
Péter Major
Mathematical Institute of the Hungarian Academy of Sciences.

The original problem: Let $\xi_1, \ldots, \xi_n$ be a sequence of i.i.d. random variables with some distribution $\mu$ on a space $(X, \mathcal{X})$, and let us consider a function $f(x_1, \ldots, x_k)$ of $k$ variables on $(X^k, \mathcal{X}^k)$.

Define with their help the $U$-statistic $I_{n,k}(f)$ of (order $k$)

$$I_{n,k}(f) = \frac{1}{k!} \sum_{1 \leq j_s \leq n, s=1, \ldots, k, j_s \neq j_{s'}} f(\xi_{j_1}, \ldots, \xi_{j_k})$$

determined by the kernel function $f(x_1, \ldots, x_k)$ and independent $\mu$-distributed independent random variables $\xi_1, \ldots, \xi_n$. 
Give a good estimate on the probabilities

\[ P \left( n^{-k/2} I_{n,k}(f) > u \right). \]

This problem is closely related to the estimate of the moments

\[ E \left( n^{-k/2} I_{n,k}(f) \right)^{2M} \]

for large integers \( M \).

In the case of \( k = 1 \) it is natural to assume that \( Ef(\xi) = 0 \). The natural multivariate counterpart of the zero expectation is the property holds of degenerate \( U \)-statistics.

**Definition of degenerate \( U \)-statistics.** Take a \( U \)-statistic \( I_{n,k}(f) \) determined by a sequence
of iid. random variables $\xi_1, \ldots, \xi_n$ with distribution $\mu$ and kernel function $f(x_1, \ldots, x_k)$. This $U$-statistic is degenerate, if the identity

$$E(f(\xi_1, \ldots, \xi_k) | \xi_1 = x_1, \ldots, \xi_{j-1} = x_{j-1},$$

$$\xi_{j+1} = x_{j+1}, \ldots, \xi_k = x_k) = 0$$

for all indices $1 \leq j \leq k$

and values $x_s \in X$, $s \in \{1, \ldots, k\} \setminus \{j\}$

holds, or in an equivalent form

$$\int f(x_1, \ldots, x_{j-1}, u, x_{j+1}, \ldots, x_k) \mu(du) = 0$$

for all indices $1 \leq j \leq k$

and values $x_s \in X$, $s \in \{1, \ldots, k\} \setminus \{j\}$.

Degenerate $U$-statistics behave like sums of independent random variables with expectation zero.

Let us also assume that

$$\int f^2(x_1, \ldots, x_k) \mu(dx_1) \ldots \mu(dx_k) \leq \sigma^2$$
with some $\sigma > 0$ (which means that
\[
E \left( n^{-k/2} I_{n,k}(f) \right)^2 \leq \frac{\sigma^2}{k!}
\].

The normalized $U$-statistic $n^{-k/2} I_{n,k}(f)$ has a limit represented by multiple Wiener–Itô integrals as $n \to \infty$. Let us first study the (moment and tail-distribution) behaviour of these random integrals which appear as limits of degenerate $U$-statistics.

A short explanation of Wiener–Itô integrals.

A Wiener–Itô integral of order $k$ is the limit of appropriate polynomials of order $k$ of independent Gaussian random variables (with expectation zero).

They are defined with the help of so-called white noise.
**Definition of the white noise.** Given a measure $\mu$ on a space $(X, \mathcal{X})$, a set of jointly Gaussian random variables indexed by the measurable sets $A \subset X$, $\mu(A) < \infty$, is called a **white noise** with reference measure $\mu$ if

$$E_{\mu_W(A)} \mu_W(B) = \mu(A \cap B) \quad \text{and} \quad E_{\mu_W(A)} = 0$$

for all measurable sets $A, B \subset X$.

This implies that for disjoint sets $A_1, \ldots, A_n$, $\mu_W(A_1), \ldots, \mu_W(A_n)$ are independent Gaussian random variables with expectation zero and variance $\mu(A_1), \ldots \mu(A_n)$. Beside this

$$\mu_W(A_1 \cup \cdots \cup A_n) = \mu_W(A_1) + \cdots + \mu_W(A_n).$$

Let us consider an **elementary** function $f(x_1, \ldots, x_k)$ of the following form.

Let $A_1, \ldots, A_n$ be disjoint subsets of $X$. Put

$$f(x_1, \ldots, x_k) = c(j_1, \ldots, j_k)$$
if
\[ x_1 \in A_{j_1}, \ldots, x_k \in A_{j_k} \quad 1 \leq j_s \leq n, \ 1 \leq s \leq k \]
with \( j_s \neq j_{s'} \) for \( s \neq s' \)

and
\[ f(x_1, \ldots, x_k) = 0 \quad \text{otherwise.} \]

For this elementary function \( f(x_1, \ldots, x_k) \)
\[
\int f(x_1, \ldots, x_k) \mu_W(dx_1) \cdots \mu_W(dx_k) \\
= \sum c(j_1, \ldots, j_k) \mu_W(A_{j_1}) \cdots \mu_W(A_{j_k}).
\]

The Wiener–Itô integral
\[
Z_{\mu,k}(f) = \int f(x_1, \ldots, x_k) \mu_W(dx_1) \cdots \mu_W(dx_k)
\]
for a general function \( f \) such that
\[
\int f^2(x_1, \ldots, x_k) \mu(dx_1) \cdots \mu(dx_k) < \infty
\]
can be defined as a limit of Wiener–Itô integrals \( \int f_n(x_1, \ldots, x_k) \mu_W(dx_1) \cdots \mu_W(dx_k) \) if the function \( f \) is the limit of the elementary functions \( f_n \) in the \( L_2(\mu^k) \)-norm.
The following limit theorem holds for degenerate \( U \)-statistics.

**Limit theorem for degenerate \( U \)-statistics.** Let us consider a sequence \( I_{n,k}(f) \), \( n = k, k+1, \ldots \), of degenerate \( U \)-statistics determined by a sequence of iid. random variables \( \xi_1, \xi_2, \ldots \), with distribution \( \mu \) on a space \((X, \mathcal{X})\) and a kernel function \( f(x_1, \ldots, x_k) \) square integrable with respect to the measure \( \mu \). The normalized degenerate \( U \)-statistics \( n^{-k/2}I_{n,k}(f) \) converge in distribution to the Wiener–Itô integral

\[
\frac{1}{k!}Z_{\mu,k}(f) = \frac{1}{k!} \int f(x_1, \ldots, x_k) \mu_W(dx_1) \ldots \mu_W(dx_k)
\]

of the function \( f \) with respect to a white noise \( \mu_W \) with reference measure \( \mu \) as \( n \to \infty \).

**Theorem. Estimate on the distribution of Wiener–Itô integrals.** Let a white noise \( \mu_W \) be given with reference measure \( \mu \) on a space
\((X, \mathcal{X})\) together with a function \(f(x_1, \ldots, x_k)\) such that
\[
\frac{1}{k!} \int f^2(x_1, \ldots, x_k) \mu(dx_1) \ldots \mu(dx_k) \leq \sigma^2
\]
The Wiener–Itô integral
\[
Z_{\mu, k}(f) = \int f(x_1, \ldots, x_k) \mu_W(dx_1) \ldots \mu_W(dx_k)
\]
satisfies the inequalities
\[
P(\left| Z_{\mu, k}(f) \right| > u) \leq C \exp \left\{ -\frac{1}{2} \left( \frac{u}{\sigma} \right)^{2/k} \right\}
\]
and
\[
E(Z_{\mu, k})^{2M} \leq C \left( \frac{2}{e} \right)^{kM} M^{kM} \sigma^{2M} \quad M = 1, 2, \ldots.
\]
for all \(u > 0\) with some constant \(C = C(k) > 0\) depending only on the multiplicity \(k\) of the integral.

**Example.** Lower bound for the tail distribution of Wiener–Itô integrals. Let \(\mu_W\) be a white noise with a measure \(\mu\) on some space
Take real valued function $f_0(x)$ on the space $(X, \mathcal{X})$ such that $\int f_0(x)^2 \mu(dx) = 1$. Introduce the function

$$f(x_1, \ldots, x_k) = \sigma f_0(x_1) \cdots f_0(x_k)$$

with some number $\sigma > 0$ and its Wiener–Itô integral $Z_{\mu,k}(f)$ with respect to $\mu_W$.

For this example

$$\int f(x_1, \ldots, x_k)^2 \mu(dx_1) \cdots \mu(dx_k) = \sigma^2,$$

and

$$P(|Z_{\mu,k}(f)| > u) \geq \frac{\tilde{C}}{(\frac{u}{\sigma})^{1/k}} \exp \left\{ -\frac{1}{2} \left( \frac{u}{\sigma} \right)^{2/k} \right\}$$

for all $u > 0$ with some constant $\tilde{C} > 0$.

A similar estimate holds for degenerate $U$-statistics.

**Theorem. Estimate on the tail distribution of degenerate $U$-statistics.** Let $\xi_1, \ldots, \xi_n$ be
a sequence of iid. random variables with some measure $\mu$. Let $f(x_1, \ldots, x_k)$ be such a function of $k$ variables for which $I_{n,k}(f)$ is a degenerate $U$-statistic, (with the random variables $\xi_1, \ldots, \xi_n$) and satisfies the inequalities

$$
\|f\|_\infty = \sup_{x_j \in X, 1 \leq j \leq k} |f(x_1, \ldots, x_k)| \leq 1, \quad (1)
$$

$$
\|f\|_2^2 = \int f^2(x_1, \ldots, x_k) \mu(dx_1) \ldots \mu(dx_k) \leq \sigma^2, \quad (2)
$$

with some number $0 < \sigma^2 \leq 1$.

There are some numbers $A = A(k) > 0$ and $B = B(k) > 0$ depending only on the order $k$ of the $U$-statistic $I_{n,k}(f)$ such that the inequalities

$$
P(k!n^{-k/2}|I_{n,k}(f)| > u) \leq A \exp\left\{-\frac{u^2/k}{2\sigma^2/k \left(1 + B \left(un^{-k/2}\sigma^{-(k+1)}\right)^{1/k}\right)}\right\}
$$
hold for all numbers $0 \leq u \leq n^{k/2} \sigma^{k+1}$, and

$$E \left( n^{-k/2} k! I_{n,k}(f) \right)^{2M} \leq A \left( 1 + B^M \left( \frac{M}{n \sigma^2} \right)^{kM} \right) \left( \frac{2}{e} \right)^{kM} (kM)^{kM} \sigma^{2M}. $$

There can be given examples of degenerate $U$-statistics which show that in the case $u > n^{k/2} \sigma^{k+1}$ only much weaker estimate hold.

The previous results yield a good estimate on the high moments and tail distribution of Wiener–Itô integrals and degenerate $U$-statistics if we have a bound on their variance. In an informal interpretation they say that the high moments or tail distribution of a Wiener–Itô integral of order $k$ or of a degenerate $U$-statistics of order $k$ with variance $\sigma^2$ satisfy such an estimate as the moments and tail distribution of $\sigma \eta^k$, where $\eta$ is a standard normal random variable.
These estimates are sharp in the following sense. There are such Wiener–Itô integrals and $U$-statistics whose moments and tail-distribution have similar lower bounds.

**Question.** What kind of improvement is possible with the help of additional information?

a) **Estimation of bilinear forms of independent standard normal random variables.** (Equivalent to the study of Wiener–Itô integrals of order 2.)

\[
Z = \sum_{j=1}^{n} \sum_{k=1}^{n} a(j, k)(\xi_j \xi_k - E\xi_j \xi_k),
\]

with coefficient $a(j, k) = a(k, j)$ for all pairs $(j, k)$, where $\xi_1, \ldots, \xi_n$ are independent, standard normal random variables.

\[
\sigma^2 = EZ^2 = 2 \sum_{j=1}^{n} \sum_{k=1}^{n} a^2(j, k).
\]
By means of diagonalization of the matrix \((a(j, k))\), \(1 \leq j, k \leq n\),

\[
Z = \sum_{j=1}^{n} \lambda_j (\eta_j^2 - E\eta_j^2),
\]

where \(\eta_1, \ldots, \eta_n\) are independent, standard normal random variables, \(\lambda_1, \ldots, \lambda_n\) are the eigenvalues of the matrix \((a(j, k))\).

\[
\sigma^2 = 2 \sum_{j=1}^{n} \lambda_j^2
\]

**Heuristic argument.** If all eigenvalues \(\lambda_j\) are small, (much smaller than \(\sigma = \sqrt{EZ^2}\)), then the central limit theorem suggests that \(Z\) is asymptotically normal with expectation zero and variance \(\sigma^2\). Hence for not too large \(M\)
\[ EZ^2 M \leq \sigma^2 M E\eta^2 M \leq C^M \sigma^2 M M^M \] with a standard normal random variable \( \eta \) and universal constant \( C \).

If \( M \) is large, then we expect (heuristically) that \( EZ^2 M \) has the magnitude

\[
\max_{1 \leq j \leq n} E[\lambda_j (\eta_j^2 - E\eta_j^2)]^2 M
\]

or the magnitude \( D^2 M E\eta^4 M \leq C^M D^2 M M^2 M \), where

\[
D = \max_{1 \leq j \leq n} |\lambda_j|.
\]

This heuristic argument suggests that

\[ EZ^2 M \leq C^M \sigma^2 M M^M \] if \( M \) is not too large,

and

\[ EZ^2 M \leq C^M D^2 M M^2 M \], if \( M \) is very large.
**Theorem.** Estimate about bilinear forms of independent Gaussian random variables. The above defined random Gaussian bilinear form

\[
Z = \sum_{j=1}^{n} \sum_{k=1}^{n} a(j, k) (\xi_j \xi_k - E\xi_j \xi_k)
\]

satisfies the inequality

\[
C^{-M} \max \left( \sigma^{2M} M^M, D^{2M} M^{2M} \right) \leq EZ^{2M} \leq C^M \max \left( \sigma^{2M} M^M, D^{2M} M^{2M} \right)
\]

for all \( M = 1, 2, \ldots \) with a universal constant \( C > 1 \) and the above defined constants \( \sigma^2 \) and \( D \).

As a consequence,

\[
P(Z > x) \leq C_1 \exp \left\{ -\min \left( \frac{Cx}{D}, \frac{Cx^2}{\sigma^2} \right) \right\}.
\]
The above result has the following counterpart for $U$-statistics of order 2.

**Theorem about the tail distribution of $U$-statistics of order 2.** Let a sequence $\xi_1, \ldots, \xi_n$ of independent $\mu$ distributed random variables be given together with a function $f(x, y)$ canonical with respect to the measure $\mu$, and consider the (degenerate) $U$-statistic $I_{n,2}(f)$ defined with the help of the above quantities. Let us assume that the function $f$ satisfies conditions

$$\int f(x, y)^2 \mu(dx) \mu(dy) \leq \sigma^2$$

and

$$\int f(x, y) g_1(x) g_2(y) \mu(dx) \mu(dy) \leq D$$

if $\int g_1^2(x) \mu(dx) \leq 1$, $\int g_2^2(y) \mu(dy) \leq 1$ with some $\sigma > 0$ and $D > 0$, and also the relations

$$\sup_x \int f^2(x, y) \mu(dy) \leq A,$$

$$\sup_y \int f^2(x, y) \mu(dx) \leq A,$$
and
\[ \sup_{x,y} |f(x, y)| \leq B \]
hold with some appropriate constants \( A > 0 \) and \( B > 0 \).

Then there exists a universal constant \( K > 0 \) such that the inequality
\[
P\left( n^{-1} |I_{n,2}(f)| > u \right) \\
\leq K \exp \left\{ -\frac{1}{K} \left( \frac{u^2}{\sigma^2}, \frac{u}{D}, \frac{n^{1/3}u^{2/3}}{A^{1/3}}, \frac{n^{1/2}u^{1/2}}{B^{1/2}} \right) \right\}
\]
is valid for all \( u > 0 \).

How to generalize the above result for general Gaussian polynomials (Wiener–Itô integrals) and degenerate \( U \)-statistics of order \( k \)?

Reformulation of the definition of \( \sigma^2 \) and \( D \) by means of a variational principle
\[
\sigma = \sum_{u(j,k): \sum u^2(j,k) \leq 1} a(j,k) u(j,k)
\]
\[ D = \sum_{(u(j),v(k)) : \sum u(j)^2 \leq 1, \sum v(k)^2 \leq 1} a(j, k) u(j) v(k) \]

or its \((U\text{-statistic or Wiener–Itô integral})\) version

\[ \sigma = \int f(x, y) u(x, y) \mu(dx) \mu(dy) \]

\[ u(x, y) : \int u^2(x, y) \mu(dx) \mu(dy) \leq 1 \]

\[ D = \int f(x, y) u(x) v(y) \mu(dx) \mu(dy) \]

\[ (u(x), v(y)) : \int u(x)^2 \mu(dx) \leq 1, \int v(y)^2 \mu(dy) \leq 1 \]

Let us define the multivariate version of the above quantities.

Put \( K = \{1, \ldots, k\} \), and let \( \mathcal{P} = \mathcal{P}(K) \) denote the set of all partitions of \( K \). For a partition \( P = \{A_1, \ldots, A_s\} \in \mathcal{P}(K) \) let \( \mathcal{F}_P \) consist of the following sequences of functions.

\[ \mathcal{F}_P = \{ g_r(x_j, j \in A_r), \ 1 \leq r \leq s : \int g_r^2(x_j, j \in A_r) \prod_{j \in A_r} \mu(dx_j) \leq 1 \text{ for all } 1 \leq r \leq s \} \]
Given the set $K = \{1, \ldots, k\}$ and a function $f(x_1, \ldots, x_k)$ a define for all partitions $P = \{A_1, \ldots, A_s\} \in \mathcal{P}(K)$ the quantity

$$V_P(f) = \sup_{(g_1, \ldots, g_s) \in \mathcal{F}_P} \int f(x_1, \ldots, x_k) \prod_{1 \leq r \leq s} g_r(x_j, j \in A_r) \prod_{j=1}^{k} \mu(dx_j).$$

Put $\mathcal{P}_s = \mathcal{P}_s(K)$ the set of partitions with $s$ elements and

$$V_s(f) = \sup_{P \in \mathcal{P}_s} V_P(f).$$

**Theorem about moment estimates of Wiener–Itô integrals.** Let a $k$-fold Wiener–Itô integral

$$k! I_k(f) = \int f(x_1, \ldots, x_k) \mu_W(dx_1) \ldots \mu_W(dx_k)$$
$k \geq 2$ be given with respect to a white noise
$\mu_W$ with reference measure $\mu$. Let its kernel
function $f$ satisfy the following conditions with
a real number $R$, $0 \leq R \leq 1$:

$$V_s(f) \leq R^{s-1} \quad \text{for all } 1 \leq s \leq k,$$

and

$$R \geq M^{-(k-1)/2}.$$

with some positive integer $M$. Then the in-
equality

$$E(k! I_k(f))^{2M} \leq C^M M^{kM} R^{2M}$$

holds with this number $M$ and some universal
constant $C$ depending only on the multiplicity
$k$ of the Wiener–Itô integral.

$$V_1(f)^2 = \int f^2(x_1, \ldots, x_k) \mu(dx_1) \cdots \mu(dx_k)$$

and the condition $V_1(f) \leq 1$ means that

$$E(I_k(f))^2 \leq \frac{1}{k!}.$$ Under this condition

$$E(k! I_k(f))^{2M} \leq C^M M^{kM}.$$
The additional conditions provide an additional factor $R^{2M}$ in the upper bound.

Latała claims a stronger estimate.

**Latała’s claim** about estimates on the moments of Wiener–Itô integrals. Let the above quantities $V_s(f)$ of the Wiener–Itô integral $I_k(f)$ satisfy the inequality

$$V_s(f) \leq M^{-(s-1)} \quad \text{for all } 1 \leq s \leq k.$$

Then there exists some universal constant $C$ depending only on the order $k$ of the Wiener–Itô integral such that

$$EI_k^{2M}(f) \leq C^M M^M.$$

The moment and tail distribution estimates can be reduced to this statement. A comparison of the different estimates:
Under the conditions of Latała’s statement the old estimate yields $E I_k^{2M}(f) \leq C^M M^{kM}$, my estimate $E I_k^{2M}(f) \leq C^M M^{(k-1)M}$, and Latała’s estimate $E I_k(f)^{2M} \leq C^M M^M$.

What is the explanation of the improvement of the moment estimates?

My explanation: The moments can be calculated by the so-called diagram formula, and the terms in this formula can be better bounded under the additional conditions.

Diagram formula.

Let us calculate $E(k! I_k(f))^{2M}$ with some kernel function $f(x_1, \ldots, x_k)$ and positive integer $M$. For this goal define diagrams of the following form. They have vertices
and edges such that from each vertex there starts exactly one edge, and edges connect vertices from different rows.

Let \( \Gamma \) denote the set of all diagrams, and define for all \( \gamma \in \Gamma \) the value \( h_\gamma(f) \) in the following way:

Enumerate the edges \( \{(j, l), (j', l')\} \) of a diagram \( \gamma \) by the numbers \( 1 \leq s \leq kM \), and put \( a_\gamma(j, l) = s \) if \( (j, l) \) is a vertex of the edge with label \( s \).
Put

\[ F(x_{(j,l)}, 1 \leq j \leq 2M, 1 \leq l \leq k) = \prod_{j=1}^{2M} f(x_{(j,1)}, \ldots, x_{(j,k)}) \]

and

\[ h_\gamma(f) = \int F(x_{a_\gamma(j,l)}, 1 \leq j \leq 2M, 1 \leq l \leq k) \prod_{s=1}^{kM} \mu(dx_s). \]

Then

\[ E(k!I_k(f))^{2M} = \sum_{\gamma \in \Gamma} h_\gamma(f). \]

We need good estimates on \( h_\gamma(f) \) for the diagrams \( \gamma \in \Gamma \).

Previous estimate:

\[ h_\gamma(f) \leq 1 \quad \text{if } V_1(f) \leq 1. \]
New estimate:

\[ h_\gamma(f) \leq R^{2M-2} \]

under the conditions of the Theorem about moment estimates of Wiener–Itô integrals for connected diagrams \( \gamma \) (with \( 2M \) rows).

It enables us to estimate the moments of \( I_k(f) \) via its semi-invariants.

The proof of this estimate is based on the following

**Main Inequality.** Let

\[ f(x_1, \ldots, x_m, v_{m+n+1}, \ldots, v_{m+n+q}) \]

and

\[ g(x_{m+1}, \ldots, x_{m+n}, v_{m+n+1}, \ldots, v_{m+n+q}) \]
be two square integrable functions with \( m + q \) and \( n + q \) variables on a measure space \((X, \mathcal{X}, \mu)\), and define the function

\[
F(x_1, \ldots, x_{m+n}) = \int f(x_1, \ldots, x_m, v_{m+n+1}, \ldots, v_{m+n+q}) \cdot g(x_{m+1}, \ldots, x_{m+n}, v_{m+n+1}, \ldots, v_{m+n+q}) \prod_{u=m+n+1}^{m+n+q} \mu(\,dv_u) .
\]

Let \( m + n \geq 1, q \geq 1 \), and let the functions \( f \) and \( g \) satisfy the relations \( V_s(f) \leq D_1 R^{s-1} \) and \( V_s(g) \leq D_2 R^{s-1} \) with some \( D_1 > 0, D_2 > 0 \) and \( 0 \leq R \leq 1 \) for all \( 1 \leq s \leq m+q \) and \( 1 \leq s \leq n+q \) respectively. Then the function \( F \) satisfies the inequality

\[
V_s(F(x_1, \ldots, x_{m+n})) \leq D_1 D_2 R \cdot R^{s-1}
\]

for all \( 1 \leq s \leq m + n \).

In the application of this estimate \( R \) plays the role of a small parameter. Its iteration gives
a good bound on the $L_2$-norm $V_1(\cdot)$. The estimation of $h_\gamma(f)$ by means of the Main inequality exploits the connectedness of $\gamma$ in the condition $q \geq 1$.

Latała’s approach.

He deals with polynomials of independent Gaussian random variables, and exploits that an estimate of de la Peña and Montgomery–Smith enables to reduce the problem to so-called decoupled $U$-statistics, where the different variables take values of independent random variables. This provides more independence, and enables us to calculate the expectation of the high power of polynomials we are interested in by calculating first their conditional expectation with respect to the values of the first coordinate.
Formulation of Latała’s statement in the case $k = 3$.

$$Z = \sum_{i,j,k} a(i, j, k) \xi_i \eta_j \zeta_k,$$

where all random variables $\xi_i$, $\eta_j$ and $\zeta_k$ have standard normal distribution, and they are independent of each other. Let the coefficients $a(i, j, k)$ of the polynomial $Z$ in formula (5.1) satisfy the following inequalities depending on the fixed parameter $M$:

$$\sum_{i,j,k} a(i, j, k) u(i, j, k) \leq 1 \quad \text{if} \quad \sum_{i,j,k} u^2(i, j, k) \leq 1,$$

$$\sum_{i,j,k} a(i, j, k) u(i, j) v(k) \leq M^{-1/2}$$

if $\sum_{i,j} u^2(i, j) \leq 1$ and $\sum_{k} v^2(k) \leq 1$,

$$\sum_{i,j,k} a(i, j, k) u(i, k) v(j) \leq M^{-1/2}$$
if \( \sum_{i,k} u^2(i,k) \leq 1 \) and \( \sum_j v^2(j) \leq 1 \),

\[
\sum_{i,j,k} a(i,j,k)u(j,k)v(i) \leq M^{-1/2}
\]

if \( \sum_{j,k} u^2(j,k) \leq 1 \) and \( \sum_i v^2(i) \leq 1 \), and

\[
\sum_{i,j,k} a(i,j,k)u(i)v(j)w(k) \leq M^{-1}
\]

if \( \sum_i u^2(i) \leq 1 \), \( \sum_j v^2(j) \leq 1 \) and \( \sum_k w^2(k) \leq 1 \).

Then the random polynomial \( Z \) satisfies the inequality

\[
EZ^{2M} \leq C^MM^M
\]

with some universal constant \( C > 0 \).

The idea of proof:

\[
E(Z^{2M} | \xi_i = x_i) = E \left( \sum_{i,j,k} a(i,j,k)x_i\eta_j\xi_k \right)^{2M}
\]
\[
E \left( \sum_{j,k} A(j, k|x) \eta_j \zeta_k \right)^{2M},
\]

where

\[ A_i(j, k|x) = A_i(j, k|x_1, x_2, \ldots) = \sum_i a(i, j, k)x_i. \]

and

\[ EZ^{2M} = E(EZ^{2M}|\xi_i = x_i). \]

The expression \( E(Z^{2M}|\xi_i = x_i) \) can be well bounded, because it is a bilinear form of independent Gaussian random variables. These estimates (together with the concentration inequalities of Ledoux and Talagrand) and the variational principle representation of the Hilbert–Schmidt norm and norm of the matrix \( (A_{j,k}|x) \) leads to some inequalities which are simple to check, and an inequality whose investigation is hard.
The hard problem:

Put

\[ Y(v, w) = \sum_{i,j,k} a(i, j, k)v(j)w(k)\xi_i \]

for all \((v(1), v(2), \ldots)\) and \((w(1), w(2), \ldots)\) such that \(\sum v(j)^2 \leq 1\) and \(\sum w^2(k) \leq 1\). Show that

\[
E \left| \sup_{v(j), w(k)} \sum_{\sum v^2(j) \leq 1, \sum w^2(k) \leq 1} Y(v, w) \right| \leq \frac{C}{\sqrt{M}}.
\]

Introduce the metric

\[
\rho((v, w), (\bar{v}, \bar{w})) = \left[ E(Y(v, w) - Y(\bar{v}, \bar{w}))^2 \right]^{1/2}
\]

in the parameter space \((v, w)\), for all \(v = (v(1), v(2), \ldots)\) and \(w = (w(1), w(2), \ldots)\) such that \(\sum v(j)^2 \leq 1\), \(\sum w^2(k) \leq 1\). A good estimate on the number \(N(\varepsilon)\) of elements of an optimal \(\varepsilon\)-dense set in this metric space would be useful.

But how to get such a bound?