Journal of Theoretical Probability, Vol. 13, No. 1, 2000

On the Asymptotic Behavior of Weighted *U*-Statistics[1]

Mohamed Rifi[2] and Frederic Utzet[2,3]

Received June 19, 1998; revised June 10, 1999

A sufficient condition for the convergence of degenerated weighted *U*-statistics of order k is proved. From this, the previous results by O'Neil and Redner[(10)] for $k=2$ and Major,[(7)] which appeared to be very different, are related.

KEY WORDS: *U*-statistics; Ito–Wiener integrals.

1. INTRODUCTION

A weighted *U*-statistic is an expression of the type

$$\mathscr{U}_n = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_n)\, f(X_{i_1},..., X_{i_k})$$

where X_n, $n \geqslant 1$, are independent and identically distributed random variables, $a\colon \mathbb{N}^k \to \mathbb{R}$ is a bounded symmetric function called the *weight* function, and $f\colon \mathbb{R}^k \to \mathbb{R}$ is a measurable and symmetric function such that

$$E[\, f(X_1,..., X_k)] = 0$$

and

$$E[\, f(X_1,..., X_k)^2] < \infty$$

[1] This work was partially supported by grant 1997SGR 144 from the CIRIT and Grant PB96-1182 from DGES.

[2] Departament de Matematiques, Universitat Autonoma de Barcelona, Spain.

[3] To whom all correspondence should be addressed. Departament de Matematiques. Facultat de Ciencies. Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. E-mail: utzet@mat.uab.es.

We will restrict to the called *degenerated* weighted U-statistics, that is, we will assume that f verifies

$$E[f(x_1,..., x_{k-1}, X_k)]=0, \qquad \forall x_1,..., x_{k-1} \in \mathbb{R}$$

(the functions f verifying this last condition are called *canonicals*). The class of weighted U-statistics is important in itself but additionally because it includes the incomplete U-statistics (for a complete survey on weighted and incomplete U-statistics see Lee[6]).

The asymptotic behavior of the weighted U-statistics has been studied by O'Neil and Redner[10] for the case $k=2$ and by Major[7] for any value of k. The origin of this work was an attempt to understand the apparent differences between both papers. The first mentioned authors impose the condition (remember $k=2$) that for all $r \geqslant 2$, there exists the limit

$$\lim_n \frac{1}{n^r} \sum_{i_1 \neq i_2 \neq \cdots \neq i_r} a(i_1, i_2)\, a(i_2, i_3) \cdots a(i_r, i_1)$$

and then they prove that

$$\lim_n \frac{1}{n} \mathscr{U}_n = Y, \qquad \text{in distribution}$$

where the random variable Y is characterized by its generating function. Major[7] defines the functions in $L^2([0, 1]^k)$

$$A_n(y_1,..., y_k)=a([ny_1],..., [ny_k])$$

and assumes that there is a continuous function A: $[0, 1]^k \to \mathbb{R}$ such that

$$\lim_n A_n = A, \qquad \text{in } L^2([0, 1]^k)$$

He then proves (among other results) that $n^{-k/2}\mathscr{U}_n$ converges in law to a limit which is given by a multiple stochastic integral of $A \otimes f$ with respect to a two parameter Wiener process.

A comparison of both papers is not easy: it is not evident that the Major's theorems restricted to $k=2$ give similar conditions to that O'Neil and Redner,[10] and neither is it clear that the limit obtained by O'Neil and Redner may be expressed as a multiple integral. We think that (for $k=2$) the result form O'Neil and Redner is the most general possible. The results from Major are very interesting, first because the techniques used by O'Neil and Redner cannot be used for $k \geqslant 3$, and secondly because the way he expresses the condition for convergence is, in our opinion, the most appropriate for the general case. In this paper we obtain sufficient conditions for the convergence for general k, from which we deduce the theorems of Major. Further, we get a relationship between our condition and the O'Neil and Redner condition which gives a better insight into this topic.

Our main idea is that, since the condition of convergence (in O'Neil and Redner and in Major) does not depend on the kernel f, take $\{X_n, n \geqslant 1\}$ i.i.d., $E[X_1]=0$, $E[X_1^2]=1$, and f as simple as possible, that is,

$$f(x_1,..., x_k) = x_1 \cdots x_k$$

The (weak) invariance principle is used in order to construct $\{Z_n, n \geqslant 1\}$ i.i.d. $\mathcal{N}(0, 1)$ (and also a copy of $\{X_n, n \geqslant 1\}$) such that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - Z_i) \xrightarrow[n \to \infty]{P} 0$$

and prove that

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)(X_{i_1} \cdots X_{i_k} - Z_{i_1} \cdots Z_{i_k})) \xrightarrow[n \to \infty]{P} 0$$

Then we will state our condition as (the vectorial version of) the converge in law of the sequence

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, Z_{i_1} \cdots Z_{i_k}$$

to a certain multiple stochastic integral. Thus we eliminate Major's condition about the continuity of the limit function A, and we change a convergence in $L^2([0, 1]^k)$ to a convergence in law.

Another way to study the problem, and to understand why the limit obtained by Major is a multiple stochastic integral with respect to a two-parameter Wiener process is that the asymptotic behavior of

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, f(X_{i_1},..., X_{i_k})$$

should be very similar to that of the sequence

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} A(i_1/n,..., i_k/n)\, f(X_{i_1},..., X_{i_k})$$

(A is the limit function introduced by Major), and this sequence should behavior as

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} A(U_{i_1},..., U_{i_k})\, f(X_{i_1},..., X_{i_k}), \tag{1.1}$$

where $\{U_n, n \geqslant 1\}$ are i.i.d. uniform in $[0,1]$, independent of $\{X_n, n \geqslant 1\}$, and (1.1) is a degenerated symmetric statistic studied by Dynkin and Mandelbaum,[4] whose limit is, precisely, a multiple stochastic integral with respect to two-parameter Wiener process. We will not follow this approach here.

2. THE MAIN THEOREM

First at all, note that if F is the distribution function of X and U is uniform in $[0, 1]$ then $F^{-1}(U) \stackrel{\mathscr{L}}{=} X$, where F^{-1} is the left continuous inverse of F. It follows that the limit in law of the sequence $n^{-k/2}\mathscr{U}_n$ should be the same as the sequence

$$\begin{aligned} n^{-k/2} &\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, f(F^{-1}(U_{i_1}),..., F^{-1}(U_{i_k})) \\ &= n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)(f \circ (F^{-1})^{\otimes n})(U_{i_1},..., U_{i_k}) \end{aligned}$$

where $\{U_n, n \geqslant 1\}$ are i.i.d. uniform in $[0, 1]$. Therefore, changing the kernel f, we may assume that X_n are uniform in $[0, 1]$.

We need introduce some notations. We will denote by $I_k(h)$ the multiple Wiener-Ito integral of a function $h \in L^2([0, 1]^k)$ with respect to a standard Wiener process, and $J_k(g)$ will be the multiple Wiener-Ito integral with respect to a two-parameter Wiener process of a function $g \in L^2([0, 1]^{2k})$ [see Dynkin and Mandelbaum[4] or Mandelbaum and Taqqu[8]]. Given to functions h and g we will put

$$(h \otimes g)(x, y) = h(x)\, g(y)$$

and

$$h^{\otimes k} = \underbrace{h \otimes \cdots \otimes h}_{k \text{ times}}$$

Further we will write all the vectors in column and given two vectors of the same dimension, $\mathbf{Y}_1$ and $\mathbf{Y}_2$, we will denote the vector obtained by their element by element product (also called Hadamard product) by $\mathbf{Y}_1 \cdot \mathbf{Y}_2$.

To finish, we will need a multidimensional Wiener process with no independent components; following Csörgő and Horváth[(3)] [p. 18], a d-dimensional Wiener process with covariance matrix $\Sigma = (\sigma_{ij})$, is a centered Gaussian random process $\mathbf{W}_\Sigma = \{(W_t^{(1)},..., W_t^{(d)})', t \in \mathbb{R}_+)$ such that

$$E[W_s^{(i)} W_t^{(j)}] = \sigma_{ij} \min(s, t)$$

Given a function $h \in L^2([0, 1]^k)$, we will denote by $\mathbf{I}_k^\Sigma(h)$ the vector $(I_k^{(1)}(h),..., I_k^{(d)}(h))'$, where $I_k^{(i)}(h)$ is the multiple Wiener integral of h with respect the i component of $\mathbf{W}_\Sigma$.

The main theorem is the following.

Theorem 1. Let $\{X_n, n \geqslant 1\}$ be a sequence of i.i.d. uniform in $[0, 1]$ random variables. Consider a function $f: \mathbb{R}^k \to \mathbb{R}$ measurable and symmetric such that

$$E[f(X_1,..., X_k)] = 0, \qquad E[f(X_1,..., X_k)^2] < \infty$$

and

$$E[f(x_1,..., x_{k-1}, X_k)] = 0, \qquad \forall x_1,..., x_{k-1} \in \mathbb{R}$$

and let $a: \mathbb{N}^k \to \mathbb{R}$ be a symmetric bounded function. Assume that there is a symmetric function $A \in L^2([0, 1])^k$ such that for every $d \geqslant 1$ and every sequence $\{\mathbf{Z}_n, n \geqslant 1\}$ of i.i.d. centered d-dimensional Gaussian random vectors with covariance matrix Σ, we have that

$$\lim_n n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k) \mathbf{Z}_{i_1} \cdots \mathbf{Z}_{i_k} = \frac{1}{k!} \mathbf{I}_k^\Sigma(A), \qquad \text{in distribution}$$

Then

$$\lim_n n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k) f(X_{i_1},..., X_{i_k})$$
$$= \frac{1}{k!} J_k(A \otimes f), \qquad \text{in distribution}$$

Remark 1. We will see that when $k = 2$ (Theorem 5) or when $f = h^{\otimes k}$ (Theorem 4) the hypothesis of the Theorem can be strongly weakened.

Remark 2. In the general case, the convergence (in distribution)

$$\lim_n n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, Z_{i_1} \cdots Z_{i_k} = \frac{1}{k!} I_k(A) \tag{2.1}$$

where $\{Z_n, n \geqslant 1\}$ is a sequence of i.i.d. standard normal random variables $\mathcal{N}(0, 1)$ is necessary for the convergence (in distribution)

$$\lim_n n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, f(X_{i_1},..., X_{i_k}) = \frac{1}{k!} J_k(A \otimes f) \tag{2.2}$$

Assume that (2.2) holds for every f verifying the conditions of the theorem. Let Φ the distribution function of a standard normal random variable, and consider

$$f(x_1,..., x_k) = \Phi^{-1}(x_1) \cdots \Phi^{-1}(x_k)$$

Then,

$$\begin{aligned} n^{-k/2} &\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, f(X_{i_1},..., X_{i_k}) \\ &\overset{\mathcal{L}}{=} n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, Z_{i_1} \cdots Z_{i_k} \end{aligned}$$

and, by (2.2), the limit of this sequence is

$$\frac{1}{k!} J_k(A \otimes (\Phi^{-1})^{\otimes k})$$

However,

$$I_k(A) \overset{\mathcal{L}}{=} J_k(A \otimes (\Phi^{-1})^{\otimes k})$$

since

$$\int_0^1 (\Phi^{-1}(x))^2\, dx = E[Z^2] = 1$$

(see Mandelbaum and Taqqu,[8] [A4] or Major,[7] [Lemma 3]). Then, we get (2.1).

Remark 3. It is important to observe that the condition for the convergence relies only on the function a, and not on f. Really, this is the point which allows the proof we are presenting. Note that when $a = \mathbf{1}_C$, $C \subset \mathbb{N}^k$, the weighted U-statistic is an incomplete U-statistic. We will begin the proof with a lemma which deals with this situation.

Lemma 1. Let Y be a symmetric bounded random variable (necessarily $E[Y] = 0$) with $E[Y^2] = 1$. Consider a set $C \subset \mathbb{N}^k$ and write

$$C_n = \{(j_1, ..., j_k) \in C,\ 1 \leqslant j_1 < \cdots < j_k \leqslant n\}$$

There exist (in some probability space) two sequences $\{Y_n,\ n \geqslant 1\}$ i.i.d. with the same law as Y, and $\{Z_n,\ n \geqslant 1\}$ i.i.d. with law $\mathcal{N}(0, 1)$, such that

$$\lim_n n^{-k/2} \sum_{(i_1, ..., i_k) \in C_n} (Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k}) = 0, \qquad \text{in } L^1$$

Proof. The proof is based in the (weak) representation of a sum of i.i.d. square integrable random variables as a sum of i.i.d. $\mathcal{N}(0, 1)$ random variables. Here we shall briefly recall some points of that construction needed in the proof (for the details see for instance Breiman,[2] [Chap. 13]. Let $\{B_t,\ t \geqslant 0\}$ be a standard Brownian motion; there exists a sequence $\{(V_n, V'_n),\ n \geqslant 1\}$ of i.i.d. random variables, independent of the Brownian motion, with $V_n < 0 < V'_n$ a.s., such that if we define

$$T_1 = \inf\{t \geqslant 0 : B_t \notin (V_1, V'_1)\}$$

then $B(T_1)$ has the law of Y. Write

$$Y_1 = B(T_1)$$

Define now

$$T_2 = \inf\{t \geqslant 0 : B(T_1 + t) - B(T_1) \notin (V_1, V'_1)\}$$

and write

$$Y_2 = B(T_1 + T_2) - B(T_1)$$

which has also the same law as Y and is independent of Y_1. And so on. On the other hand, define

$$Z_1 = B(1)$$

$$Z_n = B(n) - B(n-1), \qquad n \geqslant 2$$

Then it it proved that

$$\lim_n n^{-1/2} \sum_{i=1}^{n} (Y_i - Z_i) = 0, \qquad \text{in probability}$$

It is important to remark that the symmetry of Y implies that the variables (V_n, V'_n) verify $V_n = -V'_n$.

We are going to see that

$$E\left[\left|\sum_{(i_1,\dots,i_k)\in C_n} (Y_{i_1}\cdots Y_{i_k} - Z_{i_1}\cdots Z_{i_k})\right|\right] \leqslant E\left[\left|\sum_{1\leqslant i_1<\cdots<i_k\leqslant n} (Y_{i_1}\cdots Y_{i_k} - Z_{i_1}\cdots Z_{i_k})\right|\right]$$

To proof this, write

$$\mathscr{F}(C_n) = \sigma\{Y_{i_1}\cdots Y_{i_k}, Z_{i_1}\cdots Z_{i_k}, (i_1,\dots, i_k)\in C_n\}$$

Then

$$E\left[\sum_{1\leqslant i_1<\cdots<i_k\leqslant n} (Y_{i_1}\cdots Y_{i_k} - Z_{i_1}\cdots Z_{i_k})/\mathscr{F}(C_n)\right] = \sum_{(i_1,\dots,i_k)\in C_n} (Y_{i_1}\cdots Y_{i_k} - Z_{i_1}\cdots Z_{i_k}) + \sum_{(i_1,\dots,i_k)\notin C_n} E[(Y_{i_1}\cdots Y_{i_k} - Z_{i_1}\cdots Z_{i_k})/\mathscr{F}(C_n)]$$

Let $1\leqslant i_1<\cdots<i_k\leqslant n$ such that $(i_1,\dots, i_k)\notin C_n$.

$$E[Y_{i_1}\cdots Y_{i_k}/\mathscr{F}(C_n)] = E[Y_{i_1}\cdots Y_{i_k}/\mathscr{F}(C_n), Y_{i_2}\cdots Y_{i_k}/\mathscr{F}(C_n)] = E[Y_{i_2}\cdots Y_{i_k}E[Y_{i_1}/\mathscr{F}(C_n), Y_{i_2}\cdots Y_{i_k}]/\mathscr{F}(C_n)]$$

Given the symmetry of Y, we can change B_t by $-B_t$ and it follows that the law of the vectors

$$(Y_{j_1},\dots, Y_{j_k}, Z_{j_1},\dots, Z_{j_k}, (j_1,\dots, j_k)\in C_n)$$

and

$$(-Y_{j_1},..., -Y_{j_k}, -Z_{j_1},..., -Z_{j_k}, (j_1,..., j_k) \in C_n)$$

are the same. Hence

$$E[Y_{i_1}/\mathscr{F}(C_n), Y_{i_2} \cdots Y_{i_k}] = E[-Y_{i_1}/\mathscr{F}(C_n), Y_{i_2} \cdots Y_{i_k}]$$

(the σ-field does not change). It follows that

$$E[Y_{i_1}/\mathscr{F}(C_n), Y_{i_2} \cdots Y_{i_k}] = 0$$

and then,

$$E[Y_{i_1} \cdots Y_{i_k}/\mathscr{F}(C_n)] = 0$$

In the same way it is proved that

$$E[Z_{i_1} \cdots Z_{i_k}/\mathscr{F}(C_n)] = 0$$

Thus, we have proved that

$$\begin{aligned}\sum_{(i_1,..., i_k) \in C_n} &(Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k}) \\ &= E\left[\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k})/\mathscr{F}(C_n)\right]\end{aligned}$$

By Jensen inequality,

$$\begin{aligned}E&\left[\left|\sum_{(i_1,..., i_k) \in C_n} (Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k})\right|\right] \\ &\leqslant E\left[\left|\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k})\right|\right]\end{aligned}$$

We will see that term on the right tends to zero. Remember the so-called Newton identities [see Avram and Taqqu,[1] Appendix]:

$$\sum_{1 \leqslant j_1 < \cdots < j_k \leqslant n} x_{j_1} \cdots x_{j_k} = P_k\left(\sum_{j=1}^n x_j, \sum_{j=1}^n x_j^2,..., \sum_{j=1}^n x_j^k\right)$$

where P_k is a polynomial, $P_k(a_1,..., a_k) = 1/k!\, a_1^k + \cdots$

Therefore,

$$n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (Y_{i_1} \cdots Y_{i_k} - Z_{i_1} \cdots Z_{i_k})$$

$$= \frac{1}{k!}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i\right)^k - \frac{1}{k!}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i\right)^k \tag{2.3}$$

$$+ Q\left(\frac{1}{n} \sum_{i=1}^{n} Y_i^2, \frac{1}{n^{3/2}} \sum_{i=1}^{n} Y_i^3, \ldots, \frac{1}{n^{k/2}} \sum_{i=1}^{n} Y_i^k\right) \tag{2.4}$$

$$- Q\left(\frac{1}{n} \sum_{i=1}^{n} Z_i^2, \frac{1}{n^{3/2}} \sum_{i=1}^{n} Z_i^3, \ldots, \frac{1}{n^{k/2}} \sum_{i=1}^{n} Z_i^k\right) \tag{2.5}$$

where Q is a polynomial. By the strong law of large numbers (2.4) and (2.5) tends to $Q(1, 0, \ldots, 0)$.

The term (2.3) can be written as

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i\right)^k - \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i\right)^k$$

$$= \left\{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - Z_i)\right\} R\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i, \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i\right)$$

where R is a polynomial. By construction, the first term of the product on the right converges to 0 in probability. Using the (bivariate) central limit theorem, the second term of the product converges in law to $R(N_1, N_2)$, where (N_1, N_2) is a bidimensional Gaussian vector. It follows that the expression on the left converges to 0 in probability. The last step is to see that all the terms appearing in (2.3) and (2.4) are uniformly integrable. This is proved in Lemma 2. □

Lemma 2. With the same hypothesis as Lemma 1, for all $r \geqslant 1$,

$$\sup_n E\left[\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j\right)^r\right] < \infty$$

Proof. Let

$$S = \sum_{j=1}^{n} Y_j$$

First at all, note that from the symmetry of Y follows the symmetry of S; hence, all the odds moments of S are zero. We will compute the moment of order r of S, say μ_r^S, from the moments of Y, μ_r. Let κ_ℓ^S the cumulant of order ℓ of S, and κ_ℓ the corresponding of Y. We have that

$$\kappa_\ell^S = n\kappa_\ell$$

From the relationships between the moments and the cumulants we have

$$\mu_r = \sum_{\substack{\ell_1,\dots,\ell_j \\ \alpha_1\ell_1+\cdots+\alpha_j\ell_j=r}} b_{1,\dots,j}\kappa_{\ell_1}^{\alpha_1}\cdots\kappa_{\ell_j}^{\alpha_j}$$

Write the same expression for μ_r^S, and change all the κ_ℓ^S by $n\kappa_\ell$ By the symmetry of Y, all the odd cumulant are zero. Hence, a typical term of the expression of μ_r^S will be

$$n^{\alpha_1+\cdots+\alpha_j} b_{1,\dots,j}\kappa_{\ell_1}^{\alpha_1}\cdots\kappa_{\ell_j}^{\alpha_j}$$

with $\ell_1,\dots,\ell_j$ even. Since

$$\alpha_1\ell_1+\cdots+\alpha_j\ell_j=r$$

we deduce

$$\alpha_1+\cdots+\alpha_j \leqslant \alpha_1\frac{\ell_1}{2}+\cdots+\alpha_j\frac{\ell_j}{2} \leqslant \frac{r}{2}$$

And the lemma follows. □

Remark 4. Lemmas 1 and 2 can be extended to the case when the random variable Y is a sum of symmetric bounded random variables:

$$Y = Y^{(1)} + \cdots + Y^{(r)}$$

To proof this, consider a (weak) representation of the sum of the i.i.d. random vectors with law $(Y^{(1)},\dots, Y^{(d)})'$ as is exposed by Csörgő and Horváth,[3] [Thm. 2.1] using a multidimensional Brownian motion with the same covariance structure as the vector $(Y^{(1)},\dots, Y^{(d)})'$. That is, there are independent copies $(Y_j^{(1)},\dots, Y_j^{(d)})'$ of $(Y^{(1)},\dots, Y^{(d)})'$ and i.i.d. Gaussian random vectors $(Z_j^{(1)},\dots, Y_j^{(d)})'$ such that

$$\lim_n \frac{1}{\sqrt{n}} \sum_{j=1}^n ((Y_j^{(1)},\dots, Y_j^{(r)})' - (Z_j^{(1)},\dots, Z_j^{(r)})') = 0, \qquad \text{in probability}$$

Write

$$Y_j = Y_j^{(1)} + \cdots + Y_j^{(r)} \qquad \text{and} \qquad Z_j = Z_j^{(1)} + \cdots + Z_j^{(r)}$$

Then

$$\lim_n \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (Y_j - Z_j) = 0, \qquad \text{in probability}$$

All the steps done in the proof of Lemma 1 can be repeated in this more general setup. The extension of Lemma 2 is trivial.

Proof of the Theorem 1. The proof will be divided into 3 steps.

Step 1. Let $a = \mathbf{1}_C$, with $C \subset \mathbb{N}^k$, and $f(x_1,..., x_k) = \phi(x_1) \cdots \phi(x_k)$ such that $\phi(X)$ is a sum of bounded symmetric random variables, and $E[\phi(X_1)^2] = 1$. We have

$$\begin{aligned} n^{-k/2} & \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{1}_C(i_1,..., i_k)\, \phi(X_{i_1}) \cdots \phi(X_{i_k}) \\ &= n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{1}_C(i_1,..., i_k)(\phi(X_{i_1}) \cdots \phi(X_{i_k}) - Z_{i_1} \cdots Z_{i_k}) \\ &\quad + n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{1}_C(i_1,..., i_k)\, Z_{i_1} \cdots Z_{i_k} \end{aligned}$$

and the first term goes to zero in probability by Lemma 1. By hypothesis, the second term converges (in law) to $1/k!\, I_k(A)$. Since

$$I_k(A) \stackrel{\mathscr{L}}{=} J_k(A \otimes \phi^{\otimes k})$$

(Mandelbaum and Taqqu,[8] [A4] or Major,[7] [Lemma 3]) we obtain the result.

Step 2. Let $a\colon \mathbb{N}^k \to \mathbb{R}$ be a symmetric and bounded function, and f as in Step 1. In the measurable space $(\mathbb{N}^k, \mathscr{P}(\mathbb{N}^k))$, the function a can be approximated by elementary functions. The functions

$$a_\ell(i_1,..., i_k) = 2^{-\ell}[2^\ell a(i_1,..., i_k)]$$

($[x]$ is the integer part of x) converges uniformly to a. Note that a_ℓ takes only a finite number of values because a is bounded. Fix $\varepsilon > 0$; there exists a_ε,

$$a_\varepsilon = \sum_{j=1}^{r} b_j \mathbf{1}_{B_r}, \qquad B_r \subset \mathbb{N}^k$$

such

$$|a_\varepsilon(i_1,..., i_k) - a(i_1,..., i_k)| < \varepsilon, \qquad \forall (i_1,..., i_k)$$

Write

$$S_n = n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, \phi(X_{i_1}) \cdots \phi(X_{i_k})$$

$$T_n = n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, Z_{i_1} \cdots Z_{i_k}$$

and S_n^ε (respectively T_n^ε) as S_n (resp. T_n) changing a by a_ε. Then

$$\begin{aligned} |E[\exp\{itS_n\} - \exp\{itI_k(A)\}]| \leqslant{} & |E[\exp\{itS_n\} - \exp\{itS_n^\varepsilon\}]| && (2.6)\\ & + |E[\exp\{itS_n^\varepsilon\} - \exp\{itT_n^\varepsilon\}]| && (2.7)\\ & + |E[\exp\{itT_n^\varepsilon\} - \exp\{itT_n\}]| && (2.8)\\ & + |E[\exp\{itT_n\} - \exp\{itI_k(A)\}]| && (2.9) \end{aligned}$$

To deal with (2.6), use the inequality $|e^{ix} - 1| < |x|$, and

$$\begin{aligned} & E[(S_n - S_n^\varepsilon)^2] \\ &\quad = E\left[\left(n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (a(i_1,..., i_k) - a_\varepsilon(i_1,..., i_k))\, \phi(X_{i_1}) \cdots \phi(X_{i_k})\right)^2\right] \\ &\quad = E[\phi(X_1)^2]^k \frac{1}{n^k} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} (a(i_1,..., i_k) - a_\varepsilon(i_1,..., i_k))^2 \\ &\quad \leqslant \frac{1}{k!} \varepsilon^2 \end{aligned}$$

In the same way, the term (2.7) is bounded

$$E[(T_n - T_n^\varepsilon)^2] \leqslant \frac{1}{k!} \varepsilon^2$$

Further, for term (2.8)

$$E[|S_n^\varepsilon - T_n^\varepsilon|]$$
$$= n^{-k/2} E\left[\left|\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a_\varepsilon(i_1,..., i_k)(\phi(X_{i_1}) \cdots \phi(X_{i_k}) - Z_{i_1} \cdots Z_{i_k})\right|\right]$$
$$\leqslant \sum_{j=1}^{r} |b_j|\, n^{-k/2}\, E\left[\left|\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{1}_{B_j}(i_1,..., i_k)\right.\right.$$
$$\left.\left.\times (\phi(X_{i_1}) \cdots \phi(X_{i_k}) - Z_{i_1} \cdots Z_{i_k})\right|\right]$$

There is n_0 such that for all $n \geqslant n_0$ and $j = 1,..., s$,

$$n^{-k/2} E\left[\left|\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathbf{1}_{B_j}(i_1,..., i_k)(\phi(X_{i_1}) \cdots \phi(X_{i_k}) - Z_{i_1} \cdots Z_{i_k})\right|\right] < \frac{\varepsilon}{|b_j|}$$

The term (2.9) converges to 0 by hypothesis, independly of ε. Hence, the desired convergence follows.

Step 3. To finish, consider a general canonical f. Let $\{\psi_r, r \geqslant 1\}$ be the Haar orthonormal base of $L^2[0, 1]$; this base is usually numbered using a double index and defined in the following way: for $\ell \geqslant 1$ and r an odd integer between 0 and 2^ℓ,

$$\psi_{\ell, r}(x) = \begin{cases} 2^{(\ell-1)/2}, & x \in \left[\frac{r-1}{2^\ell}, \frac{r}{2^\ell}\right) \\ -2^{(\ell-1)/2}, & x \in \left[\frac{r}{2^\ell}, \frac{r+1}{2^\ell}\right) \\ 0, & \text{otherwise} \end{cases}$$

Observe that $\psi_r(X)$ (X uniform on $[0, 1]$) is a symmetric bounded random variable.

Given $\varepsilon > 0$, there is a function

$$\phi_\varepsilon = \sum_{j=1}^{s} \lambda_j \phi_j^{\otimes k}$$

such that

$$\|f - \phi_\varepsilon\|^2 < \varepsilon$$

where ϕ_j is a finite lineal combination of some of the ψ_i and where $\| \ \|$ denotes the norm in $L^2([0,1]^k)$ [see Giné,[5] p. 70].

Let

$$\sigma_{ij} = E[\phi_i(X_1)\,\phi_j(X_1)]$$

and $\Sigma = (\sigma_{ij})$. By the invariance principle for sums of random vectors and Lemma 1 (and Remarks 2) there is a s-dimensional Wiener process $\{(W_t^{(1)},..., W_t^{(s)}),\ t \geqslant 0\}$ with covariance matrix Σ from which we can construct copies of $(\phi_1(X_i),..., \phi_s(X_i))'$ and gaussian random vectors $(Z_i^{(1)},..., Z_i^{(s)})'$ with covariance matrix Σ such that

$$\begin{aligned}\lim_n n^{k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} & a(i_1,..., i_k)((\phi_1(X_{i_1}) \cdots \phi_1(X_{i_k}), \ \cdots, \phi_s(X_{i_1}) \cdots \phi_s(X_{i_k}))' \\ & - (Z_{i_1}^{(1)} \cdots Z_{i_k}^{(1)},..., Z_{i_1}^{(s)} \cdots Z_{i_k}^{(s)})'), \qquad \text{in } L^1\end{aligned}$$

Write

$$\begin{aligned} L_n &= n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, f(X_{i_1},..., X_{i_k}) \\ L_n^{\varepsilon} &= n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, \phi_{\varepsilon}(X_{i_1},..., X_{i_k}) \\ M_n &= n^{-k/2} \sum_{j=1}^{s} \lambda_j \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)\, Z_{i_1}^{(j)} \cdots Z_{i_k}^{(j)} \end{aligned}$$

As in the step 2, decompose

$$\begin{aligned} |E[\exp\{itL_n\} &- \exp\{itJ_k(A \otimes f)\}]| \\ &\leqslant |E[\exp\{itL_n\} - \exp\{itL_n^{\varepsilon}\}]| \qquad (2.10) \\ &\quad + |E[\exp\{itL_n^{\varepsilon}\} - \exp\{itM_n\}]| \qquad (2.11) \\ &\quad + |E[\exp\{itM_n\} - \exp\{itJ_k(A \otimes \phi_{\varepsilon}\}]| \qquad (2.12) \\ &\quad + |E[\exp\{itJ_k(A \otimes \phi_{\varepsilon}\} - \exp\{itJ_k(A \otimes f)\}]| \qquad (2.13) \end{aligned}$$

The term (2.10) is easily bounded:

$$\begin{aligned} E\left[\left(n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k)(f(X_{i_1},..., X_{i_k})) - \phi_{\varepsilon}(X_{i_1},..., X_{i_k}))\right)^2\right] \\ \leqslant \|f - \phi_{\varepsilon}\|^2\, C < C\varepsilon \end{aligned}$$

where C is the constant that bounds a.

To deal with (2.11), proceed as in term (2.8) of step 2, and use Lemma 1 to find a n_0 such that for any $n \geqslant n_0$,

$$E\left[n^{-k/2} \left| \sum_{j=1}^{r} \lambda_j \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_k) \right.\right.$$
$$\left.\left. \times (\phi_j(X_{i_1}) \cdots \phi_j(X_{i_k}) - Z_{i_1}^{(j)} \cdots Z_{i_k}^{(j)}) \right| \right] < \varepsilon$$

For (2.12), by hypotheses,

$$\lim_n n^{-k/2} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_n)\, Z_{i_1}^{(j)} \cdots Z_{i_k}^{(j)} = \frac{1}{k!} I_k(A), \qquad \text{in distribution}$$

jointly for $j = 1,..., s$. Using a vectorial version of the formula which gives the relationship between I_k and J_k, we have that

$$\mathbf{I}_k^{\Sigma}(A) \stackrel{\mathscr{L}}{=} (J_k(A \otimes \phi_1^{\otimes k}),..., J_k(A \otimes \phi_s^{\otimes k}))'$$

It follows that

$$\lim_n n^{-k/2} \sum_{j=1}^{s} \lambda_j \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} a(i_1,..., i_n)\, Z_{i_1}^{(j)} \cdots Z_{i_k}^{(j)}$$
$$= \frac{1}{k!} \sum_{j=1}^{s} \lambda_j J_k(A \otimes \phi_j^{\otimes k}) = \frac{1}{k!} J_k(A \otimes \phi_\varepsilon)$$

in distribution.

At the end, for (2.13) note that

$$E[(J_k(A \otimes f) - J_k(A \otimes \phi_\varepsilon))^2] = \|A \otimes f - A \otimes \phi_\varepsilon\|^2 < \varepsilon\, \|A\|^2$$

And the proof is finished. □

3. THE THEOREMS OF MAJOR

We are going to prove three theorems from Major.[7] In Major, the three theorems need different proofs, but we will see that all can be deduced from our Theorem 1.

Theorem 2 [Major,[7] Thm. 1]. Let $\{X_n, n \geqslant 1\}$, f and a as in Theorem 1. Define $A_n: [0, 1]^k \to \mathbb{R}$ by

$$A_n(y_1,..., y_k) = a([ny_1],..., [ny_k])$$

and assume that there is a continuous function $A\colon [0, 1]^k \to \mathbb{R}$ such that

$$\lim_n A_n = A, \qquad \text{in } L^2([0, 1]^k)$$

Then

$$\lim_n n^{-k/2} \mathcal{U}_n = J_k(A \otimes f), \qquad \text{in law}$$

Proof. Consider a sequence $\{\mathbf{Z}_n,\ n \geqslant 1\}$ of i.i.d. centered d-dimensional Gaussian random vectors with covariance matrix $\Sigma = (\sigma_{ij})$. First at all, there exist functions $\phi_1,..., \phi_d \in L^2([0, 1])$ such that

$$\int_0^1 \phi_i(t)\, \phi_j(t)\, dt = \sigma_{ij}$$

A very simple probabilistic proof is as follows: Since Σ is positive semi-definite, there is a matrix $P = (p_{ij})$ such that

$$\Sigma = PP'$$

Let $\psi_1,..., \psi_d$ be different elements of any orthonormal base of $L^2([0, 1])$; consider the Gaussian vector $\mathbf{Y} = (I_1(\psi_1),..., I_1(\psi_d))' \sim \mathcal{N}(0, Id)$. Then

$$P\mathbf{Y} \sim \mathcal{N}(0, \Sigma)$$

The ℓ component of $P\mathbf{Y}$ is

$$\sum_{i=1}^d p_{\ell i} I_1(\psi_i) = I_1\left(\sum_{i=1}^d p_{\ell i} \psi_i\right)$$

The functions which we are locking for are $\phi_\ell = \sum_{i=1}^d p_{\ell i} \psi_i$, $\ell = 1,..., d$.

From the hypotheses it follows that for $\ell = 1,..., d$,

$$\lim_n J_k(A_n \otimes \phi_\ell^{\otimes k}) = J_k(A \otimes \phi_\ell^{\otimes k}), \qquad \text{in } L^2(\Omega)$$

But note that

$$A_n(y_1,..., y_k)$$
$$= \sum_{1 \leqslant j_1,..., j_k \leqslant n} a(j_1,..., j_k)\, \mathbf{1}_{[(j_1-1)/n,\, j_1/n) \times \cdots \times [(j_k-1)/n,\, j_k/n)}(y_1,..., y_k)$$

Set

$$K_n = \{(j_1,\dots, j_k),\ 1 \leqslant j_1,\dots, j_k \leqslant n,\ \text{all indices different}\}$$

and separate the earlier sum in two parts:

$$\begin{aligned} A_n(y_1,\dots, y_k) &= \sum_{(j_1,\dots, j_k)\in K_n} a(j_1,\dots, j_k)\, \mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)}(y_1,\dots, y_k) \\ &\quad + \sum_{(j_1,\dots, j_k)\in K_n^c} a(j_1,\dots, j_k)\, \mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)}(y_1,\dots, y_k) \end{aligned}$$

Therefore,

$$\begin{aligned} J_k(A_n \otimes \phi_\ell^{\otimes k}) &= \sum_{(j_1,\dots, j_k)\in K_n} a(j_1,\dots, j_k)\, J_k(\mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)} \otimes \phi_\ell^{\otimes k}) \\ &\quad + J_k\Bigg(\sum_{(j_1,\dots, j_k)\in K_n^c} a(j_1,\dots, j_k)\, \mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)} \otimes \phi_\ell^{\otimes k}\Bigg) \\ &\overset{\mathscr{L}}{=} n^{-k/2} k! \sum_{1\leqslant j_1<\cdots<j_k\leqslant n} a(j_1,\dots, j_k)\, Z_{j_1}^{(\ell)}\cdots Z_{j_k}^{(\ell)} \qquad (3.1) \\ &\quad + J_k\Bigg(\sum_{(j_1,\dots, j_k)\in K_n^c} a(j_1,\dots, j_k)\, \mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)} \otimes \phi_\ell^{\otimes k}\Bigg) \qquad (3.2) \end{aligned}$$

We will prove that the sum in (3.2) converges to zero in $L^2(\Omega)$. The expectation of the square of a typical term in that expression is

$$\begin{aligned} &E\Bigg[\Bigg(J_k\Bigg(\sum_{j_1=j_2\neq j_3\neq\cdots\neq j_k)} a(j_1,\dots, j_k) \\ &\qquad \times \mathbf{1}_{[(j_1-1)/n,\, j_1/n)\times\cdots\times[(j_k-1)/n,\, j_k/n)} \otimes \phi_\ell^{\otimes k}\Bigg)\Bigg)^2\Bigg] \\ &= \sigma_{\ell\ell}^k \frac{1}{n^k} \sum_{j_1=j_2\neq j_3\neq\cdots\neq j_k)} a(j_1,\dots, j_k)^2 \\ &\leqslant \sigma_{\ell\ell}^k \frac{1}{n^k} \binom{n}{k-1} C \xrightarrow[n\to\infty]{} 0 \end{aligned}$$

where C is the constant that bounds a.

It follows that the hypothesis of Major $\lim_n A_n = A$ in $L^2([0,1]^k)$ implies that

$$\lim_n n^{-k/2} k! \sum_{1 \leqslant j_1 < \cdots < j_k \leqslant n} a(j_1,..., j_k)\, \mathbf{Z}_{j_1} \cdots \mathbf{Z}_{j_k} = \mathbf{I}_k^{\Sigma}(A), \qquad \text{in distribution}$$

$\square$

Remark 5. Major,[7] [Thm. 1], does not assume that the weight function a is bounded, but this assumption is implicit in his Theorem 2 and explicit in his Theorem 3.

Theorem 3 [Major,[7] Thm. 2]. Assume that the function a can be written in the form

$$a(j_1 \cdots, j_k) = u(h(j_1),..., h(j_k))$$

where $h: \mathbb{N} \to \{1,..., r\}$ with some integer r, and u is an arbitrary function on $\{1,..., r\}^k$. Assume also that the limit

$$\lim_n \frac{1}{n} \#\{j : j \leqslant n,\ h(j) = s\} = H(s)$$

exists for all $s = 1,..., r$. Then the sequence $n^{-k/2} \mathcal{U}_n$ converges in distribution to the stochastic multiple integral $1/k!\, J(f \otimes A)$, where

$$A(y_1,..., y_k) = u(j_1,..., j_k)$$

if

$$H(j_1) + \cdots + H(j_{s-1}) < y_s \leqslant H(j_1) + \cdots + H(j_s), \qquad s = 1,..., k$$

Proof. As earlier, let $\{\mathbf{Z}_n = (Z_n^{(1)},..., Z_n^{(d)})',\ n \geqslant 1\}$ be a sequence of i.i.d. random vectors $\mathcal{N}(0, \Sigma)$. Then

$$\begin{aligned} n^{-k/2} & \sum_{1 \leqslant j_1 < \cdots < j_k \leqslant n} a(j_1,..., j_k)\, Z_{j_1}^{(\ell)} \cdots Z_{j_k}^{(\ell)} \\ &= \sum_{1 \leqslant \beta_1 \leqslant \cdots \leqslant \beta_k \leqslant r} u(\beta_1,..., \beta_k)\, n^{-k/2} \\ &\quad \times \sum_{\substack{1 \leqslant j_1 < \cdots < j_k \leqslant n \\ h(j_1) = \beta_1,..., h(j_k) = \beta_k}} a(j_1,..., j_k)\, Z_{j_1}^{(\ell)} \cdots Z_{j_k}^{(\ell)} \end{aligned}$$

Consider a term of the sum with $\beta_1,..., \beta_k$ different.

$$u(\beta_1,..., \beta_k)\, n^{-k/2} \sum_{\substack{1 \leqslant j_1 < \cdots < j_k \leqslant n \\ h(j_1)=\beta_1,..., h(j_k)=\beta_k}} a(j_1,..., j_k)\, Z_{j_1}^{(\ell)} \cdots Z_{j_k}^{(\ell)}$$

$$= \frac{1}{k!}\, u(\beta_1,..., \beta_k) \left(\frac{1}{\sqrt{n}} \sum_{\substack{1 \leqslant i_1 \leqslant n \\ h(i_i)=\beta_1}} Z_{i_1}^{(\ell)} \right) \cdots \left(\frac{1}{\sqrt{n}} \sum_{\substack{1 \leqslant i_k \leqslant n \\ h(i_k)=\beta_k}} Z_{i_1}^{\ell} \right)$$

Now consider a term with coincidences in $\beta_1,..., \beta_k$. For example, $\beta_1 = \beta_2$, $\beta_3 = \beta_4$ and $\beta_5,..., \beta_k$ different (there is no loss of generality assuming the coincidences are well ordered). Then

$$u(\beta_1,..., \beta_k)\, n^{-k/2} \sum_{1 \leqslant j_1 < \cdots < j_k \leqslant n} a(j_1,..., j_k)\, Z_{j_1}^{(\ell)} \cdots Z_{j_k}^{(\ell)}$$

$$= \frac{1}{k!}\, u(\beta_1,..., \beta_k) \left(\frac{1}{n} \sum_{\substack{1 \leqslant i_1 < i_2 \leqslant n \\ h(i_i)=h(i_2)=\beta_1}} Z_{i_1}^{(\ell)} Z_{i_2}^{(\ell)} \right) \left(\frac{1}{n} \sum_{\substack{1 \leqslant i_3 < i_4 \leqslant n \\ h(i_3)=h(i_4)=\beta_3}} Z_{i_3}^{(\ell)} Z_{i_4}^{(\ell)} \right)$$

$$\times \left(\frac{1}{\sqrt{n}} \sum_{\substack{1 \leqslant i_{k-1} \leqslant n \\ h(i_{k-1})=\beta_{k_1}}} Z_{i_{k-1}}^{(\ell)} \right) \cdots \left(\frac{1}{\sqrt{n}} \sum_{\substack{1 \leqslant i_k \leqslant n \\ h(i_k)=\beta_k}} Z_{i_k}^{(\ell)} \right)$$

Develop the first two factors using the Newton identities:

$$\frac{1}{n} \sum_{\substack{1 \leqslant i_1 < i_2 \leqslant n \\ h(i_i)=h(i_2)=\beta_1}} Z_{i_1}^{(h)} Z_{i_2}^{(h)}$$

$$= \frac{1}{2} \left(\frac{1}{\sqrt{n}} \sum_{1 \leqslant i \leqslant n \,:\, h(i)=\beta_1} Z_i^{(h)} \right)^2 - \frac{1}{2} \frac{1}{n} \sum_{1 \leqslant i \leqslant n \,:\, h(i)=\beta_1} Z_i^{(h)\,2}$$

Observe that, for each $j = 1,..., r$, the random vectors $1/\sqrt{n} \sum_{i \leqslant n \,:\, h(i)=j} (Z_i^{(1)},..., Z_i^{(d)})'$ are centered Gaussian with covariance matrix

$$\frac{1}{n} (\# \{ i : 1 \leqslant i \leqslant n,\, h(i) = j \})\, \Sigma$$

which, by hypothesis, converges to $H(j)\, \Sigma$. Hence the vectors $1/\sqrt{n} \times \sum_{i \leqslant n \,:\, h(i)=j} \mathbf{Z}_i$ converge in distribution to a normal random vector with covariance matrix $H(j)\, \Sigma$. Since each of these sequences uses different elements of the original sequence $\{\mathbf{Z}_n, n \geqslant 1\}$, it follows that the dr-dimensional vector

$$\begin{pmatrix} \dfrac{1}{\sqrt{n}} \sum\limits_{\substack{1 \leqslant i \leqslant n \\ h(i)=1}} \mathbf{Z}_i \\ \vdots \\ \dfrac{1}{\sqrt{n}} \sum\limits_{\substack{1 \leqslant i \leqslant n \\ h(i)=r}} \mathbf{Z}_i \end{pmatrix}$$

converges in law to a centered multivariate gaussian random vector $(\mathbf{N}'_1, ..., \mathbf{N}'_r)'$, with $\mathbf{N}_i$, $\mathbf{N}_j$ independent random vectors, if $i \neq j$, and with variance matrix

$$\mathrm{Cov}(\mathbf{N}_j) = H(j)\, \Sigma, \qquad j = 1, ..., r$$

We can write

$$(\mathbf{N}_1, ..., \mathbf{N}_r) \stackrel{\mathscr{L}}{=} (\mathbf{I}_1^\Sigma(f_1), ..., \mathbf{I}_1^\Sigma(f_r))$$

where

$$f_j = \mathbf{1}[\, H(1) + \cdots + H(j-1), H(1) + \cdots + H(j))$$

On the other hand,

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i \leqslant n \,:\, h(i)=j} Z_i^{(l)\,2} = H(j), \qquad \text{a.s}$$

and

$$\lim_n n^{-kr/2} \sum_{i \leqslant n \,:\, h(i)=j} Z_i^{(\ell)\,r} = 0, \qquad \text{a.s.,} \quad r = 3, ..., k$$

To finish, using the formula

$$I_s(\mathbf{1}_E) \cdot I_r(\mathbf{1}_F) = I_{r+s}(\mathbf{1}_{E \times F}), \qquad \text{when} \quad E \cap F = \varnothing$$

and reordering conveniently all the limits obtained, we get precisely the hypotheses of Theorem 1. □

Theorem 4 [Major,[7] Thm. 3]. Assume that the weight function is of the form

$$a(j_1,\dots,j_k)=e(i_1)\cdots e(i_k)$$

such that the sequence $\{e(j),\ j\geqslant 1\}$ is bounded and

$$\lim_n \frac{1}{n}\sum_{j=1}^n e(j)^2=E>0$$

exists. Then $n^{-k/2}\mathscr{U}_n$ converges in distribution to $E^{k/2}I^k(f)$.

Proof. This case is the most important, but it also is the easiest because we can use the Newton identities from the very beginning. With the same notations as Theorem 3,

$$n^{-k/2}\sum_{1\leqslant j_1<\cdots<j_k\leqslant n} e(j_1)\cdots e(j_k)\,Z_{j_1}^{(\ell)}\cdots Z_{j_k}^{(\ell)}$$
$$=P_k\left(\frac{1}{\sqrt{n}}\sum_{j=1}^n e(j)\,Z_j^{(\ell)},\frac{1}{n}\sum_{j=1}^n e(j)^2\,Z_j^{(\ell)\,2},\dots,\frac{1}{n^{k/2}}\sum_{j=1}^n e(j)^k\,Z_j^{(\ell)\,k}\right)$$

From the condition

$$\lim_n \frac{1}{n}\sum_{j=1}^n e(j)^2=E>0$$

it follows that

$$\lim_n \frac{1}{\sqrt{n}}\sum_{j=1}^n e(j)\,\mathbf{Z}_j=\mathcal{N}(0,E\Sigma),\qquad \text{in distribution}$$

and

$$\lim_n \frac{1}{n}\sum_{j=1}^n e(j)^2\,Z_j^{(\ell)\,2}=E,\qquad \text{a.s}$$

All the other terms into P_k tend to zero a.s. Let ϕ be a function such that $\int_0^1\phi^2(t)\,dt=1$ then $\mathbf{I}_1^{\Sigma}(\sqrt{E}\,\phi)$ has law $\mathcal{N}(0,E\Sigma)$. Then the limit of

$$n^{-k/2}\sum_{1\leqslant j_1<\cdots<j_k\leqslant n} e(j_1)\cdots e(j_k)\,\mathbf{Z}_{j_1}\cdots\mathbf{Z}_{j_k}^{(\ell)}$$

is

$$(P_k(I_1^{(1)}(\sqrt{E}\,\phi), E, 0,..., 0),..., P_k(I_1^{(d)}(\sqrt{E}\,\phi), E, 0,..., 0))' = = \mathbf{I}_k^{\Sigma}(E^{k/2}\phi^{\otimes k})$$

By Theorem 1, the limit of the U-statistic is

$$J_k(E^{k/2}\phi^{\otimes k} \otimes f) \stackrel{\mathscr{L}}{=} E^{k/2} I_k(f) \qquad \square$$

Remark 6. Note the equivalence between

$$\lim_n \frac{1}{n} \sum_{j=1}^{n} e(j)^2 = E > 0$$

and

$$\lim_n \frac{1}{n} \sum_{j=1}^{n} e(j)\, Z_j = \mathcal{N}(0, E), \qquad \text{in distribution}$$

where $\{Z_n, n \geqslant 1\}$ are i.i.d. standard normal random variables.

4. THE CASE $k = 2$

When $k = 2$ we may use different techniques because the double Ito-Wiener stochastic integral is determinate by its moments and we know the characteristic function. Further, the most accessible structure of $L^2([0, 1]^2)$ can be used to get better versions of Theorem 1; for example, the following one

Theorem 5. Let $\{X_n, n \geqslant 1\}$ be a sequence of i.i.d. uniform in $[0, 1]$ random variables. Consider a function $f: \mathbb{R}^2 \to \mathbb{R}$ measurable and symmetric such that

$$E[f(X_1, X_2)] = 0, \qquad E[f(X_1, X_2)^2] < \infty$$

and

$$E[f(x, X_2)] = 0, \qquad \forall x \in \mathbb{R}$$

and let $a: \mathbb{N}^2 \to \mathbb{R}$ be a symmetric bounded function. Assume that there is a symmetric function $A \in L^2([0, 1])^2$ such that for every sequence $\{Z_n, n \geqslant 1\}$ of i.i.d. standard normal random variables we have that

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j)\, Z_i Z_j = \frac{1}{2} I_2(A), \qquad \text{in distribution}$$

Then

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j) \, f(X_i X_i) = \frac{1}{2} J_2(A \otimes f), \qquad \text{in distribution}$$

O'Neil and Redner[10] solved completely this case proving the following:

Theorem 6 [O'Neil and Redner,[10] Thm. 2.1]. Let $\{X_n, n \geqslant 1\}$ be a sequence of i.i.d. random variables. Consider a function $f: \mathbb{R}^2 \to \mathbb{R}$ which is measurable and symmetric such that

$$E[f(X_1, X_2)] = 0, \qquad E[f(X_1, X_2)^2] < \infty$$

and

$$E[f(X, x)] = 0, \qquad \forall x \in \mathbb{R}$$

and let $a: \mathbb{N}^2 \to \mathbb{R}$ be a symmetric and bounded function. Assume that for all $r \geqslant 2$, there exists the limit

$$\lim_n \frac{1}{n^r} \sum_{i_1 \neq i_2 \neq \cdots \neq i_r} a(i_1, i_2) \, a(i_2, i_3) \cdots a(i_r, i_1) = \omega_r$$

and $\omega_2 > 0$. Then

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j) \, f(X_i, X_j) = Y, \qquad \text{in distribution}$$

where Y is a random variable with generating function

$$\psi(t) = \exp\left\{ \sum_{r=2}^{\infty} \frac{\alpha_r \omega_r}{2r} \, t^r \right\}$$

where

$$\alpha_r = E[f(X_1, X_2) \, f(X_2, X_3) \cdots f(X_r, X_1)]$$

We will prove that Theorem 5 can be deduced from Theorem 6. Assume the hypothesis of Theorem 5:

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j) \, Z_i Z_j = \frac{1}{2} I_2(A), \qquad \text{in distribution} \tag{4.1}$$

In a similar way as in the proof of Theorem 4, write

$$\frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j)\, Z_i Z_j \overset{\mathscr{L}}{=} \frac{1}{2} I_2(B_n)$$

where

$$B_n(x, y) = \sum_{\substack{1 \leqslant i, j \leqslant n \\ i \neq j}} a(i, j)\, \mathbf{1}_{[(i-1)/n,\, i/n) \times [(j-1)/n,\, j/n)}(x, y)$$

and the convergence in (4.1) is equivalent to

$$\lim_n I_2(B_n) = I_2(A), \qquad \text{in distribution} \tag{4.2}$$

Since for all $m \geqslant 1$

$$E[(I_2(B_n))^{2m}] \leqslant \left(\frac{(2m)!}{m!\, 2^m}\right)^2 (E[(I_2(B_n))^2])^m$$

[see for instance Nualart *et al.*[9]], the sequence in (4.2) is uniformly integrable for all powers and it follows that all moments (hence, all cumulants) of $I_2(B_n)$ converge to the corresponding moments (resp. cumulants) of $I_2(A)$. Since the cumulant of order r of $I_2(B_n)$ is given by

$$\int_0^1 \cdots \int_0^1 B_n(x_1, x_2) \cdots B_n(x_r, x_1)\, dx_1 \cdots dx_r$$

$$= \frac{1}{n^r} \sum_{i_1 \neq i_2 \neq \cdots \neq i_r} a(i_1, i_2)\, a(i_2, i_3) \cdots a(i_r, i_1)$$

we get the condition of O'Neil and Redner.[10]

Remark 7. It is an open question if Theorem 5 is, in fact, equivalent to Theorem 6. If we apply Theorem 6 to the sequence $\{Z_n, n \geqslant 1\}$ i.i.d. $\mathcal{N}(0, 1)$ and $f(x, y) = xy$, we get

$$\lim_n \frac{1}{n} \sum_{1 \leqslant i < j \leqslant n} a(i, j)\, Z_i Z_j = Y, \qquad \text{in distribution}$$

where Y has generating function

$$\psi(t) = \exp\left\{\sum_{r=2}^{\infty} \frac{\omega_r}{2r} t^r\right\}$$

The problem is to see that there is a function $A \in L^2([0,1]^2)$ such that

$$Y = \tfrac{1}{2} I_2(A)$$

Observe that if we could find λ_j, $j \geqslant 1$ which would solve the system

$$\omega_2 = \sum_j \lambda_j^2$$

$$\omega_3 = \sum_j \lambda_j^3$$

then we could consider an orthonormal system in $L^2([0,1])$, $\{\phi_j,\ j \geqslant n\}$, and the function

$$A = \sum_j \lambda_j \phi_j^{\otimes 2}$$

would verify $Y \stackrel{\mathscr{L}}{=} 1/2 I_2(A)$. But to the best of our knowledgment, it is no possible to solve that system except for the trivial case

$$\omega_r = c^r, \qquad r \geqslant 2$$

which is the case studied in Theorem 4.

REFERENCES

1. Avram, F., and Taqqu, M. S. (1986). Symmetric polynomials of random variables attracted to an infinitely divisible law. *Probab. Th. Rel. Fields* **71**, 491–500.
2. Breiman, L. (1968). *Probability*, Addison-Wesley Publishing, Reading, Massachusetts.
3. Csörgő, M., and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*, John Wiley, Chichester.
4. Dynkin, B. E., and Mandelbaum, A. (1983). Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* **11**, 739–745.
5. Giné, E., Grimmett, G. R., and Saloff-Coste, L. (1997). *Lectures on Probability Theory and Statistics*, *Lecture Notes in Mathematics*, Vol. 1665. Springer, New York.
6. Lee, A. J. (1990). *U-Statistics*, Marcel Dekker, Inc., New York.
7. Major, P. (1994). Asymptotic distributions for weighted *U*-statistics. *Ann. Probab.* **22**, 1514–1535.

8. Mandelbaum, A., and Taqqu, M. S. (1984). Invariance principle for symmetric statistics. *Ann. Prob.* **12**, 483–496.
9. Nualart, D., Ustunel, A. S., and Zakai, M. (1988). On the moments of a multiple Wiener integrals and the space induced by polynomials of the integrals. *Stoch. Stoch. Rep.* **25**, 233–240.
10. O'Neil, K. A., and Redner, R. A., (1993). Asymptotic distribution of weighted *U*-statistics of degree two. *Ann. Probab.* **21**, 1159–1169.