widehat{f}\big((x,t),(y,u)\big) = f(x,y)\operatorname{sgn}(u-t) =: f(x,y)h(t,u), \tag{A.38}$$

where sgn is the sign function (2.9). Thus, in tensor notation, see (2.4), $\widehat{f} = f\otimes h$ and $T_{\widehat{f}} = T_f \otimes T_h$.

The functions $f(x,y)$ and $h(t,u) = \operatorname{sgn}(u-t)$ in (A.38) are both real-valued and antisymmetric, and thus the corresponding Hilbert–Schmidt integral operators T_f and T_h (acting on $L^2_{\mathbb{C}}(\mathcal{X})$ and $L^2_{\mathbb{C}}[0,1]$, respectively) are both anti-self-adjoint. Hence, $-\mathrm{i}T_f$ and $-\mathrm{i}T_h$ are self-adjoint, and it follows from the spectral theorem, as in the proof of Lemma A.1, that $-\mathrm{i}T_f$ and $-\mathrm{i}T_h$ have only real eigenvalues $\{\lambda^{\mathsf{a}}_q : q\in\mathcal{Q}\cup\mathcal{N}\}$ and $\{\rho_s : s\in\mathcal{S}\cup\mathcal{N}'\}$, respectively, with $\lambda^{\mathsf{a}}_q\neq 0 \iff q\in\mathcal{Q}$ and $\rho_s\neq 0 \iff s\in\mathcal{S}$ and that there are corresponding families of eigenfunctions $\{\varphi_q : q\in\mathcal{Q}\cup\mathcal{N}\}$ and $\{\psi_s : s\in\mathcal{S}\cup\mathcal{N}'\}$ which are orthonormal bases in $L^2_{\mathbb{C}}(\mathcal{X})$ and $L^2_{\mathbb{C}}[0,1]$, respectively. (However, unlike in Lemma A.1, these eigenfunctions are not real-valued.) Hence, these functions are eigenfunctions for T_f and T_h too, with eigenvalues $\mathrm{i}\lambda^{\mathsf{a}}_q$ and $\mathrm{i}\rho_s$, respectively. (The eigenvalues and eigenfunctions for T_h will be found explicitly in Lemma A.3.)

It follows that the set of all functions $\varphi_q\otimes\psi_s(x,t) := \varphi_q(x)\psi_s(t)$ is an orthonormal basis in $L^2_{\mathbb{C}}(\mathcal{X}\times[0,1])$. Furthermore, as noted in Section 2, the function $\varphi_q\otimes\psi_s$ is an eigenfunction of $T_{\widehat{f}} = T_f\otimes T_h$ with eigenvalue $-\lambda^{\mathsf{a}}_q\rho_s$. Since these functions form a basis, it follows that the set of eigenvalues of $T_{\widehat{f}}$, with multiplicities, is $\{-\lambda^{\mathsf{a}}_q\rho_s : q\in\mathcal{Q}\cup\mathcal{N}, s\in\mathcal{S}\cup\mathcal{N}'\}$. In particular, the nonzero eigenvalues $(\lambda_r)_1^R$ are

$$\{-\lambda^{\mathsf{a}}_q\rho_s : q\in\mathcal{Q}, s\in\mathcal{S}\}. \tag{A.39}$$

Recall that the nonzero eigenvalues of T_f are $\{\mathrm{i}\lambda^{\mathsf{a}}_q\}_{q\in\mathcal{Q}}$, where $\lambda^{\mathsf{a}}_q\in\mathbb{R}$. Since f is real, the complex conjugate $\overline{\varphi_q}$ is also an eigenfunction, with eigenvalue $\overline{\mathrm{i}\lambda^{\mathsf{a}}_q} = -\mathrm{i}\lambda^{\mathsf{a}}_q$. It follows that if we let $\mathcal{Q}_+ := \{q : \lambda^{\mathsf{a}}_q > 0\}$ and $\mathcal{Q}_- := \{q : \lambda^{\mathsf{a}}_q < 0\}$, then $\{\lambda^{\mathsf{a}}_q : q\in\mathcal{Q}_-\} = \{-\lambda^{\mathsf{a}}_q : q\in\mathcal{Q}_+\}$. Consequently, we may rewrite (A.39) as

$$\{\lambda_r : r\leqslant R\} = \{\pm\lambda^{\mathsf{a}}_q\rho_s : q\in\mathcal{Q}_+, s\in\mathcal{S}\}. \tag{A.40}$$

We now use Lemma A.3, which shows that the eigenvalues $\mathrm{i}\rho_s$ are

$$\Big\{\pm\frac{2\mathrm{i}}{(2k-1)\pi} : k\in\mathbb{N}\Big\}. \tag{A.41}$$

Hence, (A.35) follows from (A.40), noting that for each pair (q,k), there are two choices of signs in (A.40) and (A.41) that yield the same λ_r.

Note that each pair (q,k) thus yields 4 eigenvalues in (A.35), 2 of each sign. Hence, it follows from (A.35) that, with $\zeta_{q,k,j}\in N(0,1)$ independent,

$$\sum_{r=1}^{R}\tfrac{1}{2}\lambda_r(\zeta_r^2-1) \overset{\mathrm{d}}{=} \sum_{q\in\mathcal{Q}_+}\lambda^{\mathsf{a}}_q\sum_{k=1}^{\infty}\frac{1}{(2k-1)\pi}\big(\zeta^2_{q,k,1}+\zeta^2_{q,k,2}-\zeta^2_{q,k,3}-\zeta^2_{q,k,4}\big). \tag{A.42}$$

Consequently, (A.36) follows from Lemma 2.5. Finally, (A.37) follows from (A.35) (or from (A.40)–(A.41)) which yields

$$\sum_{r=1}^{R}\lambda_r^2=4\sum_{q\in\mathcal{Q}_+}\sum_{k=1}^{\infty}\Bigl(\frac{2}{(2k-1)\pi}\lambda_q^{\mathsf{a}}\Bigr)^2=\frac{16}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}\sum_{q\in\mathcal{Q}_+}(\lambda_q^{\mathsf{a}})^2=2\sum_{q\in\mathcal{Q}_+}(\lambda_q^{\mathsf{a}})^2, \tag{A.43}$$

together with (A.28) for $T_{\hat f}$ (or the corresponding formula for the self-adjoint operator $T_{\mathrm{i}f}$). ☐

Lemma A.3. *Let $h(t,u):=\operatorname{sgn}(t-u)$. Then the anti-self-adjoint operator T_h acting on $L^2_{\mathbb{C}}[0,1]$ has eigenvalues, all simple,*

$$\pm\frac{2\mathrm{i}}{(2k-1)\pi},\qquad k=1,2,3,\dots. \tag{A.44}$$

Proof. Suppose that φ is an eigenfunction with eigenvalue λ. Then, for a.e. $t\in[0,1]$,

$$\lambda\varphi(t)=T_h\varphi(t)=\int_0^1\operatorname{sgn}(u-t)\varphi(u)\,\mathrm{d}u=-\int_0^t\varphi(u)\,\mathrm{d}u+\int_t^1\varphi(u)\,\mathrm{d}u \tag{A.45}$$

Suppose first that $\lambda\neq0$. The right-hand side of (A.45) is a continuous function of $t\in[0,1]$, and thus φ can be assumed to be continuous. Then (A.45) holds for every $t\in[0,1]$, and the right-hand side of (A.45) is continuously differentiable in $(0,1)$; thus φ is continuously differentiable on $[0,1]$. Taking the derivative in (A.45) yields

$$\lambda\varphi'(t)=-2\varphi(t) \tag{A.46}$$

and thus (for some irrelevant $C\neq0$)

$$\varphi(t)=Ce^{i\omega t}\qquad\text{with}\qquad\omega=2\mathrm{i}/\lambda. \tag{A.47}$$

Taking $t=0$ and 1 in (A.45) yields

$$\lambda\varphi(0)=\int_0^1\varphi(u)\,\mathrm{d}u=-\lambda\varphi(1), \tag{A.48}$$

i.e., $\varphi(1)=-\varphi(0)$. This and (A.47) yield $e^{\mathrm{i}\omega}=-1$, and thus

$$\omega=\pm(2k-1)\pi,\qquad k=1,2,\dots. \tag{A.49}$$

Conversely, it is easily checked that each such ω gives an eigenfunction φ by (A.47), satisfying (A.45) with eigenvalue

$$\lambda=\frac{2\mathrm{i}}{\omega}=\pm\frac{2\mathrm{i}}{(2k-1)\pi}. \tag{A.50}$$

These are thus the nonzero eigenvalues, and we see from (A.47) that they are simple.

Finally, if (A.45) holds with $\lambda=0$, then $\int_0^t\varphi(u)\,\mathrm{d}u$ is constant, and thus $\varphi(t)=0$ a.e. Hence, 0 is not an eigenvalue. (Equivalently, the eigenfunctions $e^{\pm(2k-1)\pi\mathrm{i}}$ form an orthonormal basis on $[0,1]$, as is well-known from Fourier analysis.) ☐

Appendix B. Cumulants

The limit distributions in our theorems are, apart form the normal distribution, given by (possibly infinite) linear combinations of independent copies of the variables $\zeta^2 - 1$, η, and ϑ in Section 2.3. We give here some simple results on the cumulants of such sums; we denote the cumulants of a random variable Y by $\varkappa_m(Y)$, where $m \geqslant 1$. We write for convenience $\chi := \zeta^2 - 1$.

The cumulants $\varkappa_m(\chi)$, $\varkappa_m(\eta)$, and $\varkappa_m(\vartheta)$ are (by definition) obtained by Taylor expansions of the logarithms of the characteristic functions of χ, η, and ϑ given in (2.21) (2.23), and (2.25); note that these characteristic functions and their logarithms are analytic functions of t in a neighbourhood of 0. This yields for $m \geqslant 2$ the cumulants, using [20, 4.19.8, 4.28.9, and 24.2.2] for (B.2)–(B.3),

$$\varkappa_m(\chi) = 2^{m-1}(m-1)!, \tag{B.1}$$

$$\varkappa_m(\eta) = (-1)^{1+m/2}\frac{2^m(2^m-1)}{m}B_m = \frac{2^m(2^m-1)}{m}|B_m|, \tag{B.2}$$

$$\varkappa_m(\vartheta) = \tfrac{1}{2}\varkappa_m(\eta) = (-1)^{1+m/2}\frac{2^{m-1}(2^m-1)}{m}B_m = \frac{2^{m-1}(2^m-1)}{m}|B_m|, \tag{B.3}$$

where B_m denotes the Bernoulli numbers [20, Chapter 24]. For $m = 1$, we have $\varkappa(\chi)_1 = \varkappa(\eta)_1 = \varkappa(\vartheta)_1 = 0$. (These are just the means.) Note that thus $\varkappa_m(\eta) = \varkappa_m(\vartheta) = 0$ for all odd m, which reflects the fact that η and ϑ have symmetric distributions.

Sums of the type $\sum_{r=1}^R \lambda_r\chi_r$, $\sum_{r=1}^R \lambda_r\eta_r$, and $\sum_{r=1}^R \lambda_r\vartheta_r$ (where χ_r are independent copies of χ, and so on) appear frequently above; they have cumulants that can be expressed in terms of the sums $\sum_{r=1}^R \lambda_r^m$, since the cumulant of a sum of independent variables equals the sum of their cumulants. Hence, for any finite or infinite sequence $(\lambda_r)_1^R$ with $\sum_{r=1}^R \lambda_r^2 < \infty$,

$$\varkappa_m\Bigl(\sum_{r=1}^R \lambda_r\chi_r\Bigr) = \sum_{r=1}^R \lambda_r^m \cdot \varkappa_m(\chi), \tag{B.4}$$

with $\varkappa_m(\chi)$ given by (B.1), and similarly for $\sum_{r=1}^R \lambda_r\eta_r$ and $\sum_{r=1}^R \lambda_r\vartheta_r$.

For example, since $\eta \overset{\mathrm{d}}{=} \vartheta_1 + \vartheta_2$, we have $\varkappa_m(\eta) = 2\varkappa_m(\vartheta)$ for all $m \geqslant 1$, as is seen in (B.2)–(B.3). Similarly, since (2.28) can be written

$$\vartheta \overset{\mathrm{d}}{=} \sum_{k=-\infty}^{\infty} \frac{1}{(2k-1)\pi}\chi_k, \tag{B.5}$$

we have, for all even $m \geqslant 2$, using the standard formula [20, 25.6.2] for $\zeta(m)$,

$$\begin{aligned}\varkappa_m(\vartheta) &= \sum_{k=-\infty}^{\infty} \frac{1}{((2k-1)\pi)^m}\varkappa_m(\chi) = \pi^{-m}2\bigl(1-2^{-m}\bigr)\zeta(m)\varkappa_m(\chi)\\ &= \frac{2^m\bigl(1-2^{-m}\bigr)}{m!}|B_m|\varkappa_m(\chi) = \frac{2^m\bigl(1-2^{-m}\bigr)}{m!}|B_m|2^{m-1}(m-1)!,\end{aligned} \tag{B.6}$$

which agrees with (and thus gives a proof of) (B.3) and (B.2). (Recall that the odd cumulants of ζ and η are 0.)

For a final example, the limit distribution $\sum_{k=1}^{\infty} \frac{1}{\pi k}\eta_k$ in (8.31) has cumulants, using (B.2) and, again, [20, 25.6.2],

$$\varkappa_m\Bigl(\sum_{k=1}^{\infty} \frac{1}{\pi k}\eta_k\Bigr) = \sum_{k=1}^{\infty} \frac{1}{(\pi k)^m}\varkappa_m(\eta) = \pi^{-m}\zeta(m)\varkappa_m(\eta) = \frac{2^{2m}(2^m-1)}{2m\cdot m!}B_m^2, \tag{B.7}$$

which agrees with the cumulants given (implicitly) in [6, Corollary 2]. (Note that the variable W there is twice as big, because of different normalizations.)

Appendix C. A proof of (8.17)–(8.19)

We give here, as another illustration of the theorems and methods in the paper, a proof of (8.19) and thus (2.27) using Theorem 3.1 and eigenvalue calculations; this is hardly new, but we do not know a reference. We omit some details.

As said in Remark 8.3, we leave it as an open problem to do similar calculations for the operator (6.7) in Theorem 6.1, which ought to lead to an explicit (more or less complicated) formula for the joint characteristic function of the limits in distribution of $n^{-1}U_n^{+-}(f_{\mathsf{s}})$ and $n^{-1}U_n^{+-}(f_{\mathsf{a}})$,

Let f_{s} and f_{a} be as in (8.15)–(8.16), and let $s,\tau\in\mathbb{R}$. (We use τ here, since we want to use $t\in I$ as one of the coordinates in $\mathcal{X}\times[0,1]$.) Take, suppressing the argument $((x_1,x_2),(x_1',x_2'))$,

$$f_{s,\tau} := s\cdot 2f_{\mathsf{s}} + \tau\cdot 2f_{\mathsf{a}} = ax_1x_2' + bx_2x_1', \tag{C.1}$$

where $a := s+\tau$, $b := s-\tau$. Similarly as in (8.8)–(8.10), we see that an eigenfunction of $T_{\widehat{f_{s,\tau}}}$ with nonzero eigenvalue λ has to be of the form $x_1\psi_1(t)+x_2\psi_2(t)$, and the eigenvalue equation is equvalent to the system

$$\lambda\psi_1(t) = a\int_t^1 \psi_2(u)\,\mathrm{d}u + b\int_0^t \psi_2(u)\,\mathrm{d}u, \tag{C.2}$$

$$\lambda\psi_2(t) = b\int_t^1 \psi_1(u)\,\mathrm{d}u + a\int_0^t \psi_1(u)\,\mathrm{d}u. \tag{C.3}$$

This is in turn equivalent to the system of differential equations

$$\lambda\psi_1'(t) = (b-a)\psi_2(t), \tag{C.4}$$

$$\lambda\psi_2'(t) = (a-b)\psi_1(t), \tag{C.5}$$

with the initial values

$$\lambda\psi_1(0) = a\int_0^1 \psi_2(u)\,\mathrm{d}u, \qquad \lambda\psi_2(0) = b\int_0^1 \psi_1(u)\,\mathrm{d}u. \tag{C.6}$$

Assume $s\tau\neq 0$, or equivalently $a\neq\pm b$. (This excludes the cases f_{s} and f_{a} already studied, which are somewhat special.) Let

$$\omega = \frac{a-b}{\lambda}. \tag{C.7}$$

The general solution to (C.4)–(C.5) then is, for some real (or complex) A and B,

$$\psi_1(t) = A\cos(\omega x) + B\sin(\omega x), \tag{C.8}$$

$$\psi_2(t) = A\sin(\omega x) - B\cos(\omega x), \tag{C.9}$$

and (C.6) yields, using (C.7) and (C.4)–(C.5), the conditions

$$(a-b)\psi_1(0) = \omega\lambda\psi_1(0) = a\omega\int_0^1 \psi_2(u)\,\mathrm{d}u = a\psi_1(0) - a\psi_1(1), \tag{C.10}$$

$$(a-b)\psi_2(0) = \omega\lambda\psi_2(0) = b\omega\int_0^1 \psi_1(u)\,\mathrm{d}u = b\psi_2(1) - b\psi_2(0), \tag{C.11}$$

which simplify to

$$b\psi_1(0) = a\psi_1(1), \qquad a\psi_2(0) = b\psi_2(1) \tag{C.12}$$

or

$$bA = a(A\cos\omega + B\sin\omega), \tag{C.13}$$

$$-aB = b(A\sin\omega - B\cos\omega). \tag{C.14}$$

Regarding (C.13)–(C.14) as a system of linear equations in (A,B), it follows that there is a solution to (C.2)–(C.3), and thus an eigenvalue λ, if and only if the determinant of the system (C.13)–(C.14) is 0, i.e., if

$$0 = (a\cos\omega - b)(a - b\cos\omega) - ab\sin^2\omega = (a^2+b^2)\cos\omega - 2ab. \tag{C.15}$$

Furthermore, since we assume $a \neq \pm b$, it is easily seen that then the system (C.13)–(C.14) has rank 1 and thus a one-dimensional space of solutions (A,B); hence, the eigenvalue λ is simple. Let

$$\omega_0 := \arccos\frac{2ab}{a^2+b^2}. \tag{C.16}$$

Then the complete set of solutions ω to (C.15) is $\{\pm\omega_0 + 2k\pi\}$, $k\in\mathbb{Z}$, and hence, by (C.7), the nonzero eigenvalues of $T_{\widehat{f_{s,\tau}}}$ are (all simple)

$$\pm\frac{a-b}{\omega_0+2k\pi} = \pm\frac{2\tau}{\omega_0+2k\pi}, \qquad k\in\mathbb{Z}. \tag{C.17}$$

Consequently, (3.6) and (2.21) yield

$$n^{-1}U_n(s\cdot 2f_{\mathsf{s}} + \tau\cdot 2f_{\mathsf{a}}) = n^{-1}U_n(f_{s,\tau}) \xrightarrow{\mathrm{d}} W_{s,\tau} \tag{C.18}$$

where, using also the product expansion for cosine [20, 4.22.2],

$$\begin{aligned}
\mathbb{E}\,[e^{\mathrm{i}W_{s,\tau}}] &= \prod_\lambda e^{-\mathrm{i}\lambda/2}(1-\mathrm{i}\lambda)^{-1/2} = \prod_{k=-\infty}^{\infty}\Bigl(\Bigl(1-\frac{2\mathrm{i}\tau}{\omega_0+2\pi k}\Bigr)\Bigl(1+\frac{2\mathrm{i}\tau}{\omega_0+2\pi k}\Bigr)\Bigr)^{-1/2}\\
&= \prod_{k=-\infty}^{\infty}\Bigl|1+\frac{2\mathrm{i}\tau}{\omega_0+2\pi k}\Bigr|^{-1} = \prod_{k=-\infty}^{\infty}\left|\frac{1+\frac{\omega_0+\pi}{(2k-1)\pi}}{1+\frac{\omega_0+\pi+2\mathrm{i}\tau}{(2k-1)\pi}}\right|\\
&= \prod_{k=1}^{\infty}\left|\frac{1-\frac{(\omega_0+\pi)^2}{(2k-1)^2\pi^2}}{1-\frac{(\omega_0+\pi+2\mathrm{i}\tau)^2}{(2k-1)^2\pi^2}}\right| = \Bigl|\frac{\cos(\omega_0/2+\pi/2)}{\cos(\omega_0/2+\pi/2+\mathrm{i}\tau)}\Bigr| = \Bigl|\frac{\sin(\omega_0/2)}{\sin(\omega_0/2+\mathrm{i}\tau)}\Bigr|\\
&= \Bigl|\frac{\sin(\omega_0/2)}{\sin(\omega_0/2)\cosh(\tau)+\mathrm{i}\cos(\omega_0/2)\sinh(\tau)}\Bigr|\\
&= \bigl(\cosh^2(\tau)+\cot^2(\omega_0/2)\sinh^2(\tau)\bigr)^{-1/2}.
\end{aligned} \tag{C.19}$$

Furthermore, by (C.16),

$$\cot^2(\omega_0/2) = \frac{1+\cos(\omega_0)}{1-\cos(\omega_0)} = \frac{(a+b)^2}{(a-b)^2} = \frac{s^2}{\tau^2}. \tag{C.20}$$

It follows from (C.18)–(C.20) that (8.17) and (8.18) hold jointly, with limits W_{s} and W_{a} having the joint characteristic function (8.19). (The cases $s=0$ or $\tau=0$, implicit in (8.18) and (8.17), follow by continuity.)

References

[1] Patrick Billingsley: *Convergence of Probability Measures.* Wiley, New York, 1968. MR 0233396

[2] Denis Chebikin: Variations on descents and inversions in permutations. *Electron. J. Combin.* **15** (2008), no. 1, Research Paper 132, 34 pp. MR 2448882

[3] John B. Conway: *A Course in Functional Analysis*, 2nd ed. Springer-Verlag, New York, 1990. MR 1070713

[4] Harald Cramér: Über eine Eigenschaft der normalen Verteilungsfunktion. *Math. Z.* **41** (1936), no. 1, 405–414. MR 1545629

[5] E. B. Dynkin & A. Mandelbaum: Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* **11** (1983), no. 3, 739–745. MR 0707925

[6] Chaim Even-Zohar: The writhe of permutations and random framed knots. *Random Structures Algorithms* **51** (2017), no. 1, 121–142. MR 3668848

[7] Gavin G. Gregory: Large sample theory for U-statistics and tests of fit. *Ann. Statist.* **5** (1977), no. 1, 110–123.

[8] Allan Gut: *Probability: A Graduate Course*, 2nd ed. Springer, New York, 2013. MR 2977961

[9] Peter Hall: On the invariance principle for U-statistics. *Stochastic Process. Appl.* **9** (1979), no. 2, 163–174. MR 0548836

[10] Wassily Hoeffding: A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19** (1948), 293–325. MR 0026294

[11] Wassily Hoeffding: The strong law of large numbers for U-statistics. Institute of Statistics, Univ. of North Carolina, Mimeograph series 302 (1961). `https://repository.lib.ncsu.edu/handle/1840.4/2128`

[12] Svante Janson: *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, UK, 1997. MR 1474726

[13] Svante Janson: Renewal theory for asymmetric U-statistics. *Electron. J. Probab.* **23** (2018), Paper No. 129, 27 pp. MR 3896866

[14] Svante Janson & Michael J. Wichura: Invariance principles for stochastic area and related stochastic integrals. *Stochastic Process. Appl.* **16** (1984), no. 1, 71–84. MR 0723644

[15] Olav Kallenberg: *Foundations of Modern Probability*, 2nd ed. Springer, New York, 2002. MR 1876169

[16] Peter D. Lax: *Functional Analysis.* Wiley, 2002. MR 1892228

[17] Paul Lévy: Wiener's random function, and other Laplacian random functions. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pp. 171–187, Univ. California Press, Berkeley-Los Angeles, Calif., 1951. MR 0044774

[18] R. G. Miller, Jr. & Pranab Kumar Sen: Weak convergence of U-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.* **43** (1972), 31–41.

[19] Georg Neuhaus: Functional limit theorems for U-statistics in the degenerate case. *J. Multivariate Anal.* **7** (1977), no. 3, 424–439. MR 0455084

[20] *NIST Handbook of Mathematical Functions.* Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert and Charles W. Clark. Cambridge Univ. Press, 2010.
Also available as *NIST Digital Library of Mathematical Functions*, `http://`

dlmf.nist.gov/ MR 2723248
[21] Philip E. Protter: *Stochastic Integration and Differential Equations*. 2nd ed. Springer-Verlag, Berlin, 2004. MR 2020294
[22] H. Rubin & R.A. Vitale: Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** (1980), 165–170. MR 0557561
[23] Angus E. Taylor: *Introduction to Functional Analysis*. John Wiley & Sons, Inc., New York, 1958. MR 0098966
[24] Marc Yor: Remarques sur une formule de Paul Lévy. *Séminaire de Probabilités XIV (Paris, 1978/1979)*, pp. 343–346, Lecture Notes in Math., 784, Springer, Berlin, 1980. MR 0580140

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Email address: svante.janson@math.uu.se
URL: http://www2.math.uu.se/~svante/